

IoT 数据分析

用户指南

文档版本 01
发布日期 2023-02-28



版权所有 © 华为技术有限公司 2023。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暗示的声明或保证。

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 从这里开始	1
2 服务控制台总览	2
3 数据源配置	4
3.1 数据源配置概述	4
3.2 数据源--添加 IoTDA 实时数据源	4
3.3 数据源--添加 DIS 数据源	5
3.4 数据源--添加 API 数据源	7
4 数据管道	9
4.1 创建数据管道	9
4.1.1 使用 IoTDA 数据源	9
4.1.2 使用 DIS 数据源	13
4.1.3 其他数据源的使用	14
4.2 编辑处理流程	14
4.2.1 数据源配置	14
4.2.2 数据清洗配置	16
4.2.3 数据输出配置	23
4.3 调试数据管道	25
4.3.1 调试概述	25
4.3.2 界面介绍	27
4.3.3 操作指导	29
4.3.4 数据模板	33
4.4 运行数据管道	36
4.4.1 运行数据管道	36
4.4.2 运行过程监控与数据跟踪（敬请期待）	37
4.5 算子使用指南	37
4.5.1 输入类算子	37
4.5.1.1 IoT 设备接入数据源	37
4.5.1.2 DIS 数据源	39
4.5.1.3 API 数据源	41
4.5.2 输出类算子	43
4.5.2.1 数据存储输出	44
4.5.2.2 DIS 输出	46

4.5.2.3 实时分析输出.....	47
4.5.3 转换类算子.....	48
4.5.3.1 产品过滤.....	48
4.5.3.2 数据计算.....	49
4.5.3.3 相邻消息计算.....	51
4.5.3.4 数据扁平.....	54
4.5.3.5 删除属性.....	55
4.5.3.6 选择属性.....	56
4.5.3.7 数据过滤.....	57
4.5.3.8 超时插补.....	58
4.5.3.9 数据聚合.....	60
4.5.4 表达式说明.....	63
5 存储管理.....	66
5.1 存储.....	66
5.1.1 创建存储.....	66
5.1.2 删除存储.....	66
5.2 存储组.....	67
5.2.1 修改数据保存时间.....	67
6 资产建模.....	69
6.1 简介.....	69
6.1.1 基本概念介绍.....	69
6.1.2 主要功能简介.....	71
6.2 模型管理.....	72
6.2.1 创建资产模型.....	72
6.2.2 删除资产模型.....	73
6.2.3 添加属性信息.....	73
6.2.4 修改属性信息.....	75
6.2.5 删除属性信息.....	76
6.2.6 添加分析任务.....	77
6.2.7 修改分析任务.....	77
6.2.8 删除分析任务.....	79
6.3 资产管理.....	79
6.3.1 创建资产.....	80
6.3.2 删除资产.....	80
6.3.3 修改名称.....	81
6.3.4 工作台.....	82
6.3.4.1 创建子资产.....	82
6.3.4.2 移动子资产.....	83
6.3.4.3 删除子资产.....	83
6.3.4.4 配置设备 ID.....	84
6.3.4.5 配置静态值.....	84
6.3.4.6 配置分析任务输入.....	85

6.3.4.7 资产状态.....	86
6.3.5 监控.....	87
6.3.6 日志.....	88
6.3.7 导入资产.....	89
6.3.8 导出资产.....	90
6.4 分析任务定义.....	90
6.4.1 简介.....	90
6.4.2 分析任务类型.....	91
6.4.3 输入参数.....	91
6.4.4 计算公式.....	93
6.4.4.1 简介.....	93
6.4.4.2 字面量.....	93
6.4.4.3 操作符.....	94
6.4.4.4 函数.....	95
6.4.4.5 其他.....	98
6.4.5 分析任务输出.....	98
7 实时分析.....	100
7.1 作业开发.....	100
7.1.1 作业开发概述.....	100
7.1.2 新建作业.....	100
7.1.3 编排作业.....	102
7.1.4 保存作业.....	109
7.2 作业调试.....	109
7.2.1 调试概述.....	109
7.2.2 界面介绍.....	111
7.2.3 操作指导.....	113
7.2.4 数据模板.....	118
7.3 作业运行.....	121
7.3.1 作业运行概述.....	121
7.3.2 启动作业.....	122
7.3.3 查询作业状态.....	123
7.3.4 查看作业操作记录.....	123
7.3.5 停止作业.....	124
8 时序分析.....	125
8.1 时序分析简介.....	125
8.1.1 概念简介.....	125
8.1.2 功能简介.....	125
8.2 资产时序探索.....	126
8.2.1 选择根资产.....	126
8.2.2 选择资产.....	127
8.2.3 选择属性.....	128
8.3 设置时间窗.....	129

8.4 时序洞察.....	130
8.4.1 预览列表展示.....	130
8.4.2 原始数据查看.....	131
8.4.3 图表探索分析.....	132
8.5 收藏夹.....	136
8.6 自动刷新.....	137
9 审计.....	138
9.1 支持云审计的关键操作.....	138
9.2 如何查看审计日志.....	140

1 从这里开始

开通服务

步骤1 用户首次进入IoT数据分析服务控制台后，将会自动开通IoT数据分析服务。

图 1-1 总览页面



步骤2 基础配置

为了能够开始正常使用IoT数据分析服务，需要用户完成如下基础配置：

1. 配置数据源：通过配置数据源，将待分析的数据注入IoT数据分析服务。用户可以根据自己的情况选择不同类型的数据源配置。具体内容可参考[数据源配置](#)
2. 配置数据管道：通过配置数据管道，用户可指定数据存储的位置，如果需要，也可以在管道中实现对数据的实时清洗。具体内容可参考[数据管道](#)
3. 配置存储：通过配置存储，用户可指定存储管理的策略，比如设置温存储各自的老化周期。具体内容可参考[存储管理](#)

完成上述3步后，最基础的配置就完成，IoT数据分析服务已经准备好接收用户的数据了。

用户后续可根据数据分析的需要，再逐步配置更多的内容，比如资产模型或者实时流作业等。

----结束

2 服务控制台总览

数据分析服务控制台总览页为您提供数据分析流程介绍以及业务资源使用情况总览。

可以通过资产模型使用量、数据存储量和计算单元使用量，监控当前业务资源的使用情况。

流程介绍

图 2-1 流程介绍



步骤1 配置数据源

数据分析服务会将数据源的数据存储并进行后续分析。点击图标即可进入数据源配置界面。

步骤2 创建数据管道

数据管道是数据分析的基础功能，对进入数据分析的数据进行格式标准化、清洗等数据预处理。点击图标即可进入数据管道配置界面。

步骤3 配置存储

数据的存储是数据分析的基础，点击图标即可进入存储配置界面。

步骤4 资产建模

根据您的业务场景创建相应的数字资产模型，并基于该模型进行后续的数据分析，点击图标即可进入资产建模界面。

步骤5 数据分析

可以结合业务数据进行针对性的业务洞察，发现业务价值。点击图标即可进入分析界面。

步骤6 结果导出

您可以根据需要导出所需的原始以及分析后的数据。

----结束

资源使用量展示

此处展示您在IoT数据分析服务的资源使用情况，包括资产模型资源、存储资源和计算资源，可直观判断数据分析使用情况。

模型资源使用量：从服务开通到目前的资产模型资源的累计使用情况。

数据存储量：从服务开通到目前的数据类型的累计存储总量。

计算单元使用量：从服务开通到目前的计算资源的累计使用情况。



3 数据源配置

3.1 数据源配置概述

IoT数据分析服务支持多种数据源类型，用户可选择合适的数据源并进行配置：

- IoTDA实时数据源：将IoT设备接入服务（IoTDA）作为数据源，IoT设备接入服务将转发设备上报的数据到IoT数据分析服务，之后您可以通过IoT数据分析服务分析设备上报数据。
- DIS数据源：将华为云DIS服务作为数据源，您可以通过将IoT数据实时推送到DIS服务，然后IoT数据分析服务会从DIS中获取该数据。
- API数据源：IoT数据分析服务提供API接口，您可以通过调用API直接向IoT数据分析服务发送数据。

3.2 数据源--添加 IoTDA 实时数据源

前提条件

您已开通IoT设备接入服务

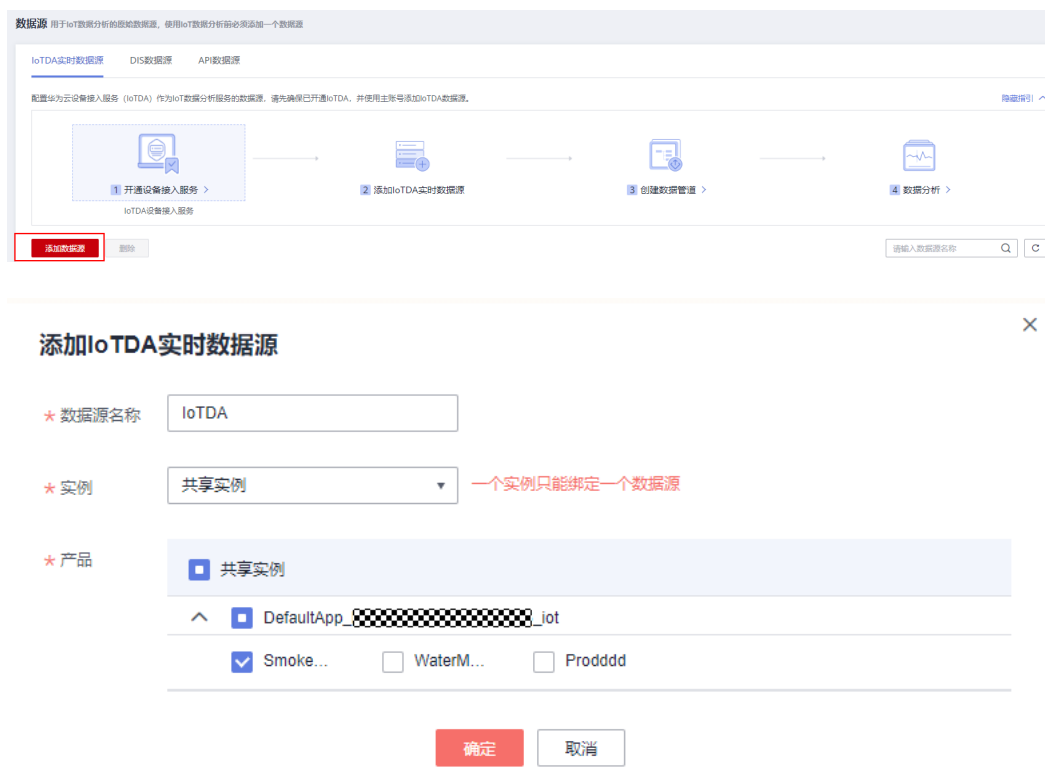
操作步骤

步骤1 使用华为云租户主账号登录进入“IoT数据分析”服务控制台。

步骤2 在左侧导航栏选择“数据源>IoTDA实时数据源”，点击页面中“立即授权”按钮进行确认授权。以便后续IoT数据分析服务与IoTDA服务能够互相访问租户数据。



步骤3 确认授权后，点击“添加数据源”，输入数据源名称、在实例下拉列表中选择1个您名下的IoT设备接入服务实例（IoTDA实例）、并在产品栏选择您需要把该IoTDA实例下哪些产品的数据上报到数据分析服务中，然后点击“确定”。



----结束

⚠ 注意

1个IoT设备接入服务实例只能对应于1个IoTDA实时数据源，即不允许2个IoTDA实时数据源选择同一个IoT设备接入服务实例。

IoTDA实时数据源的“立即授权”、以及添加数据源功能由于需要访问IAM服务，因此必须使用华为云租户主账号进行操作。

3.3 数据源--添加 DIS 数据源

前提条件

您已购买华为云DIS服务。

操作步骤

步骤1 点击IoT数据分析服务控制台左侧导航栏的数据源，进入数据源页面，点击“实时数据源”。



步骤2 实时点击“添加数据源”，在弹出的页面中输入消息通知服务的连接参数，各参数说明详见表1。



表 3-1 添加实时数据源参数说明

参数名	说明	取值样例
数据源名称	数据源的名称	DISNotify
项目ID	项目ID	根据实际值填写
通道名称	DIS通道	根据实际值填写
访问密钥(AK)	访问密钥的KEY	根据实际值填写

参数名	说明	取值样例
私有访问密钥 (SK)	访问密钥的密码	根据实际值填写

步骤3 点击“连接测试”按钮，测试数据源是否可访问。如果测试失败，请确认项目ID、访问密钥、私有访问密钥是否正确，修改后请重新测试。

步骤4 点击“确定”按钮，完成数据接入服务数据源的添加。

----结束

📖 说明

当前实时数据源仅支持对接华为云DIS服务，添加前请您先购买DIS服务。

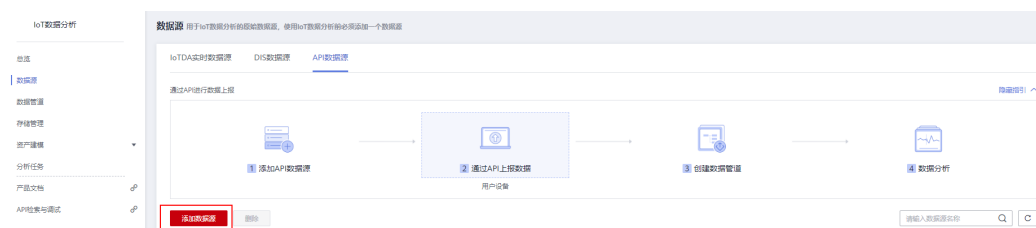
3.4 数据源--添加 API 数据源

操作步骤

步骤1 登录进入“IoT数据分析”服务控制台。



步骤2 在左侧导航栏选择“数据源>API数据源”进入页面。



步骤3 点击“添加数据源”按钮，在弹出的页面中输入数据源名称，点击“确定”按钮，完成API数据源的添加。您可通过第三方APP（例如postman）向这个API数据源的URL发送数据。



---结束

4 数据管道

4.1 创建数据管道

4.1.1 使用 IoTDA 数据源

前置条件

需要在IoT数据分析服务 - 数据源环节已创建好IoTDA实时数据源。

操作步骤

- 步骤1** 访问“IoT数据分析”服务产品页，点击“立即使用”进入“IoT数据分析”服务控制台。
- 步骤2** 在左侧导航栏选择“数据管理”，进入数据管道页面，在“管理列表”点击“创建数据管道”。



步骤3 弹出“创建数据管道”对话框，您可自行填写“作业名称”（注意：同一用户所创建的作业名称不可重复）、“描述”，并选择数据源（若列表为空，请跳转到“创建数据源”链接创建数据源，详见数据源章节），如下图所示。

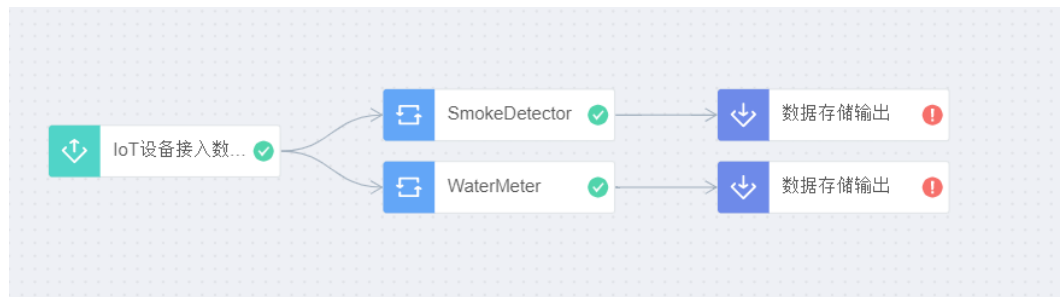
The screenshot shows a dialog box titled "创建数据管道" (Create Data Pipeline) with a close button (X) in the top right corner. It contains three main input areas:

- 作业名称 (Job Name):** A text input field with the placeholder "请输入作业名称" (Please enter job name).
- 描述 (Description):** A text area with the placeholder "请添加一句话简短说明此流程的用途" (Please add a sentence to briefly describe the purpose of this process) and a character count "0/50" at the bottom right.
- 选择数据源 (Select Data Source):** A dropdown menu with a "创建数据源" (Create Data Source) link to its right.

At the bottom of the dialog, there are two buttons: "确定" (Confirm) and "取消" (Cancel).

----结束

使用IoTDA数据源时，系统会自动读取该数据源的相关信息及相关的产品信息，生成作业执行流程，如图所示。



上图所示示例中，所选择的IoTDA数据源包含两种设备：烟感设备与水表，系统自动生成了两个处理分支，分别用于处理这两种设备上报的数据。

这两个分支各自以一个“产品过滤”算子作为起始点，分别过滤出这两种设备的数据，以便于后续处理流程仅处理某一种产品的数据，如图所示。

产品过滤 [?]

^ 基础配置

算子名称 [?]

SmokeDetector

* IoTDA上的资源空间ID [?]

f8c039fbf58741f8b8fecc11794359c6

* 产品ID [?]

5f3a24d604feea02bae78fda

产品过滤 [?]

^ 基础配置

算子名称 [?]

WaterMeter

* IoTDA上的资源空间ID [?]

f8c039fbf58741f8b8fecc11794359c6

* 产品ID [?]

5f3a24b9aea52102badffa4a

系统还自动根据这两种设备的产品信息（该产品信息以安全的方式从IoTDA获取），生成了默认的输出配置。

数据存储输出 [?]

∨ 输入

∧ 基础配置

算子名称 [?]

数据存储输出

数据存储 [?]

* 存储组 [?]

* 数据存储名称 [?]

* 数据存储属性 [?]

* 设备id来源 [?]

header.device_id

* 属性 [?]

* 属性名称 [?]	* 属性类型 [?]	* 源属性 [?]	标签 [?]	单位 [?]	操作
SmokeDete...	INTE... ▾	body.servic...	▾	%	删除
SmokeDete...	INTE... ▾	body.servic...	▾	V	删除
SmokeDete...	INTE... ▾	body.servic...	▾	%	删除
SmokeDete...	INTE... ▾	body.servic...	▾		删除
SmokeDete...	INTE... ▾	body.servic...	▾	%	删除

请用户自行选择存储组，并填写存储名称，补齐配置内容即可保存，如图所示。

* 存储组 [?]

IoTDA 设备存储 ×

* 数据存储名称 [?]

Smoke

4.1.2 使用 DIS 数据源

点击导航栏上的“数据管理”标签，进入数据管道页面，在“管理列表”部分有一按钮“创建数据管道”，如图所示：



点击“创建数据管道”按钮，弹出“创建数据管道”对话框，用户可自行填写“作业名称”（注意：同一用户所创建的作业名称不可重复）、“描述”，并选择数据源（若列表为空，请跳转到“创建数据源”链接创建数据源，详见数据源章节），如下图所示：

创建数据管道 ×

★ 作业名称

描述

★ 选择数据源 [创建数据源](#)

目前可支持的数据源两种类型：IoTDA数据源、DIS数据源。

使用DIS数据源时，需要用户自行配置数据源及数据处理流程，如图所示：

DIS数据源 [?]

^ 基础配置

算子名称 [?]

DIS数据源

* 数据源 [?]

iotadis ×

事件时间 [?]

JSON路径 [?]

eventtime

时间样式 [?]

yyyyMMdd'THHmmss'Z'

时区 [?]

- 8 +

且不再需要“产品过滤”算子：



4.1.3 其他数据源的使用

其他类型数据源的数据管道的创建操作与IoTDA实时数据源、DIS数据源操作类似，因此不再详述。

4.2 编辑处理流程

4.2.1 数据源配置

1. IoTDA数据源

在画布上选中“IoT设备接入数据源”算子，在右侧配置面板处，可修改算子名称、选择数据源、填写事件时间字段路径、时间格式，如图。

IoT设备接入数据源 ⓘ

^ 基础配置

算子名称 ⓘ

* 数据源 ⓘ

 X

⚠ 注意

使用IoTDA数据源时，系统会自动根据所选择的数据源生成这些配置，用户可检查这些配置是否正确，根据自己的需要进行修改。

2. DIS数据源配置

在画布上选中“DIS数据源”算子，在右侧配置面板处，可修改算子名称、选择数据源、填写事件时间字段路径、时间格式、时区，如图。

DIS数据源 ⓘ

^ 基础配置

算子名称 ⓘ

* 数据源 ⓘ

 X

事件时间 ⓘ

JSON路径 ⓘ

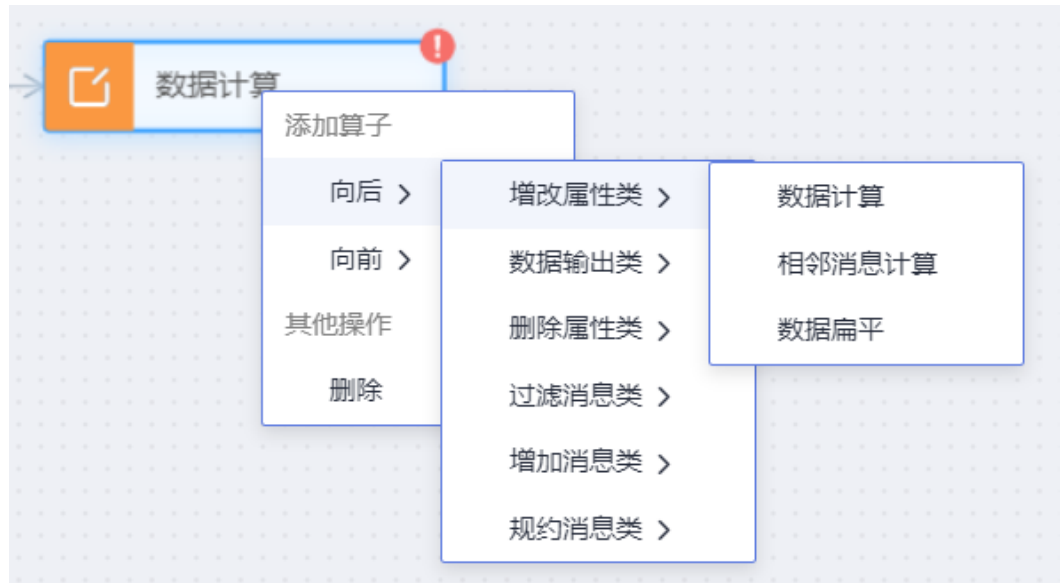
时间样式 ⓘ

时区 ⓘ

 - 8 +

4.2.2 数据清洗配置

目前提供了产品过滤、增改属性类、删除属性类、过滤消息类、增加消息类、规约消息类几个类型的数据清洗算子。增加清洗步骤的操作方式为，在算子节点上点击右键，可以看到“向前”、“向后”两个菜单，如图所示：



📖 说明

- 部分算子上仅显示“向前”或“向后”菜单，则说明该算子上仅可“向前”添加或“向后”添加算子。

接下来，以产品过滤、数据计算、选择属性、数据过滤、超时插补、数据聚合为例，分别介绍一下不同类型的算子配置。各个算子配置的详细说明请参考“[算子使用指南](#)”。

1. 产品过滤

“产品过滤”仅可添加在“IoT设备接入数据源”算子后，因此，仅在“IoT设备接入数据源”算子上点击右键时，可以看到“产品过滤”算子。

添加好算子节点后，在算子节点上点击鼠标左键，即可打开右侧的算子配置面板，如图：

产品过滤 ②

^ 基础配置

算子名称 ②

SmokeDetector

* IoTDA上的资源空间ID ②

f8c039fbf58741f8b8fecc11794359c6

* 产品ID ②

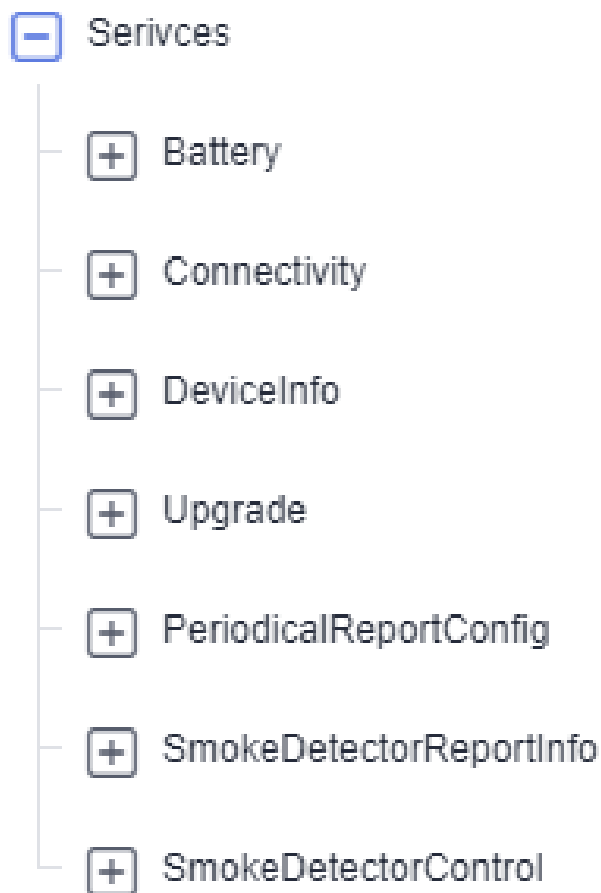
5f3a24d604feea02bae78fda

∨ 高级配置

∨ 输出

系统会根据所选择的产品，自动生成该算子的输出内容：

^ 输出



2. 数据计算

按照表达式进行数值计算，计算的结果可以赋值给源属性，也可以自定义属性名称及路径。

添加好算子节点后，在算子节点上点击鼠标左键，即可打开右侧的算子配置面板，如图：

数据计算 ②

∨ 输入

∧ 基础配置

算子名称 ②

数据计算

变量 ②

添加属性

* 变量名称 ②	* 取值 ②	操作
batteryVoltage	body.services[?(@.sen	删除
batteryLevel	body.services[?(@.sen	删除

* 计算配置 ②

添加属性

* 属性 ②	* 表达式 ②	操作
body.services[?(@.sen	batteryLevel == 1	删除
body.services[?(@.sen	batteryLevel == 2	删除

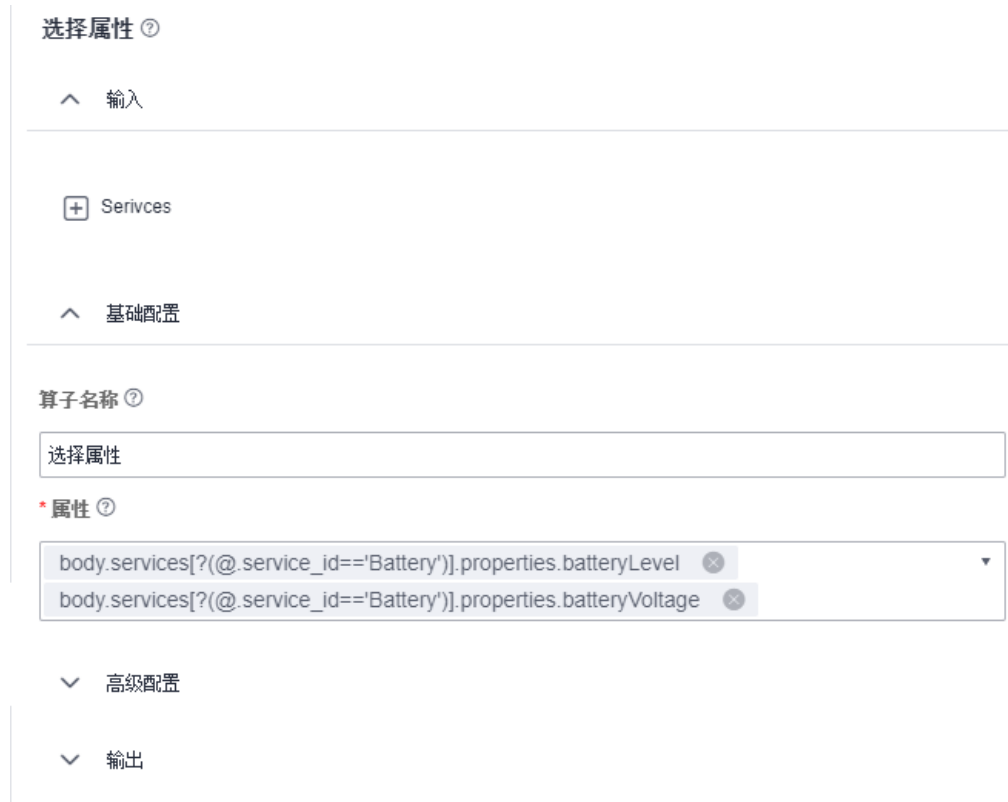
∨ 高级配置

∨ 输出

3. 选择属性

在输入数据中选择保留的属性，未选择的属性被删除。当数据字段较多时，用户可使用该功能筛选出业务所关系的字段。

添加好算子节点后，在算子节点上点击鼠标左键，即可打开右侧的算子配置面板，如图：



使用该配置筛选出batteryLevel与batteryVoltage两个字段，因为后续处理流程仅使用这两个字段。

4. 数据过滤


根据设置的条件过滤数据，满足条件的数据则输出，不满足条件的数据将被丢弃。

添加好算子节点后，在算子节点上点击鼠标左键，即可打开右侧的算子配置面板，如图：


数据过滤

∨ 输入

∧ 基础配置

算子名称 

数据过滤

条件关系 

AND

* 过滤条件 

[添加属性](#)

* 属性 	* 计算符号 	* 数值 	操作
body.services[?(@.	MORE_T... 	100	删除
body.services[?(@.	NOT_EQ... 	0	删除

∨ 高级配置

∨ 输出



使用该配置过滤出batteryLevel不等于0且batteryVoltage大于100的数据。
数据过滤算子不会修改数据输出，用户可展开“输出”项检查输出内容。

5. 超时插补

提供基于时间的数据插补功能。时间支持延迟等待时长，以及指定固定时间点。
插补数据为上一条实际数据，但是事件时间为插补计算出的时间。

添加好算子节点后，在算子节点上点击鼠标左键，即可打开右侧的算子配置面板，如图：

超时插补 [?]

∨ 输入

∧ 基础配置

算子名称 [?]

超时插补

* 分区键 [?]

添加属性

body.services[?(@.service_id=='ATXT')].properties.SI

删除

* 插补时间类型

RELATIVE

相对时间插补周期 [?]

—

60

+

绝对时间插补周期 [?]

00***?

绝对周期的等待时长 [?]

—

+

存活时长 [?]

—

∧ 高级配置

并行数 [?]

—

∧ 输出

Services

└─ ATXT

6. 数据聚合

数据聚合算子,对数据流中的数据根据窗口周期进行汇聚计算。

添加好算子节点后，在算子节点上点击鼠标左键，即可打开右侧的算子配置面板，如图：

数据聚合 ?

▼ 输入

^ 基础配置

算子名称 ?

数据聚合

分区字段 ? 添加属性

body.services[?(@.service_id=='ATXT')].properties.VBGH 删除

• 窗口配置 ?

• 窗口大小 ?

— 10 +

• 最大消息数 ?

— 10,000 +

• 聚合字段 ? 添加属性

• 字段名称 ?	• 字段类型 ?	• 被汇聚字段 ?	• 聚合类型 ?	操作
body.services[?(DOUBLE ▼	body.services[?(AVERAGE ▼	删除
body.services[?(DOUBLE ▼	body.services[?(SUM ▼	删除

^ 高级配置

并行数 ?

—

^ 输出

[-] Services

 [-] ATXT

4.2.3 数据输出配置

1. 数据存储输出

将数据输出到IoT数据分析服务的存储，以便进一步分析（资产建模、实时分析、时序分析）使用。

用户可以配置算子名称、选择存储组、填写存储名称，并选择所需要的属性及属性类型等信息，如图所示：

数据存储输出 [?]

∨ 输入

∧ 基础配置

算子名称 [?]

Prodddd产品的数据输出

* 数据存储 [?]

* 存储组名称 [?]

IoTDA设备存储

* 数据存储名称 [?]

Prodddd产品

* 数据存储属性 [?]

* 设备id来源 [?]

header.device_id

* 属性 [?]

[添加属性](#)

* 属性名称 [?]	* 属性类型 [?]	* 源属性 [?]	标签 [?]	单位 [?]	操作
ATXT_SDFSDF	INTEGER	body.services...	FALSE		删除
ATXT_CVXVC...	DOUBLE	body.services...	FALSE		删除
ATXT_BNGBGH	STRING	body.services...	FALSE		删除
ATXT_GHJJ	STRING	body.services...	FALSE		删除
ATXT_VBGH	OBJECT	body.services...	FALSE		删除

2. DIS输出

将数据输出到DIS云服务，IoT数据分析服务仅在华为云部署时支持输出到DIS云服务。

用户在“数据源”页面配置好数据源后，在该算子配置处选择即可，如图所示：

DIS输出 ②

▼ 输入

▲ 基础配置

算子名称 ①

* DIS数据源 ②

 ×

▼ 高级配置

4.3 调试数据管道

4.3.1 调试概述

- 为什么管道任务没能输出期望的结果？
- 管道任务中哪个算子环节出错了？
- 算子处理后的数据结果是什么样的？

以上问题是我们在开发阶段经常遇到的问题。为了解决以上问题，我们提供了调试数据管道工具。用户可以上报测试数据对管道业务流程进行测试，用户可直观查看测试数据在每个算子输入和输出的内容。

工作原理：

启动测试模式后，输入类算子从数据源读取数据，将数据发送到下游的转化类算子，最终专递到输出类算子。调试工具会在每个算子的输入输出端监听数据流，最终将监听收集到的数据在页面进行展示。我们可以根据算子的输入输出数据定位以上我们所遇到的问题。



[界面介绍](#)

[启动按钮](#)

[任务状态栏](#)

[输入输出](#)

[数据监控](#)

[操作指导](#)

[创建任务](#)

[启动测试任务](#)[下载数据模板](#)[修改数据模板](#)[上传数据文本](#)[查看测试结果](#)[结束测试任务](#)[数据模板](#)[IoT设备接入数据源](#)[API接入数据源](#)

表 4-1 算子支持表

算子类别	算子中文名称	算子英文名称	是否支持调试	是否展示输入数据	是否展示输出数据
输入类算子	IoT设备接入数据源	IoTDataSource	是	否	是
	API接入数据源	APISource	是	否	是
	DIS数据源	DISSource	否	否	否
转化类算子	增加属性	AddProperties	是	是	是
	数据计算	Calculate	是	是	是
	删除属性	DeleteProperties	是	是	是
	过滤算子	Filter	是	是	是
	数据扁平	Flatten	是	是	是
	产品过滤	ProductFilter	是	是	是
	选择属性	SelectProperties	是	是	是
	相邻消息计算	ContiguousCalculate	是	是	是
	超时插补	TimeoutInterpolation	是	是	是
输出类算子	数据存储输出	DataStoreSink	是	是	否
	DIS输出	DISSink	是	是	否
	实时分析输出	RTASink	是	是	否

📖 说明

输入类算子只记录输出数据，输出类算子只记录输入数据。

4.3.2 界面介绍

- 启动按钮
 - 启动调试任务



- 任务状态栏
 - 查看任务状态。



状态说明

状态名称	描述
就绪	任务完成编辑后, 无错误配置, 等待启动。
提交中	任务提交后台服务中, 此时任务处于初始化中, 无法处理上报的测试数据
运行中	任务已完成初始化, 等待或正在处理上报测试数据, 除当任务异常或人为停止, 否则将一直正常运行。
停止中	当任务异常或人为停止任务时, 任务将停止处理测试数据。
停止	任务正常停止, 可编辑。

- 输入输出
 - 上报测试数据的控制页面。用户可在数据源类型选择对应数据源; 下载对应数据源的测试数据模板文本并修改数据内容; 最后上传测试数据文本。




- 数据监控

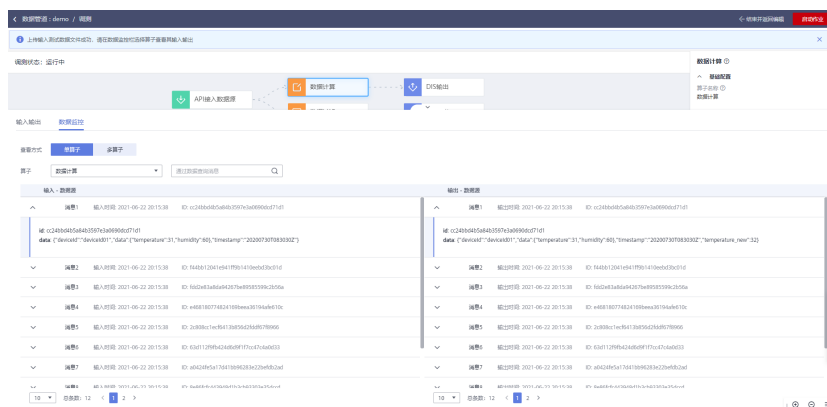
- 数据监控也是查看数据处理结果的查询页面。页面查看方式包括单算子查询和多算子查询。

- 单算子查询

- 下拉“算子”下拉框选择查看“数据计算”。


- 点击搜索输入框  按钮搜索查询，搜索框支持以消息id或data内容匹配查询。

-  和  分别显示“数据计算”算子结果。

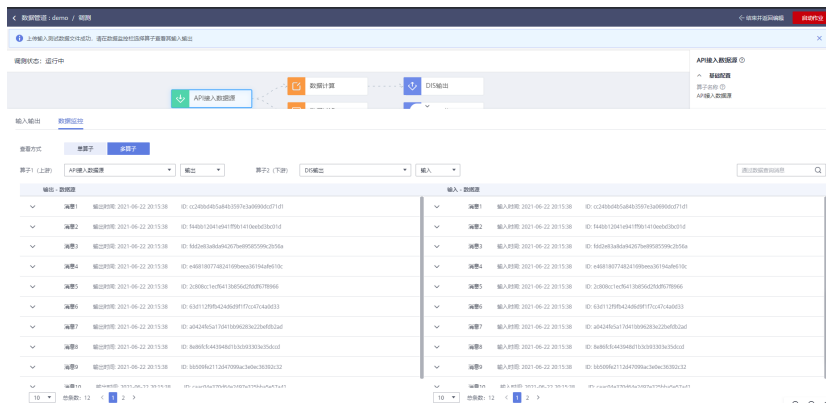


- 多算子查询

- 下拉“算子1（上游）”下拉框选择查看“API接入数据源”的“输出”；下拉“算子2（下游）”下拉框选择查看“DIS输出”的“输入”。

- 点击搜索输入框  按钮搜索查询，搜索框支持以消息id或data内容匹配查询。

- **输入 - 数据源** 显示“API接入数据源”的“输出”结果。
- **输出 - 数据源** 显示“DIS输出”的“输入”结果。



4.3.3 操作指导

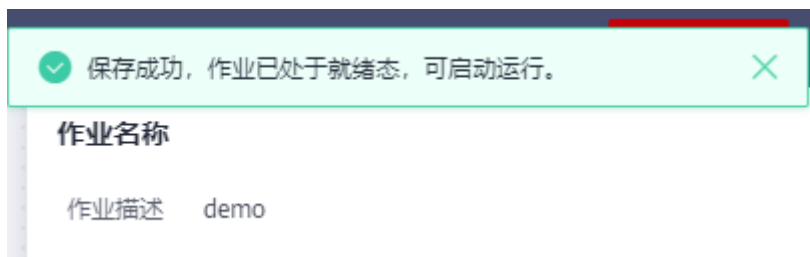
使用测试功能主要包含启动测试任务，上报数据，查看测试结果，停止测试任务。其中上报数据，查看测试结果可重复执行，不必反复启停任务。

步骤1 创建任务

1. 创建并编辑任务。
2. 保存任务。



3. 确保任务为“就绪”态。



步骤2 启动测试任务

1. 点击测试按钮，等待测试状态为“运行中”。目前部分算子暂不支持测试功能(见[算子支持表](#))，测试按钮将不可点击。



2. 自动跳转测试页面，当前状态为提交中，需要耐心等待。状态包含：“就绪”、“提交中”、“运行中”、“停止中”、“停止”。



3. 当调测状态为“运行中”时，完成启动测试任务操作。



步骤3 下载数据模板

1. 点击下载测试模板。



📖 说明

不同的输入算子的测试模板内容格式会不一样，建议在模板的基础上添加测试数据。

步骤4 修改数据模板

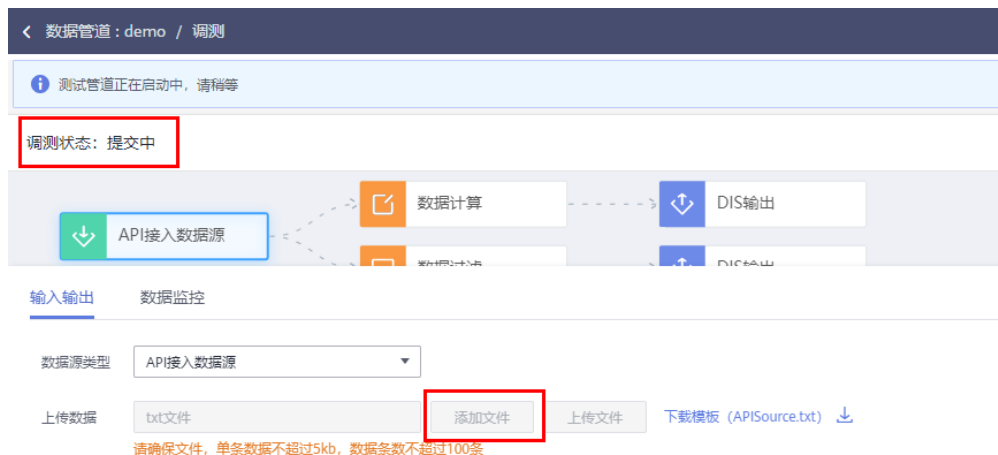
1. 删除多余不需要的数据，写入业务数据内容，参考[数据模板](#)章节。

📖 说明

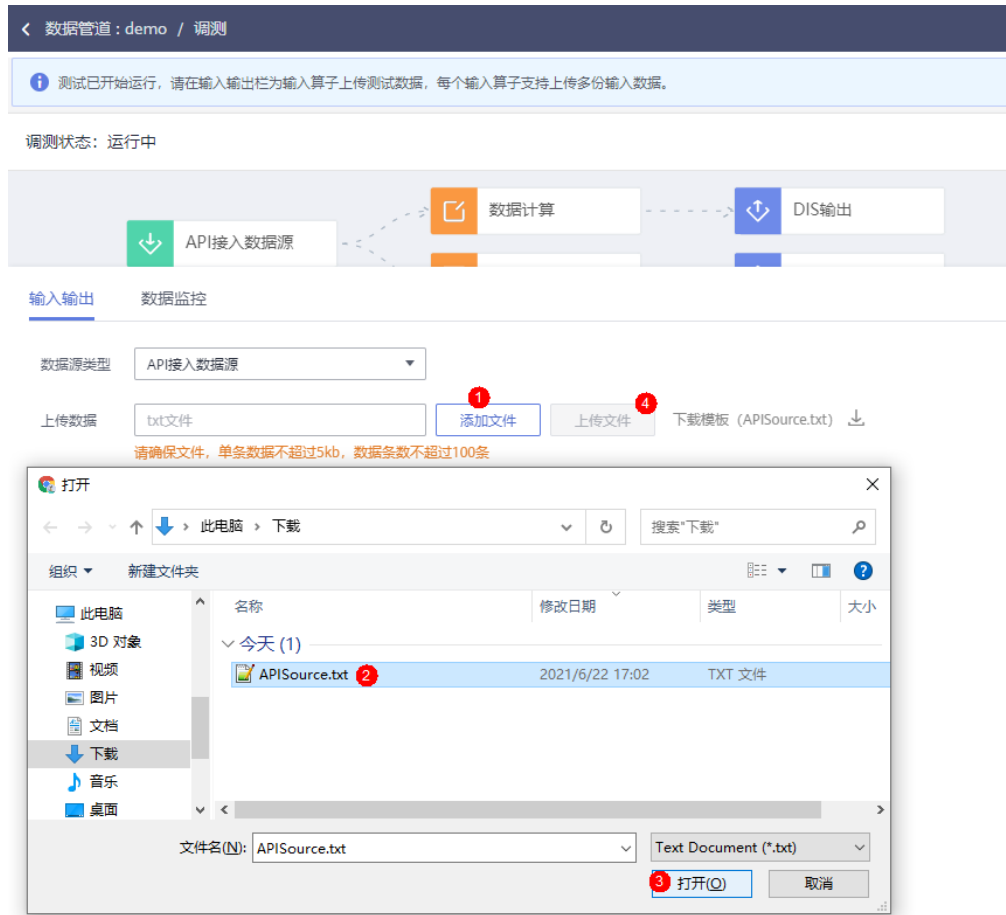
- 测试数据文本支持UTF-8编码。
- 测试数据必须是合法Json格式，且不换行。
- 测试数据文本可写多条数据，每条数据以换行符分隔，即一条数据必须写在一行文本里。
- 每条数据大小不能超过5kb。
- 一个文本里不能超过100条测试数据。
- 测试数据文本中必填项不能缺损、错误，建议参考测试模板内容。
- 注释内容不会上报

步骤5 上传数据文本

1. 当调试状态为“就绪”或“提交中”时，暂不能添加上传测试数据。



2. 当调测状态为“运行中”时，才能添加上传测试数据。



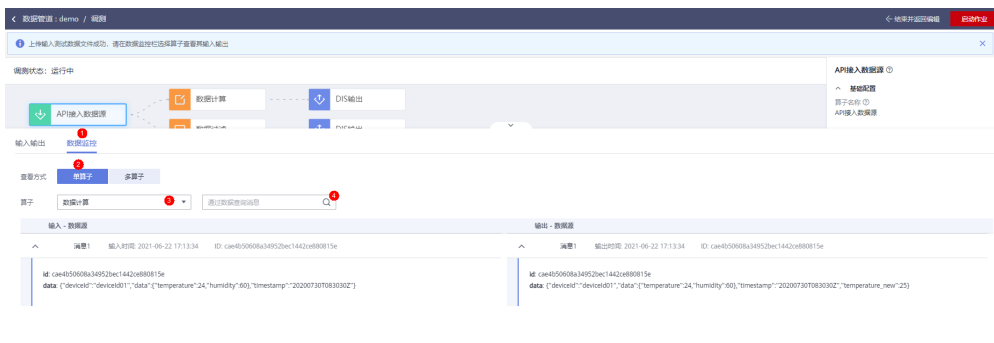
3. 上传成功



步骤6 查看测试结果

1. 切换到数据监控页签
2. 选择查看模式
3. 选择目标算子

4. 点击查询按钮



步骤7 结束测试任务

1. 点击右上角“结束并返回编辑”按钮



- 步骤8 如果需要重复测试，可重复步骤5、步骤6。

----结束

说明

- 调试数据管道需要占用计算资源。
- 一小时内算子没有新增测试数据，测试数据将会被老化。重启测试任务后，测试数据也不会保留。
- 同一个任务只能处于调测状态或正常运行态。

4.3.4 数据模板

表 4-2 IoT 设备接入数据源

数据模板	字段名	字段说明	是否必填
{ "header": {	header	测试数据头字段，必须是json格式。	是

数据模板	字段名	字段说明	是否必填
<pre> "projectId": "080dd91ecc8026bxxxxx1be5f9d68d", "dataSourceId": "0b0c2bf036a5xxxxxc8d2ed602d401f", "pipelineId": "", "dataStoreId": "" }, "body": { "header": { "app_id": "", "device_id": "", "node_id": "", "product_id": "", "gateway_id": "", "tags": [{ "tag_key": "", "tag_value": "" }], "timestamp": "" }, "body": { "services": [{ "data": { "Temperature": "" }, "service_id": "", "service_type": "", "event_time": "20210623T163200+0800" }] } } </pre>	projectId	租户id，必须模板默认值，用户不能修改。	是
	dataSourceId	数据源id，必须模板默认值，用户不能修改。	是
	body	数据体字段，必须是json格式，格式不能修改。	是
	body.header	IoTDA数据消息头，内部app_id, device_id node_id、product_id、gateway_id、tags等字段用户可根据自己业务需求修改字段值。	是
	body.body	IoTDA数据消息体，必须是json格式。	是
body.body.services	IoTDA数据消息内容，必须是json数组格式。用户可自定义数组内元素Json内容和元素个数。	是	

表 4-3 API 接入数据源

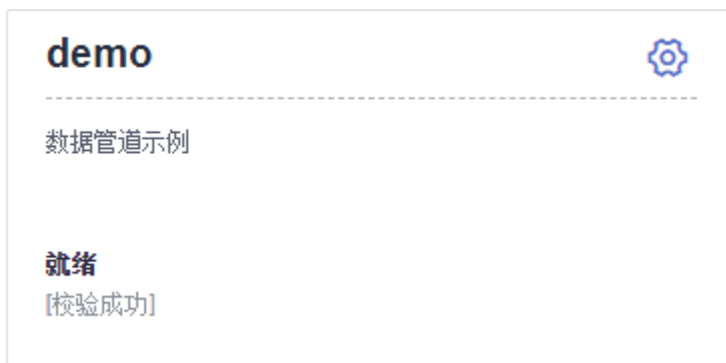
数据模板	字段名	字段说明	是否必填
<pre>{ "header": { "projectId": "080dd91ecc8026bxxxxx1be5f9d 68d", "dataSourceId": "0b0c2bf036a5xxxxxc8d2ed602d 401f", "pipelineId": "", "dataStoreId": "" }, "body": { "header": { "app_id": "", "device_id": "", "node_id": "", "product_id": "", "gateway_id": "", "tags": [{ "tag_key": "", "tag_value": "" }], "timestamp": "" }, "body": { "services": [{ "data": { "Temperature": "" }, "service_id": "", "service_type": "" }] } } }</pre>	header	测试数据头字段，必须是json格式。	是
	projectId	租户id，必须模板默认值，用户不能修改。	是
	dataSourceId	数据源id，必须模板默认值，用户不能修改。	是
	body	数据体字段，必须是json格式，用户可自定义自己内容。	是
	body.header	样例数据，用户可自定义自己内容。	否
	body.body	样例数据，用户可自定义自己内容。	否

数据模板	字段名	字段说明	是否必填
<pre>"event_time": "20210623T163200+0800" }] } } }</pre>	body.body.services	样例数据，用户可自定义自己内容。	否

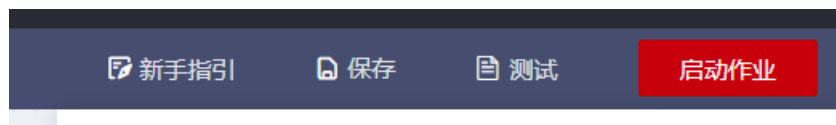
4.4 运行数据管道

4.4.1 运行数据管道

数据管道编辑完成后，若存在错误，作业状态会变成“草稿”态，请检查作业配置；若不存在错误，作业状态会变为“就绪”态，表示该作业可以运行，如图所示：



“就绪”态的作业即可运行：在“管理列表页”点击作业卡片，进入“作业详情页”，然后点击右上角的红色按钮“启动作业”。



启动作业时，系统会自行判断用户的作业是否需要实时计算资源，若需要实时计算资源，则会弹出如下对话框供用户自行选择：

* RTU数 ?

* 并行数 ?

计算资源：一个计算资源对应的物理资源为：1核CPU、4G内存，可实时清洗的数据量为1000条/S（仅供参考，请结合各自的业务场景与数据大小进行评估）。

并行数：管道作业运行时的实际并发能力，不得超过“计算资源 - 1”的4倍。

4.4.2 运行过程监控与数据跟踪（敬请期待）

4.5 算子使用指南

4.5.1 输入类算子

4.5.1.1 IoT 设备接入数据源

算子简介

- 名称：IoT设备接入数据源
- 功能说明：使用IoTDA云服务作为pipeline的数据源，接入IoTDA的数据。
- 约束：下游仅允许接产品过滤算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式： $^[\u4E00-\u9FA5A-Za-z0-9_-]{1,64}$$

数据源ID：

- 配置项英文名：datasourceId
- 说明：数据源ID，指定Pipeline作业的IoTDA数据源ID。
数据源ID可以通过[数据源接口](#)：`/v1/${project_id}/datasources?type=IoTDA` 获取。

- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: `^[a-zA-Z0-9][w_]{1,64}$`

事件时间:

- 配置项英文名: `eventTime`
- 说明: 事件时间, 以JSON Path格式来指定事件时间的获取来源。如果没有指定, 则使用接受消息时的系统时间作为消息产生的事件时间。
- 必选: 是
- 子配置项: 事件时间包括如下子配置项: JSON路径、时间样式。

JSON路径:

- 配置项英文名: `jsonPath`
- 说明: 事件时间字段的Json路径, 以JSON Path格式来指定事件时间的获取位置。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: `^{1,1000}$`

时间样式:

- 配置项英文名: `pattern`
- 说明: 时间格式, 事件时间的样例格式, 如`yyyyMMdd'T'HHmmss'Z'`、`MM/dd/yyyy hh:mm`等。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: `^{1,100}$`

高级配置项

分配时间戳提取器和水印:

- 配置项英文名: `assignerWithPeriodicWatermarksType`
- 说明: 时间戳分配器, 目前提供3种方式, `BoundedOutOfOrdernessTimestampExtractor`、`AscendingTimestampExtractor`、`IngestionTimeExtractor`。
 - `BoundedOutOfOrdernessTimestampExtractor`: 允许固定时间延迟的分配器, 适用于乱序但预先知道或只允许在流中将会遇到的最大延迟的场景。
 - `AscendingTimestampExtractor`: 递增时间戳的分配器, 适用于每个元素(elements)的时间在每个并行任务(parallel task)里是单调递增的场景。
 - `IngestionTimeExtractor`: 根据获取到数据时的系统时间分配时间戳, 适用于无事件时间的场景。
- 类型: STRING
- 必选: 否, 默认采用 `BoundedOutOfOrdernessTimestampExtractor`。
- 约束: 必须为`BoundedOutOfOrdernessTimestampExtractor`、`AscendingTimestampExtractor`、`IngestionTimeExtractor` 之一。

允许的最大延迟时间:

- 配置项英文名：maxOutOfOrdernessInSeconds
- 说明：允许的最大延迟时间，该配置仅作用于 BoundedOutOfOrdernessTimestampExtractor，具体是在EventTime类型窗口中使用。
- 类型：STRING
- 必选：否，默认为300秒。
- 约束：最小值0，最大值3600秒。

是否拆分services数组：

- 配置项英文名：isSplitServices
- 说明：是否拆分services数组，当数据内容是数组时，用户可以指定是否需要将一条数组类型的数据拆分为多条数据。
- 类型：STRING
- 必选：否，默认为false，即不拆分。
- 约束：仅可填写 TRUE 或 FALSE 。

启动配置：

- 配置项英文名：startup
- 说明：启动配置，用于配置启动时的配置参数。
- 必选：否。
- 子配置项：启动配置包括如下子配置项：启动模式。

启动模式：

- 配置项英文名：startupMode
- 说明：Pipeline启动时数据消费模式，LATEST表示读取最新数据，EARLIEST表示读取最旧的数据。
- 类型：STRING
- 必选：否，默认为 LATEST。
- 约束：仅可填写 LATEST 与 EARLIEST 。

4.5.1.2 DIS 数据源

算子简介

- 名称：DIS数据源，IoT数据分析服务仅在华为云部署形态支持DIS数据源。
- 功能说明：使用DIS云服务作为pipeline的数据源，从DIS中获取数据。
- 约束：下游不允许对接数据源，目前有：IoT设备接入数据源、DIS数据源，及产品过滤算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。

- 类型: STRING
- 必选: 否
- 约束: 需要符合正则表达式: $^[\u4E00-\u9FA5A-Za-z0-9_-]{1,64}$$

数据源ID:

- 配置项英文名: datasSourceId
- 说明: 数据源ID, 指定Pipeline作业的IoTDA数据源ID。
数据源ID可以通过[数据源接口](#): /v1/\${project_id}/datasources?type=DIS获取。
- 类型: STRING
- 必选: 否
- 约束: 需要符合正则表达式: $^[\u4E00-\u9FA5A-Za-z0-9_-]{1,64}$$

事件时间:

- 配置项英文名: eventTime
- 说明: 事件时间, 以JSON Path格式来指定事件时间的获取来源。如果没有指定, 则使用接受消息时的系统时间作为消息产生的事件时间。
- 必选: 是
- 子配置项: 事件时间包括如下子配置项: JSON路径、时间样式、时区。

JSON路径:

- 配置项英文名: jsonPath
- 说明: 事件时间字段的Json路径, 以JSON Path格式来指定事件时间的获取位置。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $^.{1,100}$$

时间样式:

- 配置项英文名: pattern
- 说明: 时间格式, 事件事件的样例格式, 如yyyyMMdd'T'HHmmss'Z'、MM/dd/yyyy hh:mm等。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $^.{1,100}$$

时区:

- 配置项英文名: timeZone
- 说明: 时区, 以UTC偏移量表示, 请填写-12 ~ 12的数字; 若不填写, 则默认为UTC时间。
- 类型: INTEGER
- 必选: 否, 默认为0时区。
- 约束: 最小值-12, 最大值12。

高级配置项

分配时间戳提取器和水印：

- 配置项英文名：assignerWithPeriodicWatermarksType
- 说明：时间戳分配器，目前提供3种方式，BoundedOutOfOrdernessTimestampExtractor、AscendingTimestampExtractor、IngestionTimeExtractor。
 - BoundedOutOfOrdernessTimestampExtractor：允许固定时间延迟的分配器，适用于乱序但预先知道或只允许在流中将会遇到的最大延迟的场景。
 - AscendingTimestampExtractor：递增时间戳的分配器，适用于每个元素(elements)的时间在每个并行任务(parallel task)里是单调递增的场景。
 - IngestionTimeExtractor：根据获取到数据时的系统时间分配时间戳，适用于无事件时间的场景。
- 类型：STRING
- 必选：否，默认采用 BoundedOutOfOrdernessTimestampExtractor 。
- 约束：必须为BoundedOutOfOrdernessTimestampExtractor、AscendingTimestampExtractor、IngestionTimeExtractor 之一 。

允许的最大延迟时间：

- 配置项英文名：maxOutOfOrdernessInSeconds
- 说明：允许的最大延迟时间，该配置仅作用于 BoundedOutOfOrdernessTimestampExtractor，具体是在EventTime类型窗口中使用。
- 类型：STRING
- 必选：否，默认为300秒。
- 约束：最小值0，最大值3600秒。

启动配置：

- 配置项英文名：startup
- 说明：启动方式，用于配置启动时的配置参数。
- 必选：否。
- 子配置项：启动配置包括如下子配置项：启动模式。

启动模式：

- 配置项英文名：startupMode
- 说明：Pipeline启动时数据消费模式，GROUP_OFFSETS表示从上次读取到的位置继续读取，LATEST表示读取最新数据，EARLIEST表示读取最旧的数据。
- 类型：STRING
- 必选：否，默认为 GROUP_OFFSETS。
- 约束：仅可填写GROUP_OFFSETS、LATEST 与 EARLIEST 。

4.5.1.3 API 数据源

算子简介

- 名称：API接入数据源

- 功能说明：通过API进行数据上报到pipeline的数据源。
- 约束：下游不允许对接数据源。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式：`^\u4E00-\u9FA5A-Za-z0-9_-]{1,64}$`

数据源ID：

- 配置项英文名：datasourceId
- 说明：指定管道的输入IoT数据接入云服务数据源。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^[a-zA-Z0-9][\w_-]{1,64}$`

事件时间：

- 配置项英文名：eventTime
- 说明：事件时间，以JSON Path格式来指定事件时间的获取来源。如果没有指定，则使用接受消息时的系统时间作为消息产生的事件时间。
- 必选：是
- 子配置项：事件时间包括如下子配置项：JSON路径、时间样式。

JSON路径：

- 配置项英文名：jsonPath
- 说明：事件时间字段的Json路径，以JSON Path格式来指定事件时间的获取位置。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^{1,1000}$`

时间样式：

- 配置项英文名：pattern
- 说明：时间格式，事件时间的样例格式，比如yyyyMMdd'T'HHmmss'Z'或yyyy-MM-dd'T'HH:mm:ss.SSS'Z'或yyyy-MM-dd'T'HH:mm:ss'Z'等。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^{1,100}$`

高级配置项

分配时间戳提取器和水印：

- 配置项英文名：assignerWithPeriodicWatermarksType
- 说明：时间戳分配器，目前提供3种方式，BoundedOutOfOrdernessTimestampExtractor、AscendingTimestampExtractor、IngestionTimeExtractor。
 - BoundedOutOfOrdernessTimestampExtractor：允许固定时间延迟的分配器，适用于乱序但预先知道或只允许在流中将会遇到的最大延迟的场景。
 - AscendingTimestampExtractor：递增时间戳的分配器，适用于每个元素(elements)的时间在每个并行任务(parallel task)里是单调递增的场景。
 - IngestionTimeExtractor：根据获取到数据时的系统时间分配时间戳，适用于无事件时间的场景。
- 类型：STRING
- 必选：否，默认采用 BoundedOutOfOrdernessTimestampExtractor 。
- 约束：必须为BoundedOutOfOrdernessTimestampExtractor、AscendingTimestampExtractor、IngestionTimeExtractor 之一 。

允许的最大延迟时间：

- 配置项英文名：maxOutOfOrdernessInSeconds
- 说明：允许的最大延迟时间，该配置仅作用于 BoundedOutOfOrdernessTimestampExtractor，具体是在EventTime类型窗口中使用。
- 类型：STRING
- 必选：否，默认为300秒。
- 约束：最小值0，最大值3600秒。

启动配置：

- 配置项英文名：startup
- 说明：启动方式，用于配置启动时的配置参数。
- 必选：否。
- 子配置项：启动配置包括如下子配置项：启动模式。

启动模式：

- 配置项英文名：startupMode
- 说明：Pipeline启动时数据消费模式，LATEST表示读取最新数据，EARLIEST表示读取最旧的数据。
- 类型：STRING
- 必选：否，默认为 LATEST。
- 约束：仅可填写 LATEST 与 EARLIEST 。

4.5.2 输出类算子

4.5.2.1 数据存储输出

算子简介

- 名称：数据存储输出
- 功能说明：数据输出到指定的数据存储。
- 约束：下游不允许接任何算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式：`^\u4E00-\u9FA5A-Za-z0-9_]{1,64}$`

数据存储：

- 配置项英文名：dataStore
- 说明：数据存储，如果需要把管道输出数据保存，则需要指定存储。如果没有指定存储，管道输出数据则不保存。
- 必选：是
- 子配置项：数据存储配置包含如下子配置项：存储组ID、存储组名称、数据存储ID、数据存储名称、数据存储属性。

存储组ID：

- 配置项英文名：dataStorageGroupId
- 说明：指定输出到数据存储的存储组ID，当使用Console进行操作时，该选项会自动填充，因此Console中不显示存储组ID配置项；当使用API时，需要用户自行输入。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^[a-zA-Z0-9_-][\w-]{0,256}$`

存储组名称：

- 配置项英文名：dataStorageGroupName
- 说明：指定存储组的名称。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^[-a-zA-Z0-9_\u4E00-\u9FA5]{1,64}$`

数据存储ID：

- 配置项英文名：dataStoreId

- 说明：指定某个数据存储，当使用Console进行操作时，该选项会自动填充，因此Console中不显示存储ID配置项；当使用API时，需要用户自行输入。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^[a-zA-Z0-9_-][\w-]{0,256}$`

数据存储名称：

- 配置项英文名：dataStoreName
- 说明：指定数据存储的名称，若存储ID不存在，则使用该名称创建新的存储。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^[-a-zA-Z0-9_\u4E00-\u9FA5]{1,64}$`

数据存储属性：

- 配置项英文名：dataStoreProperty
- 说明：数据存储属性，这些属性即为管道的输出数据。
- 必选：是
- 子配置项：数据存储属性配置包含如下子配置项：设备ID来源、属性。

设备ID来源：

- 配置项英文名：deviceId
- 说明：设备id来源，以JSON Path格式来指定设备ID的获取来源。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^{1,1000}$`

属性：

- 配置项英文名：properties
- 说明：管道输出的属性列表，仅在此列表的属性才能输出。
- 必选：是
- 子配置项：属性配置包含如下子配置项：属性名称、属性类型、源属性、标签、单位。

属性名称：

- 配置项英文名：name
- 说明：属性名称，对输出属性进行命名。属性名称不允许为deviceId, eventTime, time和sys_前缀,列名不能重复，只能包含数字、字母和下划线，不能以下划线或数字开头，长度为1~64。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^[a-zA-Z0-9_]{1,64}$`。

属性类型：

- 配置项英文名: type
- 说明: 属性类型, 指定属性的格式类型。类型转换失败时, 字段不输出。
- 类型: STRING
- 必选: 是
- 约束: 只支持4种数据类型INTEGER、DOUBLE、STRING、OBJECT。

源属性:

- 配置项英文名: existingProperty
- 说明: 源属性, 以JSON Path格式来指定源属性, 以作为输出属性的取值。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $^{\{1,1000\}}$

标签:

- 配置项英文名: tag
- 说明: 是否作为时序分析的标签。其对应属性的类型必须为STRING, 如果属性的类型不STRING, 此配置忽略。
- 类型: STRING
- 必选: 否, 默认为FALSE, 即不作为标签。
- 约束: 只可以是TRUE或FALSE。

单位:

- 配置项英文名: unit
- 说明: 指定属性的的单位, 仅用于用户填写后传给存储创建DataStore用。
- 类型: STRING
- 必选: 否。
- 约束: 需要符合正则表达式: $^{\{1,50\}}$

4.5.2.2 DIS 输出

算子简介

- 名称: DIS输出, IoT数据分析服务仅在华为云部署形态支持DIS数据源
- 功能说明: 数据输出到DIS云服务。
- 约束: 下游不允许接任何算子。

算子配置

基础配置项

算子名称:

- 配置项英文名: name
- 说明: 算子名称, 对算子进行个性化的命名, 以辅助增加可读性。支持中文、英文、数字、中划线、下划线, 长度小于等于64个字符。
- 类型: STRING

- 必选：否
- 约束：需要符合正则表达式：`^\u4E00-\u9FA5A-Za-z0-9_{1,64}$`

数据源ID：

- 配置项英文名：datasourceId
- 说明：DIS数据源ID，此数据源必须在本服务被授权访问。
数据源ID可以通过[数据源接口](#)：`/v1/${project_id}/datasources?type=DIS`获取。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^[a-zA-Z0-9][\w_-]{1,64}$`

高级配置项

并行数：

- 配置项英文名：parallelism
- 说明：并行数，指定算子执行时的并行任务个数。对于性能要求越高的算子，可以设置较高的并行数，但是占用资源越多。如果此参数不设置，那么作业运行时本算子使用作业的并行数设置。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值 1024。

分区策略：

- 配置项英文名：partitioner
- 说明：分区策略，用于将数据有序地按分区策略写入DIS。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式：`^[a-z_A-Z]\w{0,254}$`。

4.5.2.3 实时分析输出

算子简介

将经过数据管道清洗后的数据输出到实时分析，作为实时分析的数据来源。

算子配置

算子配置项如图所示：

实时分析输出 ?

^ 基础配置

算子名称 ?

实时分析输出

• 管道输出数据名称 ?

fef_rta_out

• 属性 ? 添加属性

• 属性名称 ?	• 属性类型 ?	• 源属性 ?	操作
device_id	STRING	device_id	删除

- 算子名称：用户指定这个算子的名称。
- 管道输出数据名称：用户声明这个输出的名称，以便在实时分析作业的“管道数据输入”算子中使用。
- 属性：用户选择需要将哪些属性输出给实时分析进行后续的分析任务。

4.5.3 转换类算子

4.5.3.1 产品过滤

算子简介

- 名称：产品过滤
- 功能说明：根据IoT设备接入服务（IoTDA）定义产品过滤出指定产品ID的设备数据。此算子仅在数据源为IoTDA时使用。
- 约束：只允许跟在IoT设备接入数据源算子之后。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式： $^[\u4E00-\u9FA5A-Za-z0-9_-]{1,64}$$

IoTDA上的资源空间ID：

- 配置项英文名: appld
- 说明: IoTDA上的资源空间ID。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $^{\{1,64\}}$

产品ID:

- 配置项英文名: productId
- 说明: IoTDA上的产品ID。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $^{\{[a-zA-Z0-9][\w_]{1,64}\}}$

高级配置项**并行数:**

- 配置项英文名: parallelism
- 说明: 并行数, 指定算子执行时的并行任务个数。对于性能要求越高的算子, 可以设置较高的并行数, 但是占用资源越多。如果此参数不设置, 那么作业运行时本算子使用作业的并行数设置。
- 类型: INTEGER
- 必选: 否
- 约束: 最小值1, 最大值 1024。

4.5.3.2 数据计算

算子简介

- 名称: 数据计算
- 功能说明: 按照表达式进行数值计算, 计算的结果赋值给某个属性。举例: 原消息中有温度属性, 其数值是以摄氏度数值表示, 可以通过本算子设置计算公式, 将摄氏温度计算转换成华氏度读数再赋予给原来的温度属性, 或者可以选择将计算转换后的数值赋予一个新属性。
- 约束: 下游不允许对接数据源及产品过滤算子。

算子配置**基础配置项****算子名称:**

- 配置项英文名: name
- 说明: 算子名称, 对算子进行个性化的命名, 以辅助增加可读性。支持中文、英文、数字、中划线、下划线, 长度小于等于64个字符。
- 类型: STRING
- 必选: 否
- 约束: 需要符合正则表达式: $^{\{[\u4E00-\u9FA5A-Za-z0-9_]{1,64}\}}$

变量:

- 配置项英文名: variables
- 说明: 变量, 把输入消息中的属性作为变量在表达式中使用。对变量进行定义, 设置变量名称和变量的取值。
- 必选: 否
- 子配置项: 变量配置包括如下子配置项: 变量名称、变量取值。

变量名称:

- 配置项英文名: variable
- 说明: 变量名称, 用户可自定义一个变量名称, 计算时使用, 以免配置计算表达式时使用jsonPath。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $^[a-zA-Z]\w{0,20}$$

变量取值:

- 配置项英文名: value
- 说明: 变量取值, 以JSON Path格式来指定变量从消息中指定的JSON Path属性获取的取值。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $^{\{0,1000\}}$$

计算配置:

- 配置项英文名: calculates
- 说明: 计算配置, 配置属性的计算表达式。
- 必选: 是
- 子配置项: 计算配置包括如下子配置项: 属性、表达式。

属性:

- 配置项英文名: property
- 说明: 属性名, 表示计算表达式所计算的结果赋予的对象。用户可以选择消息中已有的某个属性, 或者自行输入新的内容, 该内容需要满足Json path语法的要求。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $^{\{1,1000\}}$$

表达式:

- 配置项英文名: expression
- 说明: 计算表达式, 支持的表达式详细说明请参考: [表达式说明](#)。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $^{\{0,1000\}}$$

高级配置项

并行数：

- 配置项英文名：parallelism
- 说明：并行数，指定算子执行时的并行任务个数。对于性能要求越高的算子，可以设置较高的并行数，但是占用资源越多。如果此参数不设置，那么作业运行时本算子使用作业的并行数设置。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值 1024。

4.5.3.3 相邻消息计算

算子简介

- 名称：相邻消息计算
- 功能说明：基于前一消息和当前消息，按照表达式进行数值计算，计算的结果赋值给当前输入消息的属性。
举例：消息中有上报机器的产品总产量，但没有相对上一个上报周期的增量产量。通过相邻消息计算算子，可以用本消息中的产品总量减去上一个消息中的产品总产量，得到的计算结果赋予一个新的属性，比如产品增量。那么经过本算子清洗后的消息中，都将携带产品增量属性了。
- 约束：下游不允许对接数据源及产品过滤算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式：`^[u4E00-\u9FA5A-Za-z0-9_-]{1,64}$`

分区键：

- 配置项英文名：partitionKeys
- 说明：分区键，用于作为消息分区/分组的标识，具体填入的内容为消息中获取的JSON Path，比如可以选择消息中设备ID所在的JSON Path。设置后，拥有该分区键数值相同的所有消息会进入一个分区，比如分区键选择了设备ID后，那么相同设备ID的消息进入一个分区。当选择多个字段时，按指定字段顺序组合成的字符作为算子并行处理的分区的Key。相邻消息计算会针对每个分区中的消息单独作用；举例：1) 选择将设备ID作为分区键，那么本算子会将相同设备ID的消息进行分区，然后进行相邻消息计算。比如设备ID为1的所有消息会进入一个分区，设备ID为2的所有消息会进入另一个分区，计算策略会审视每一个分区中的消息进行单独相邻消息计算；2) 选择消息中的设备ID与设备状态为分区键，那么相同设备ID与设备状态的消息会进入一个分区，然后进行单独的相邻计算。比如设备ID为1、设备状态为正常的所有消息会进入一个分区，设备ID为1、设备状态为异常的

所有消息进入另一个分区，计算策略会审视每一个分区中的消息进行相邻消息计算。

- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $\wedge\{1,1000\}$

前一变量:

- 配置项英文名: previousVariables
- 说明: 前一变量, 把前一条消息中的属性作为变量在表达式中使用。对变量进行定义, 设置变量名称和变量的取值。
- 必选: 是
- 子配置项: 前一变量包括如下子配置项: 变量名称、变量取值。

变量名称:

- 配置项英文名: variable
- 说明: 变量名称, 用户可自定义一个变量名称, 计算时使用, 以免以配置计算表达式时使用jsonPath。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $\wedge[a-zA-Z]\w{0,20}$

变量取值:

- 配置项英文名: value
- 说明: 取值内容, 以JSON Path格式来指定变量从消息中指定的JSON Path属性获取的取值。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $\wedge\{0,1000\}$

当前变量:

- 配置项英文名: variables
- 说明: 当前变量, 把输入消息中的属性作为变量在表达式中使用。对变量进行定义, 设置变量名称和变量的取值。
- 必选: 是
- 子配置项: variables包括如下子配置项: 变量名称、变量取值。

变量名称:

- 配置项英文名: variable
- 说明: 变量名称, 用户可自定义一个变量名称, 计算时使用, 以免以配置计算表达式时使用jsonPath。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $\wedge[a-zA-Z]\w{0,20}$

变量取值:

- 配置项英文名: value
- 说明: 取值内容, 以JSON Path格式来指定变量从消息中指定的JSON Path属性获取的取值。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $\wedge\{0,1000\}$

计算表达式:

- 配置项英文名: calculates
- 说明: 以表达式来计算属性的值。
- 必选: 是
- 子配置项: 计算配置包括如下子配置项: 属性、表达式。

属性:

- 配置项英文名: property
- 说明: 属性名, 表达式计算结果赋值给此属性, 属性以JSON Path格式来表示, 请填写JSON Path。。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $\wedge\{1,1000\}$

表达式:

- 配置项英文名: expression
- 说明: 计算表达式, 支持的表达式详细说明请参考: [表达式说明](#)。
- 类型: STRING
- 必选: 是
- 约束: 需要符合正则表达式: $\wedge\{0,1000\}$

存活时长:

- 配置项英文名: ttl
- 说明: 存活时长, 配置前一变量的存活时长。如果超过此时长没有接收到新的消息更新前一变量取值, 则前一变量将被老化清除。存活时长的单位为秒, 最小清除时间为1s, 最大清除时间为30天, 如果不设置此参数表示永远不清除。当算子每次接收了消息, 更新延长存活时长。
- 类型: INTEGER
- 必选: 否
- 约束: 最小值1, 最大值 2592000, 即30天。

使用输出:

- 配置项英文名: usingOutput
- 说明: 使用输出, 是以算子计算的输出消息来获取前一变量, 还是算子输入的原始消息来获取前一变量。通过该参数的设置, 实际控制了上一条消息的缓存策

略，即缓存上一条消息的原始状态，还是缓存上一条消息经过相邻计算算子处理后的状态，默认为以算子计算更新后的输出消息来获取前一变量。算子输出消息会被表达式修改，会影响前一变量的取值。

- 类型：STRING
- 必选：否，默认为TRUE
- 约束：仅可填写TRUE 或 FALSE

高级配置项

并行数：

- 配置项英文名：parallelism
- 说明：并行数，指定算子执行时的并行任务个数。对于性能要求越高的算子，可以设置较高的并行数，但是占用资源越多。如果此参数不设置，那么作业运行时本算子使用作业的并行数设置。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值 1024。

4.5.3.4 数据扁平

算子简介

- 名称：数据扁平
- 功能说明：数据扁平算子，把数据流中的嵌套Json字段转换为多个独立字段。
- 约束：下游不允许对接数据源及产品过滤算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式： $^[\u4E00-\u9FA5A-Za-z0-9_-]{1,64}$$

扁平字段配置：

- 配置项英文名：flatFiedConfig
- 说明：扁平字段配置，配置需要进行扁平处理的字段列表。
- 必选：是
- 子配置项：扁平字段配置包括如下子配置项：源数据路径、字段名称。

源数据路径：

- 配置项英文名：path

- 说明：源数据路径，以JSON Path格式来指定变量从消息中指定的JSON Path属性获取的取值。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^.{1,1000}$`

字段名称：

- 配置项英文名：field
- 说明：字段名称，输出字段的字段名称。如果字段名与上游输入字段名相同，则此字段被重新赋值。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^[a-z_A-Z]\w{0,64}$`

高级配置项

并行数：

- 配置项英文名：parallelism
- 说明：并行数，指定算子执行时的并行任务个数。对于性能要求越高的算子，可以设置较高的并行数，但是占用资源越多。如果此参数不设置，那么作业运行时本算子使用作业的并行数设置。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值 1024。

4.5.3.5 删除属性

算子简介

- 名称：删除属性
- 功能说明：从数据中删除指定的属性。
- 约束：下游不允许对接数据源及产品过滤算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式：`^[\u4E00-\u9FA5A-Za-z0-9_-]{1,64}$`

属性：

- 配置项英文名：properties
- 说明：属性，以JSON Path格式来指定删除的属性。可以允许设置多个属性，表示删除多个属性。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式： $\wedge\{1,1000\}\$$ 。

高级配置项

并行数：

- 配置项英文名：parallelism
- 说明：并行数，指定算子执行时的并行任务个数。对于性能要求越高的算子，可以设置较高的并行数，但是占用资源越多。如果此参数不设置，那么作业运行时本算子使用作业的并行数设置。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值 1024。

4.5.3.6 选择属性

算子简介

- 名称：选择属性
- 功能说明：在输入数据中选择保留的属性，未选择的属性被删除。
- 约束：下游不允许对接数据源及产品过滤算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式： $\wedge[\u4E00-\u9FA5A-Za-z0-9_-\]{1,64}\$$

属性：

- 配置项英文名：properties
- 说明：属性，以JSON Path格式来指定保留的属性。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式： $\wedge\{1,1000\}\$$ 。

高级配置项

并行数：

- 配置项英文名：parallelism
- 说明：并行数，指定算子执行时的并行任务个数。对于性能要求越高的算子，可以设置较高的并行数，但是占用资源越多。如果此参数不设置，那么作业运行时本算子使用作业的并行数设置。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值 1024。

4.5.3.7 数据过滤

算子简介

- 名称：数据过滤
- 功能说明：根据设置的条件过滤数据，满足条件的数据则输出，不满足条件的数据将被丢弃。
- 约束：下游不允许对接数据源及产品过滤算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式： $^[\u4E00-\u9FA5A-Za-z0-9_-]{1,64}$$

条件关系：

- 配置项英文名：conditionRelationship
- 说明：条件关系，用于指定conditions下的各个条件之间的关系，可以是与、或。
- 类型：STRING
- 必选：否，默认为 AND
- 约束：仅可填写 AND 或 OR

过滤条件：

- 配置项英文名：conditions
- 说明：过滤条件，可以配置多个，多个过滤条件之间的关系由conditionRelationship指定。
- 必选：是
- 子配置项：conditions包括如下子配置项：属性、计算符号、数值。

属性：

- 配置项英文名：property

- 说明：属性，以JSON Path格式来指定属性。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式： $\wedge\{1,1000\}\$$

计算符号：

- 配置项英文名：symbol
- 说明：计算符号，如大于、小于、等于等，目前支持：MORE_THAN、MORE_THAN_EQUAL、LESS_THAN、LESS_THAN_EQUAL、EQUAL、NOT_EQUAL
- 类型：STRING
- 必选：是
- 约束：必须是MORE_THAN、MORE_THAN_EQUAL、LESS_THAN、LESS_THAN_EQUAL、EQUAL、NOT_EQUAL中的一个。

数值：

- 配置项英文名：value
- 说明：数值，即条件的右值，如条件：大于100中的100即为value的值。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式： $\wedge\{0,50\}\$$

高级配置项

并行数：

- 配置项英文名：parallelism
- 说明：并行数，指定算子执行时的并行任务个数。对于性能要求越高的算子，可以设置较高的并行数，但是占用资源越多。如果此参数不设置，那么作业运行时本算子使用作业的并行数设置。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值 1024。

4.5.3.8 超时插补

算子简介

- 名称：超时插补
- 功能说明：提供基于时间的数据插补功能。时间支持延迟等待时长，以及指定固定时间点。插补数据为上一条实际数据，但是事件时间为插补计算出的时间。
- 约束：下游不允许对接数据源及产品过滤算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式：`^\[u4E00-\u9FA5A-Za-z0-9_-\]{1,64}$`

分区键：

- 配置项英文名：keys
- 说明：分区键，用于作为消息分区/分组的标识，具体填入的内容为消息中获取的JSON Path，比如可以选择消息中设备ID所在的JSON Path。设置后，拥有该分区键数值相同的所有消息会进入一个分区，比如分区键选择了设备ID后，所有相同设备ID的消息进入一个分区。当选择多个字段时，按指定字段顺序组合成的字符作为算子并行处理的分区的Key。插补策略会针对每个分区中的消息单独作用；举例：1) 选择将设备ID作为分区键，那么本算子会将相同设备ID的消息进行分区，然后进行单独插补。比如设备ID为1的所有消息会进入一个分区，设备ID为2的所有消息会进入另一个分区，插补策略会审视每一个分区中的消息进行单独插补；2) 选择消息中的设备ID与设备状态为分区键，那么相同设备ID与设备状态的消息会进入一个分区，然后进行单独插补。比如设备ID为1、设备状态为正常的消息会进入一个分区，设备ID为1、设备状态为异常的所有消息进入另一个分区，插补策略会审视每一个分区中的消息进行单独插补。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式：`^\{1,1000}$`

插补时间类型：

- 配置项英文名：timeType
- 说明：插补时间类型，目前支持两种：
 - RELATIVE：相对时间，表示按相对前一条数据偏移时长插补。
 - ABSOLUTELY：绝对时间，表示按照固定时间插补。
- 类型：STRING
- 必选：是
- 约束：仅可填写RELATIVE 或 ABSOLUTELY。

相对时间插补周期：

- 配置项英文名：relativePeriod
- 说明：相对时间插补周期，在插补时间类型为“相对时间”时有效，单位为S。
- 类型：INTEGER
- 必选：否
- 约束：最小值0，最大值为2147483647。

绝对时间插补周期：

- 配置项英文名：absolutelyPeriod
- 说明：绝对时间插补周期，参考java或crontab定时器时间表达式，在插补时间类型为“绝对时间”时有效。

- 类型：INTEGER
- 必选：否
- 约束：需要符合正则表达式： $\wedge\{1,100\}$

绝对周期的等待时长：

- 配置项英文名：delay
- 说明：绝对周期的等待时长，单位秒。即在绝对时间等待一段后再插补消息，避免由于真实消息传送延迟，在插补时间类型为“绝对时间”时有效。
- 类型：INTEGER
- 必选：否
- 约束：最小值0，最大值为2147483647。

存活时长：

- 配置项英文名：ttlTime
- 说明：如果一个设备长期没有报消息，pipeline要停止上一个上报消息的插补定时器。存活时长的单位为秒，最小清除时间为1s，最大清除时间为30天，如果不设置此参数表示永远不清除。当算子每次接收了消息，更新延长存活时长。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值为2592000。

高级配置项

并行数：

- 配置项英文名：parallelism
- 说明：并行数，指定算子执行时的并行任务个数。对于性能要求越高的算子，可以设置较高的并行数，但是占用资源越多。如果此参数不设置，那么作业运行时本算子使用作业的并行数设置。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值 1024。

4.5.3.9 数据聚合

算子简介

- 名称：数据聚合
- 功能说明：数据聚合算子，将数据按窗口周期进行汇聚计算，以达到规约数据的目的。
- 约束：下游不允许对接数据源及产品过滤算子。

算子配置

基础配置项

算子名称：

- 配置项英文名：name
- 说明：算子名称，对算子进行个性化的命名，以辅助增加可读性。支持中文、英文、数字、中划线、下划线，长度小于等于64个字符。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式： $^[\u4E00-\u9FA5A-Za-z0-9_]{1,64}$$

分区字段：

- 配置项英文名：keyFields
- 说明：分区键，用于作为消息分区/分组的标识，具体填入的内容为消息中获取的JSON Path，比如可以选择消息中设备ID所在的JSON Path。设置后，拥有该分区键数值相同的所有消息会进入一个分区，比如分区键选择了设备ID后，所有相同设备ID的消息进入一个分区。当选择多个字段时，按指定字段顺序组合成的字符作为算子并行处理的分区的Key。数据聚合策略会针对每个分区中的消息单独作用；举例：1) 选择将设备ID作为分区键，那么本算子会将相同设备ID的消息进行分区，然后进行聚合。比如设备ID为1的所有消息会进入一个分区，设备ID为2的所有消息会进入另一个分区，该算子会对每一个分区中的消息进行单独聚合；2) 选择消息中的设备ID与设备状态为分区键，那么相同设备ID与设备状态的消息会进入一个分区，然后进行单独聚合。比如设备ID为1、设备状态为正常的所有消息会进入一个分区，设备ID为1、状态为异常的所有消息进入另一个分区，该算子会对每一个分区中的消息进行单独聚合。
- 类型：STRING
- 必选：否
- 约束：需要符合正则表达式： 1,1000$ 。

窗口配置：

- 配置项英文名：windowAssigner
- 说明：指定窗口的参数。
- 必选：是
- 子配置项：windowAssigner包括如下子配置项：[size](#)、[messageLimit](#)。

窗口大小：

- 配置项英文名：size
- 说明：窗口长度，单位为秒。
- 类型：INTEGER
- 必选：是
- 约束：最小值1，最大值3600。

最大消息数：

- 配置项英文名：messageLimit
- 说明：最大消息数，窗口内消息数据限制，取值范围1~2147483647。当时间窗口内接收数据数达到限制，则会触发一次统计并清除当前接收数据缓存（不纳入下次计算），窗口内可触发多次统计；当窗口大小满足且消息数据不满足最大消息数时，会触发一次统计并结束当前窗口计算。
- 类型：INTEGER

- 必选：是
- 约束：最小值1，最大值2147483647。

聚合字段：

- 配置项英文名：aggregateFields
- 说明：聚合字段，配置聚合字段列表。
- 必选：是
- 子配置项：aggregateFields包括如下子配置项：**field**、**type**、**aggregateField**、**aggregateType**。

字段名称：

- 配置项英文名：field
- 说明：字段名称，表达式计算结果赋值给此属性，属性以JSON Path格式来表示，请填写JSON Path。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式： $\wedge\{1,1000\}\$$ 。

字段类型：

- 配置项英文名：type
- 说明：字段类型，输出字段的数据类型。类型转换失败时，字段取值为null。
- 类型：STRING
- 必选：是
- 约束：必须为INTEGER、LONG、FLOAT、DOUBLE之一。

被汇聚字段：

- 配置项英文名：aggregateField
- 说明：被汇聚字段，指定需要进行汇聚计算的字段。
- 类型：STRING
- 必选：是
- 约束：需要符合正则表达式： $\wedge\{1,1000\}\$$ 。

聚合类型：

- 配置项英文名：aggregateType
- 说明：聚合类型，指定聚合计算的方式。SUM、AVERAGE聚合计算仅支持INTEGER、LONG、FLOAT、DOUBLE类型的字段。
- 类型：STRING
- 必选：是
- 约束：必须为SUM、AVERAGE之一。

高级配置项

并行数：

- 配置项英文名：parallelism

- 说明：并行数，指定算子执行时的并行任务个数。对于性能要求越高的算子，可以设置较高的并行数，但是占用资源越多。如果此参数不设置，那么作业运行时本算子使用作业的并行数设置。
- 类型：INTEGER
- 必选：否
- 约束：最小值1，最大值 1024。

4.5.4 表达式说明

操作符	说明	示例
&&	逻辑与	设：cond1 = true, cond2 = false, 表达式：cond1 && cond2, 返回 false。
and	逻辑与	设：cond1 = true, cond2 = false, 表达式：cond1 and cond2, 返回 false。
	逻辑或	设：cond1 = true, cond2 = false, 表达式：cond1 cond2, 返回 true。
or	逻辑或	设：cond1 = true, cond2 = false, 表达式：cond1 or cond2, 返回 true。
!	逻辑非	设：cond1 = true, 表达式：!cond1, 返回false。
not	逻辑非	设：cond1 = true, 表达式：not cond1, 返回false。
&	二进制位与	表达式：33 & 4, 会将33与4转换成2进制，并按位进行与运算：0010 0001 & 0000 0100 = 0。
	二进制位或	表达式：33 4, 会将33与4转换成2进制，并按位进行或运算：0010 0001 0000 0100 = 0010 0101 = 37。
^	二进制位异或	表达式：33 ^ 4, 会将33与4转换成2进制，并按位进行异或运算：0010 0001 ^ 0000 0100 = 0010 0101 = 37。
~	二进制位取反	表达式：~33, 会将33转换成2进制，并按位进行二进制取反运算：~0010 0001 = 1101 1110 = -34
?:	三元运算符	把两个结果中其中一个符合运算逻辑的值返回。如：condition ? ifTrue : ifFalse, 当条件condition为true时，返回ifTrue，否则返回ifFalse。
=	等于	设：val1 = 1, val2 = 1, val3 = 2, 则：表达式val1 == val2返回 true, 表达式val1 == val3返回false。
eq	等于	设：val1 = 1, val2 = 1, val3 = 2, 则：表达式val1 eq val2返回 true, 表达式val1 eq val3返回false。
!=	不等于	设：val1 = 1, val2 = 1, val3 = 2, 则：表达式val1 != val2返回 false, 表达式val1 != val3返回true。

操作符	说明	示例
ne	不等于	设: val1 = 1, val2 = 1, val3 = 2, 则: 表达式val1 ne val2返回false, 表达式val1 ne val3返回true。
<	小于	设: val1 = 1, val2 = 1, val3 = 2, 则: 表达式val1 < val2返回false, 表达式val1 < val3返回true。
lt	小于	设: val1 = 1, val2 = 1, val3 = 2, 则: 表达式val1 lt val2返回false, 表达式val1 lt val3返回true。
<=	小于等于	设: val1 = 1, val2 = 1, val3 = 2, 则: 表达式val1 <= val2返回true, 表达式val1 <= val3返回true。
le	小于等于	设: val1 = 1, val2 = 1, val3 = 2, 则: 表达式val1 le val2返回true, 表达式val1 le val3返回true。
>	大于	设: val1 = 1, val2 = 1, val3 = 2, 则: 表达式val2 > val1返回false, 表达式val3 > val1返回true。
gt	大于	设: val1 = 1, val2 = 1, val3 = 2, 则: 表达式val2 gt val1返回false, 表达式val3 gt val1返回true。
>=	大于等于	设: val1 = 1, val2 = 1, val3 = 2, 则: 表达式val2 >= val1返回true, 表达式val3 >= val1返回true。
ge	大于等于	设: val1 = 1, val2 = 1, val3 = 2, 则: 表达式val2 ge val1返回true, 表达式val3 ge val1返回true。
=~	匹配	使用表达式表示"abcdef"匹配正则表达式"abc.*": "abcdef" =~ "abc.*", 返回true。
!~	不匹配	使用表达式表示"abcdef"不匹配正则表达式"abc.*": "abcdef" !~ "abc.*", 返回false。
=^	以某字符串开头	使用表达式表示"abcdef"以"abc"开头: "abcdef" ^= "abc", 返回true。
!^	不以某字符串开头	使用表达式表示"abcdef"以"abc"开头: "abcdef" !^ "abc", 返回false。
= \$	以某字符串结尾	使用表达式表示"abcdef"以"def"结尾: "abcdef" =\$ "def", 返回true。
!\$	不以某字符串结尾	使用表达式表示"abcdef"以"def"结尾: "abcdef" !\$ "def", 返回false。
+	加法	设: val1 = 1, val2 = 2, 则: 表达式val2 + val1返回3。
-	减法	设: val1 = 1, val2 = 2, 则: 表达式val2 - val1返回1。
*	乘法	设: val1 = 1, val2 = 2, 则: 表达式val2 * val1返回2。

操作符	说明	示例
/	除法	设: val1 = 1, val2 = 2, 则: 表达式val2 / val1返回2。
div	除法	设: val1 = 1, val2 = 2, 则: 表达式val2 div val1返回2。
%	取模	设: val1 = 1, val2 = 2, 则: 表达式val2 % val1返回0。
mod	取模	设: val1 = 1, val2 = 2, 则: 表达式val2 mod val1返回0。
+=	加等于	设: val = 1, 则表达式: val += 1将val赋值为2。
-=	减等于	设: val = 1, 则表达式: val -= 1将val赋值为0。
*=	乘等于	设: val = 1, 则表达式: val *= 1将val赋值为1。
/=	赋值商	设: val = 1, 则表达式: val /= 1将val赋值为1。
%=	赋值求余	设: val = 1, 则表达式: val %= 1将val赋值为0。
&=	赋值与	设: val = 33, 则表达式: val &= 4将val赋值为0。
=	赋值或	设: val = 33, 则表达式: val = 4将val赋值为37。
^=	赋值异或	设: val = 33, 则表达式: val ^= 4将val赋值为37。
-	取负数	设: val = 1, 则表达式: -val返回-1。

5 存储管理

IoT数据分析服务的存储管理统一展示了 IoT 数据分析服务上所有保存数据的存储，包括资产模型存储和由数据管道创建的存储，并支持对这些存储中的数据配置保存时间。

5.1 存储

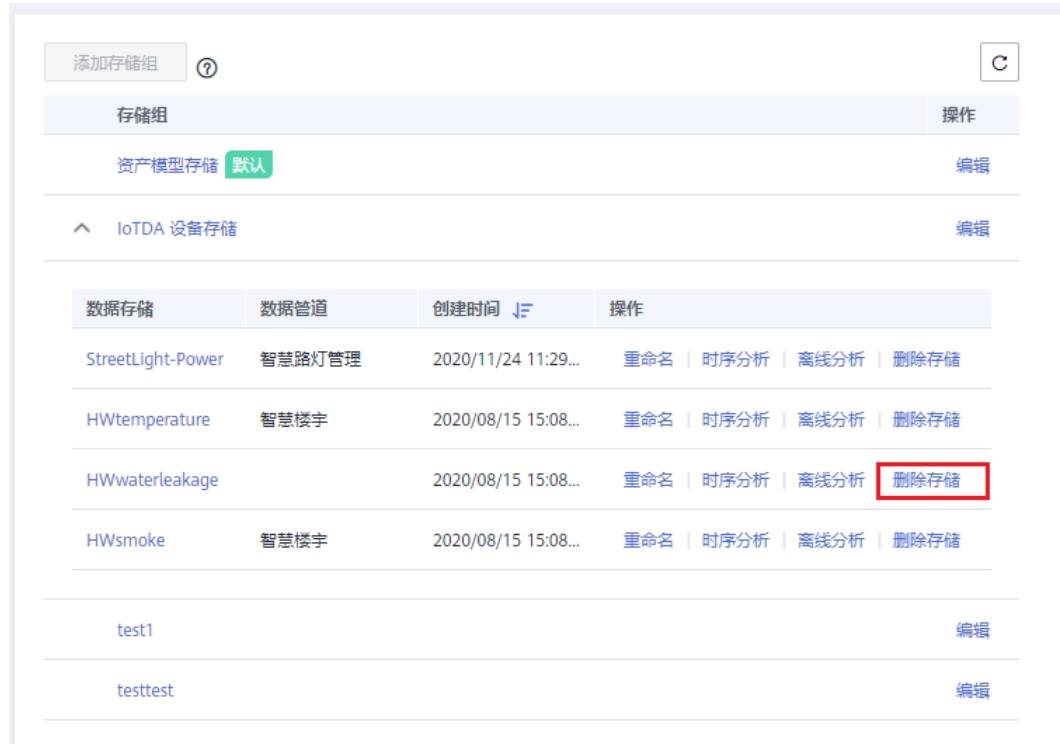
存储是保存数据的对象，通常来说，同一种设备的数据会保存在同一个存储中。存储中定义了设备的标签与设备的属性。其中设备ID（deviceid）就是最常用的一种标签（tag）。一个标签包括标签名与标签值，例如设备ID：标签名为 deviceid，而它的标签值则是具体的某个设备ID，设备的ID是通过设备数据上报到IoT数据分析服务而保存下来的。在查询中指定标签名与某个标签值（例如具体的某个设备ID），可以快速地找到对应的设备数据。

5.1.1 创建存储

在数据管道中配置存储输出算子时，如果未选择已有存储，则将创建新的存储。在数据管道中通过存储输出算子创建的存储将会展示在存储管理中的对应的存储组之下。数据管道的数据存储输出算子配置详见[数据存储输出](#)。

5.1.2 删除存储

删除存储前，需要先在数据管道删除对应的存储输出算子，或者修改对应的存储输出算子使其输出到其它存储中，然后点击存储管理页面上的“删除存储”按钮。



5.2 存储组

存储组可以将多个存储按照相同的数据保存策略（数据保存时间）统一管理。一个存储组内可以包含多个存储，一个存储只能属于一个存储组并且不可修改为别的存储组。

说明

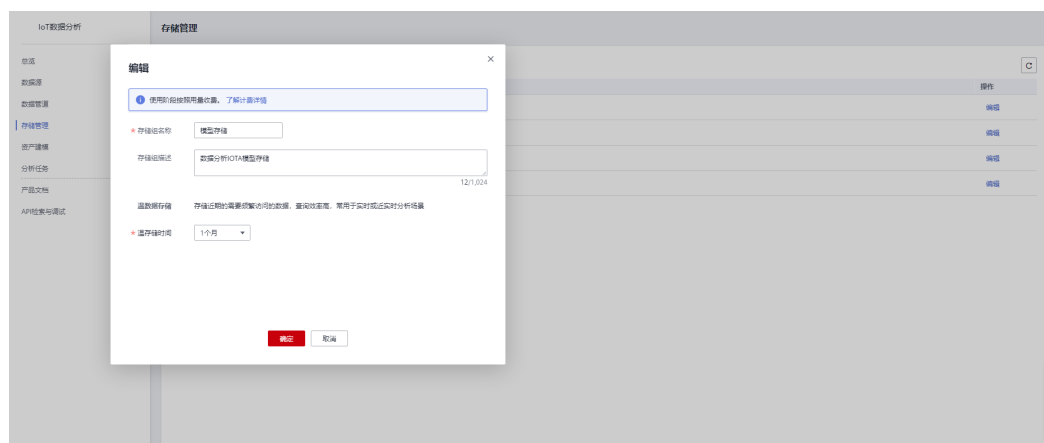
IoT数据分析服务共享版限制每个租户只有1个存储组，即为默认存储组。

5.2.1 修改数据保存时间

点击存储组上的编辑按钮，可以调整存储组下的所有存储中的数据保存时间。如下图：

添加存储组 ? C

存储组	操作		
资产模型存储 默认	编辑		
^ IoTDA 设备存储 编辑			
数据存储	数据管道	创建时间	操作
StreetLight-Power	智慧路灯管理	2020/11/24 11:29...	重命名 时序分析 离线分析 删除存储
HWtemperature	智慧楼宇	2020/08/15 15:08...	重命名 时序分析 离线分析 删除存储
HWwaterleakage	智慧楼宇	2020/08/15 15:08...	重命名 时序分析 离线分析 删除存储
HWsmoke	智慧楼宇	2020/08/15 15:08...	重命名 时序分析 离线分析 删除存储
test1			编辑
testtest			编辑



随着时间流逝，当存储中的数据时间超出了保存时间，超出保存时间范围内的数据将会被自动清除。

6 资产建模

6.1 简介

6.1.1 基本概念介绍

为了构建物联网领域业务对象的数字孪生模型，IoT数据分析服务参考软件工程中面向对象的思想提供了“资产建模”能力，其中核心概念包括资产模型和资产，下面将对此概念进行介绍。

资产模型 (AssetModel)

资产模型是对物理世界中同类事物的抽象定义，资产模型中包括属性和分析任务。对物理世界事物构建数字资产模型时，必须先定义好资产模型、然后再创建资产。

属性 (Property)

属性是指物理世界事物的状态数据，属性有3种类型，包括静态配置、测量数据和分析任务。

- **静态配置类属性**：是指一旦配置后就不变化或者变化不频繁的数据，如设备SN号等。
- **测量数据类属性**：是指由传感器、仪器仪表等设备测量所得的数据，如温度、湿度、电流、电压等。
- **分析任务类属性**：是指由IoT数据分析服务中用户所配置的分析任务所计算得到的数据，如求平均分析任务计算得到平均温度等。

分析任务 (Analysis)

分析任务是指对属性数据进行计算分析的任务，分析任务类型有3种，包括转换计算、聚合计算、流计算。

- **转换计算类分析任务**是指，当参数所引用的属性值发生变化时，对单个或多个属性值进行常规数学表达式计算，表达式中不能使用时间聚合公式。

例子：(a+b)/2;

- 聚合计算类分析任务是指，对资产的属性值进行周期性的时间聚合计算或空间聚合计算，表达式中可以使用时间聚合算子，其中空间聚合计算是指多个资产属性时间聚合算子表达式再进行联合预算（此时各属性时间聚合计算的时间范围和周期相同），如下例子所示。

例子：StateTime(machineState, 100) + StateTime(machineState, 101)，时间范围: 15min, 周期: 15min;

- 流计算类分析任务是指，以资产模型中的属性作为输入、调用IoT数据分析服务的实时分析作业进行计算、输出结果返回到资产模型的属性值，数据流示意图如下。



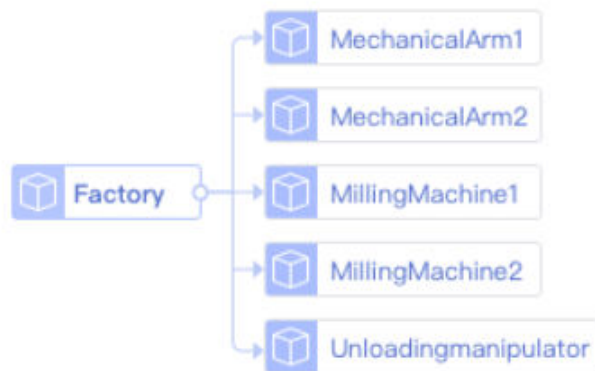
资产 (Asset)

资产是指物理世界中需要被管理的、具有唯一标识的事物或业务对象，如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）。在资产建模时，必须利用资产模型来创建资产，资产中的属性、分析任务均与其所属资产模型保持一致。

例子，如下图所示为现实世界中一个工厂的示意图



该工厂例子中有工厂、机械臂、铣床、卸货机械手四类物理对象，因此需构建 Factory、MechanicalArm、MillingMachine、Unloadingmanipulator 四个资产模型，然后利用各模型创建相应资产，并根据现实世界中资产间拓扑关系建立数字世界中资产树拓扑关系，得到如下图所示的数字资产模型



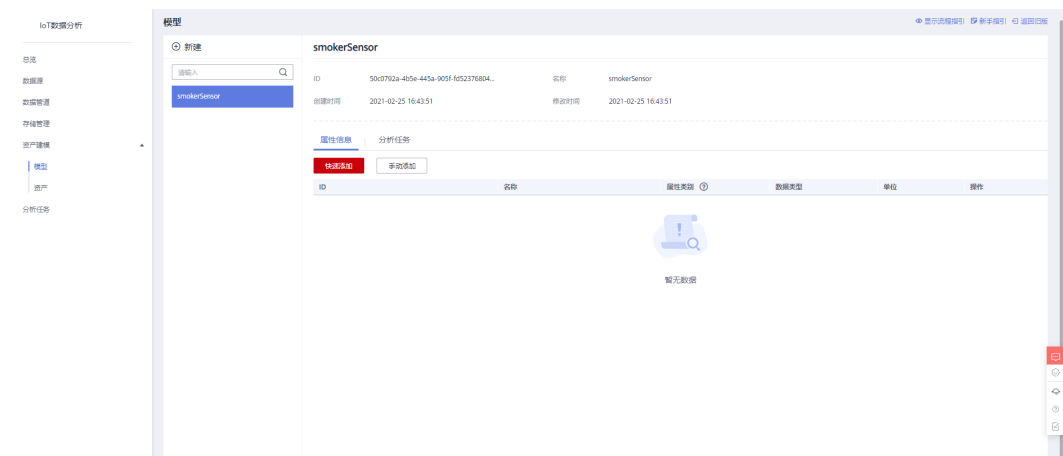
6.1.2 主要功能简介

资产建模主要有2部分功能：模型管理、资产管理。

模型管理

在资产模型管理页，您可以对资产模型进行创建、修改、删除、查询等操作。

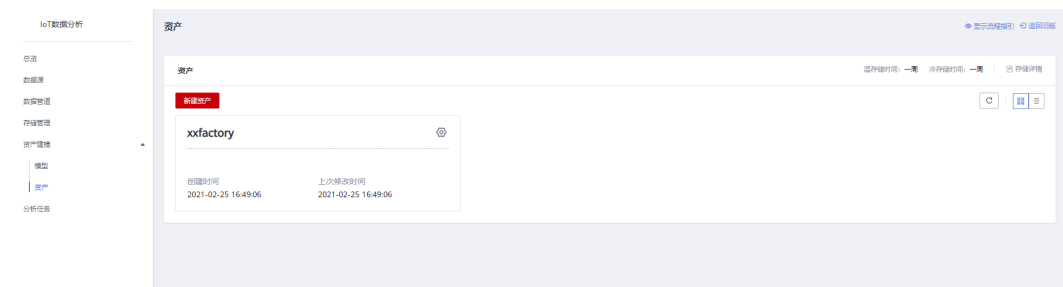
图 6-1 模型管理



资产管理

在资产管理页，您可以对资产进行创建、修改、删除、查询等操作。

图 6-2 资产管理

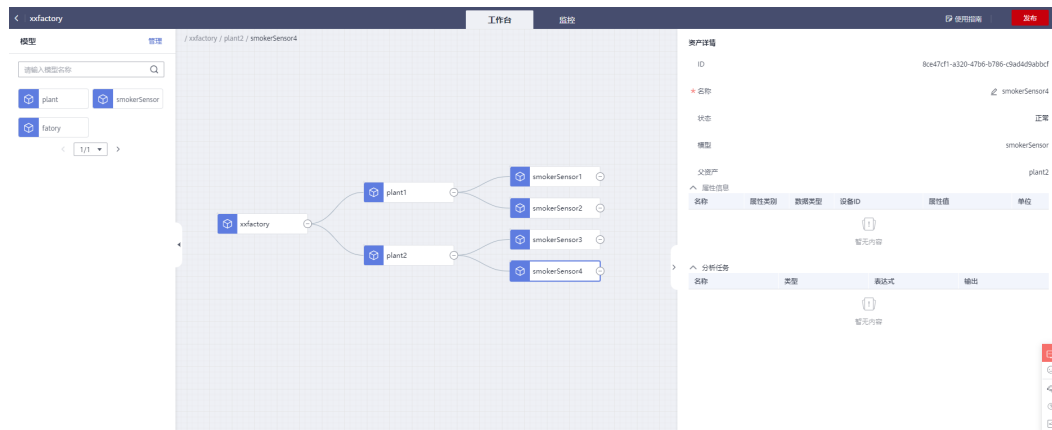


资产管理有工作台和监控两个工具，工作台用于构建资产的草稿态，在草稿态下可对资产进行修改编辑；当构建完成后，将工作台中的草稿态资产发布运行，即可在监控台中进行实时监控，监控台下的资产不能修改编辑。

工作台

在资产工作台页，您可以对资产进行更详细的操作，包括创建子资产和移动资产（即对资产间拓扑关系进行建模）、修改资产详情和删除子资产等操作。

图 6-3 资产工作台



监控

在资产监控页，您可以查看已发布的资产进行实时监控，比如资产中属性值将会根据设备上上报数据进行刷新、或根据分析任务计算结果进行刷新。

图 6-4 资产监控



6.2 模型管理

6.2.1 创建资产模型

步骤1 点击“新建”，如下图1中①所示

步骤2 输入模型名称、显示名称，点击确定，如下图1中②所示

图 6-5 新增模型

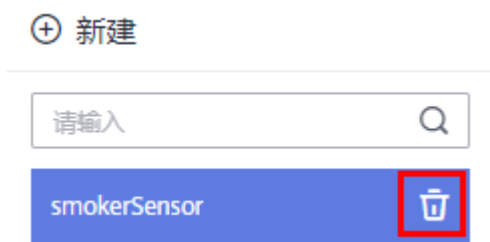


----结束

6.2.2 删除资产模型

步骤1 点击所需删除资产模型右边的删除图标，如下图1所示

图 6-6 删除资产模型



----结束

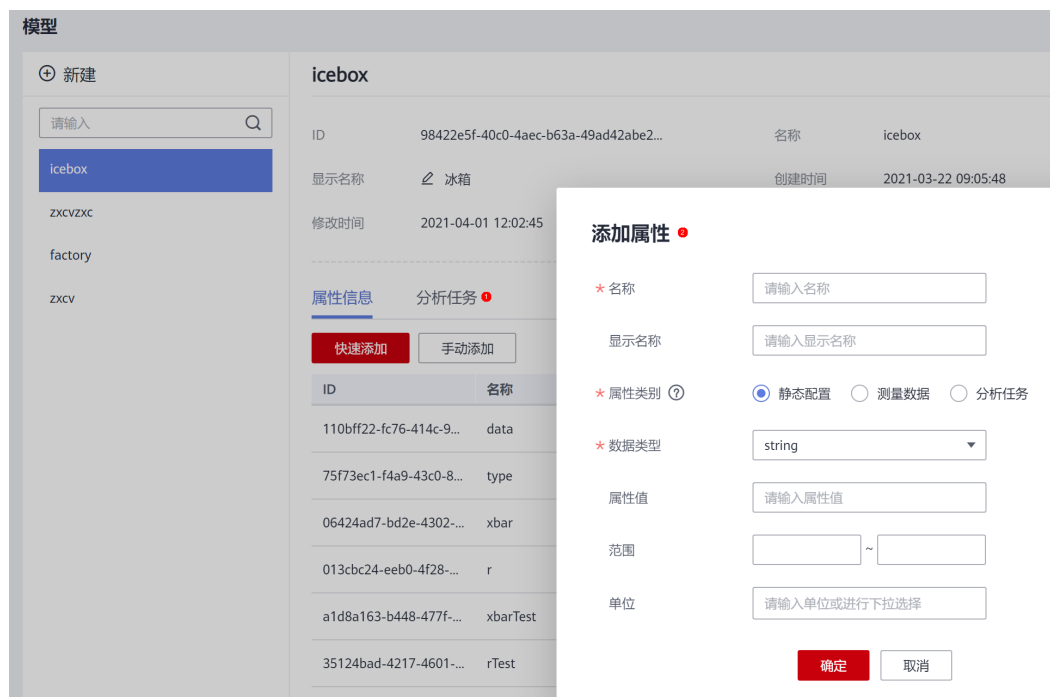
6.2.3 添加属性信息

手动添加

步骤1 点击手动添加，如下图1中①所示

步骤2 输入名称、显示名称、选择属性类别、数据类型、属性值（指属性的默认值）、范围和单位，点击确定，如下图1中②所示；在单位定义时，系统已预置一系列常用单位，若无合适单位、用户可手动输入自定义单位。

图 6-7 手动增加属性信息



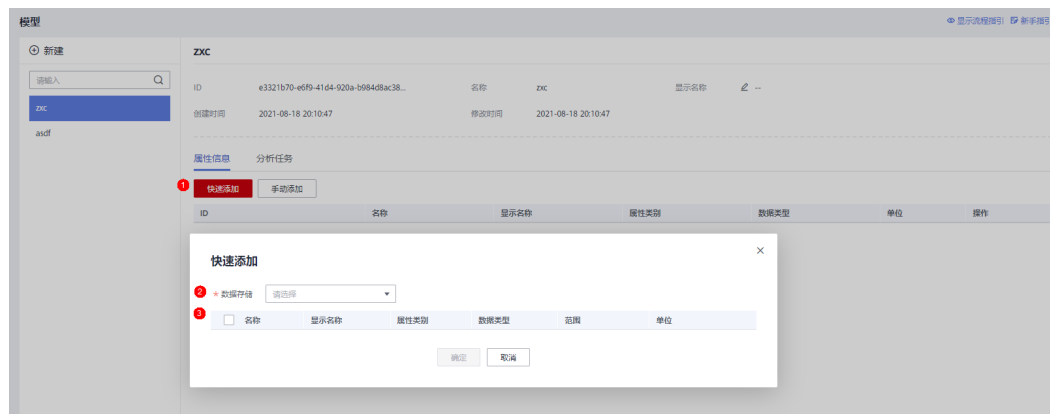
----结束

快速添加

IoT数据分析服务提供了快速批量添加属性的功能，用户可以将存储在存储管理中定义的物的各种属性快速引入到资产模型中，引用时需要指定存储位置，具体步骤如下：

- 步骤1** 点击快速添加，如下图2中①所示
- 步骤2** 选择数据存储，如下图2中②所示
- 步骤3** 输入范围，勾选要添加的属性，点击确定，如下图2中③所示。

图 6-8 快速添加



----结束

6.2.4 修改属性信息

修改属性信息后，修改的内容将自动同步到该模型未发布的资产上（工作台中的资产为未发布资产），已发布的资产不受影响。

步骤1 选择要修改的属性，点击编辑，如下图1中①所示。

步骤2 修改属性的显示名称。

步骤3 修改属性类别，只有测量数据和分析任务类别可以相互切换，静态配置类别不能切换为其他类别。

步骤4 修改范围，只有integer、double、string数据类型可以修改范围，object类型无范围。

步骤5 修改属性的单位。

步骤6 修改属性值，只有静态配置类别的属性可以修改（该属性值为默认值，如果对应资产上配置过属性值则资产属性值保持不变，如果未配置过则会刷新成该默认值），如下图2中①所示。

图 6-9 编辑测量数据类属性

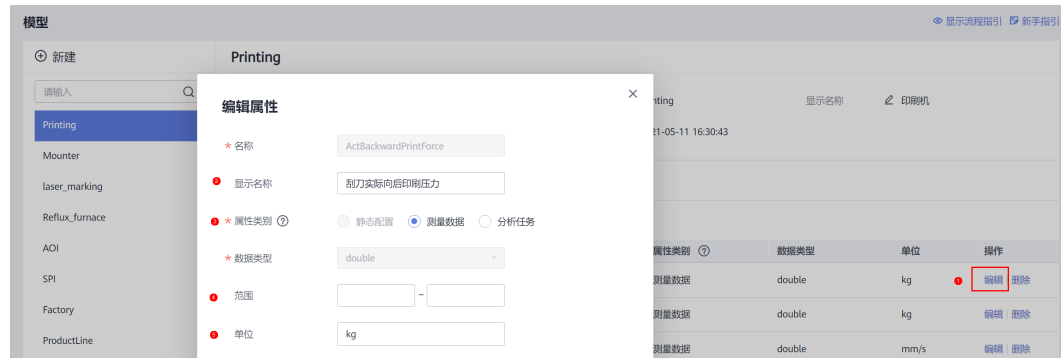


图 6-10 编辑静态配置类属性

编辑属性

* 名称	PlanningOutputPeriod
显示名称	设备理想产品生产时长
* 属性类别 ?	<input checked="" type="radio"/> 静态配置 <input type="radio"/> 测量数据 <input type="radio"/> 分析任务
* 数据类型	double
1 属性值	0.06
2 范围	<input type="text"/> ~ <input type="text"/>
单位	min

----结束

注意

修改属性信息可能导致该模型对应的资产处于异常状态，资产异常状态的场景见[资产状态](#)。

6.2.5 删除属性信息

删除属性信息后，将自动同步到该模型未发布的资产上，已发布的资产不受影响。

步骤1 选择要删除的属性，点击删除。

图 6-11 删除属性信息

plant																								
ID	ccf84221-7b51-4ec6-8bda-749c5fed0e12	名称	plant																					
创建时间	2021-02-25 16:52:11	修改时间	2021-02-27 10:09:00																					
<div style="display: flex; justify-content: space-between;"> 属性信息 分析任务 </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 快速添加 手动添加 </div> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>ID</th> <th>名称</th> <th>属性类别</th> <th>数据类型</th> <th>单位</th> <th>操作</th> </tr> </thead> <tbody> <tr> <td>427e60a7-282f-472f-a748-e62b0fb2e2a7</td> <td>test</td> <td>分析任务</td> <td>Integer</td> <td>NA</td> <td>编辑 删除</td> </tr> <tr> <td>ad4bd618-c8cb-490e-991f-ae364dc20293</td> <td>test2</td> <td>静态配置</td> <td>string</td> <td>NA</td> <td>编辑 删除</td> </tr> </tbody> </table>							ID	名称	属性类别	数据类型	单位	操作	427e60a7-282f-472f-a748-e62b0fb2e2a7	test	分析任务	Integer	NA	编辑 删除	ad4bd618-c8cb-490e-991f-ae364dc20293	test2	静态配置	string	NA	编辑 删除
ID	名称	属性类别	数据类型	单位	操作																			
427e60a7-282f-472f-a748-e62b0fb2e2a7	test	分析任务	Integer	NA	编辑 删除																			
ad4bd618-c8cb-490e-991f-ae364dc20293	test2	静态配置	string	NA	编辑 删除																			

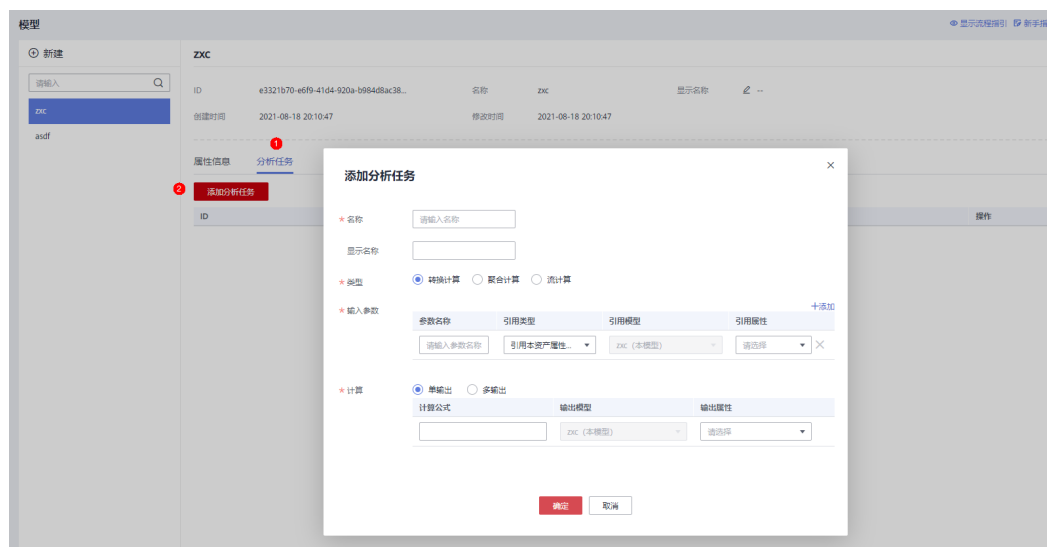
----结束

6.2.6 添加分析任务

步骤1 点击添加分析任务，如下图1中①所示

步骤2 输入名称，选择类型，完成输入参数、表达式的定义，并把分析任务的输出映射到具体所需的属性后点击确定即可，如下图1中②所示；分析任务概念详细介绍请见[分析任务定义](#)。

图 6-12 添加分析任务



----结束

6.2.7 修改分析任务

修改分析任务后，修改的内容将自动同步到该模型未发布的资产上，已发布的资产不受影响。

步骤1 选择要修改的分析任务，点击编辑。

步骤2 修改类型，类型可以任意切换。

步骤3 修改输入参数，转换计算和聚合计算类型可以添加和修改参数，流计算类型由于参数是实时分析作业定义的，只能修改属性引用类型和选择属性名称。

- 步骤4** 修改公式，可以重新定义公式。
- 步骤5** 修改输出属性，可以重新选择输出属性。
- 步骤6** 修改时间范围，只有聚合计算类型的分析任务可以修改。
- 步骤7** 修改周期，只有聚合计算类型的分析任务可以修改。
- 步骤8** 修改流计算任务，只有流计算类型的分析任务可以修改。

图 6-13 修改“转换计算”分析任务

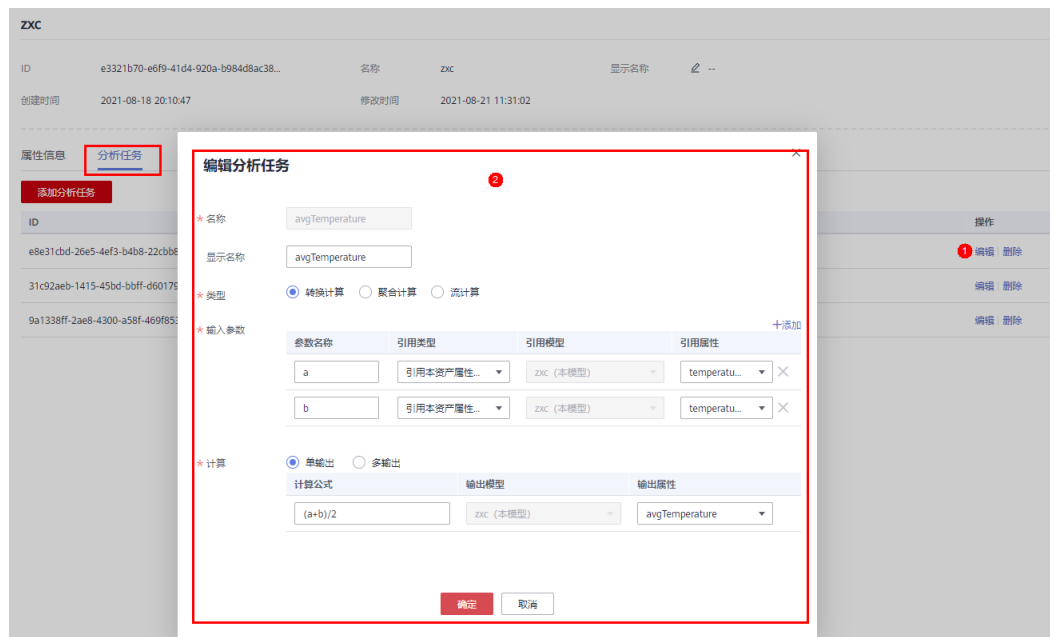


图 6-14 修改“聚合计算”分析任务

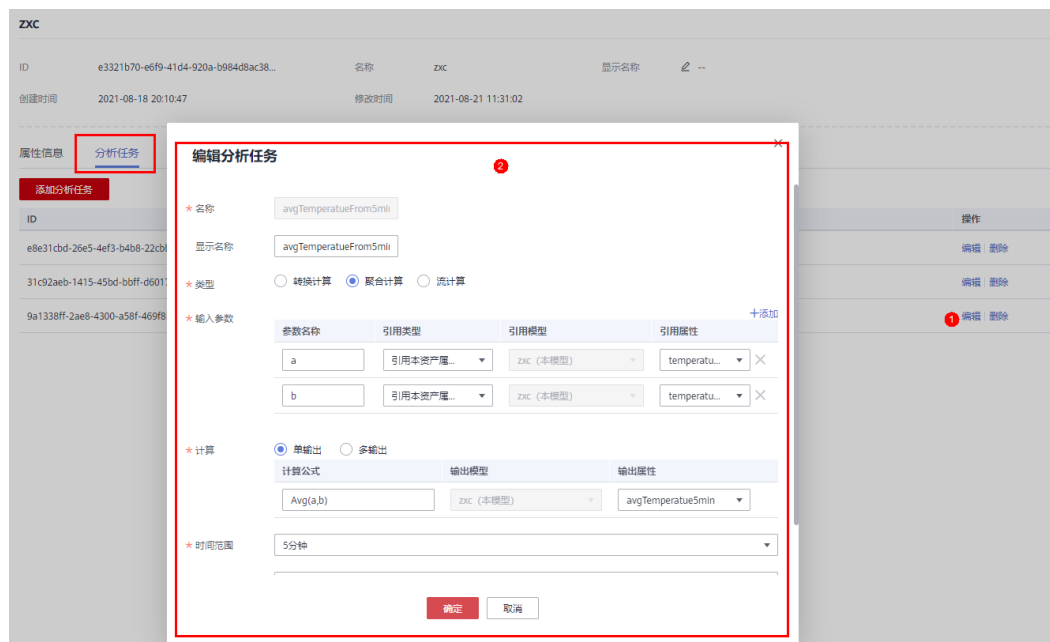
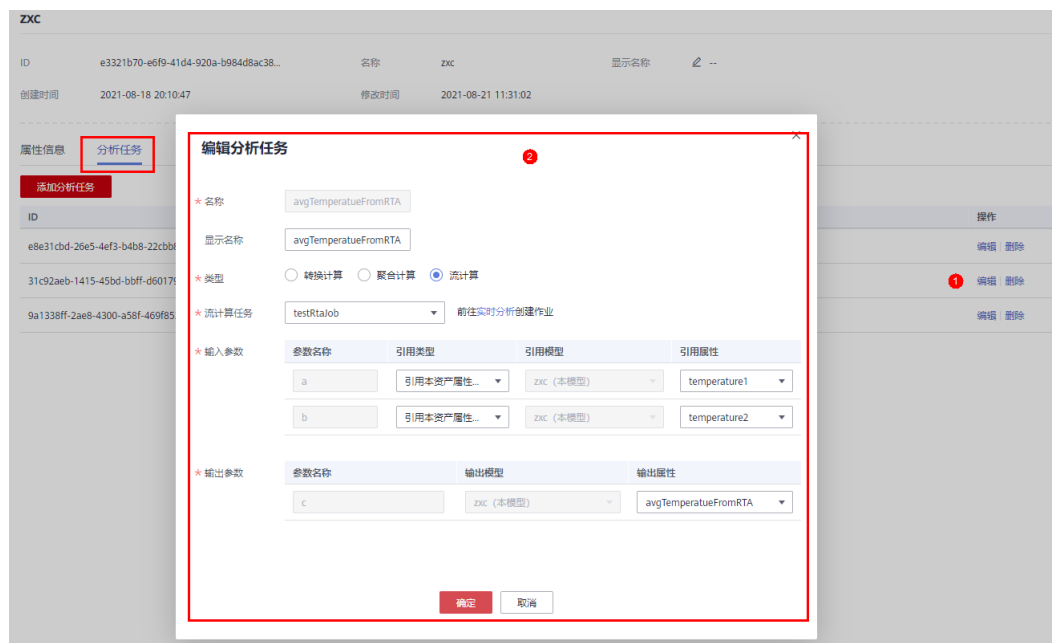


图 6-15 修改“流计算”分析任务



----结束

注意

修改分析任务可能导致该模型对应的资产处于异常状态，资产异常状态的场景见[资产状态](#)。

6.2.8 删除分析任务

删除分析任务后，将自动同步到该模型未发布的资产上，已发布的资产不受影响。

步骤1 选择要删除的分析任务，点击删除。

图 6-16 删除分析任务



----结束

6.3 资产管理

6.3.1 创建资产

步骤1 在资产管理页，点击“新建资产”，如下图1中①所示

步骤2 输入名称，选择模型，点击确定，如下图1中②所示

图 6-17 新建资产

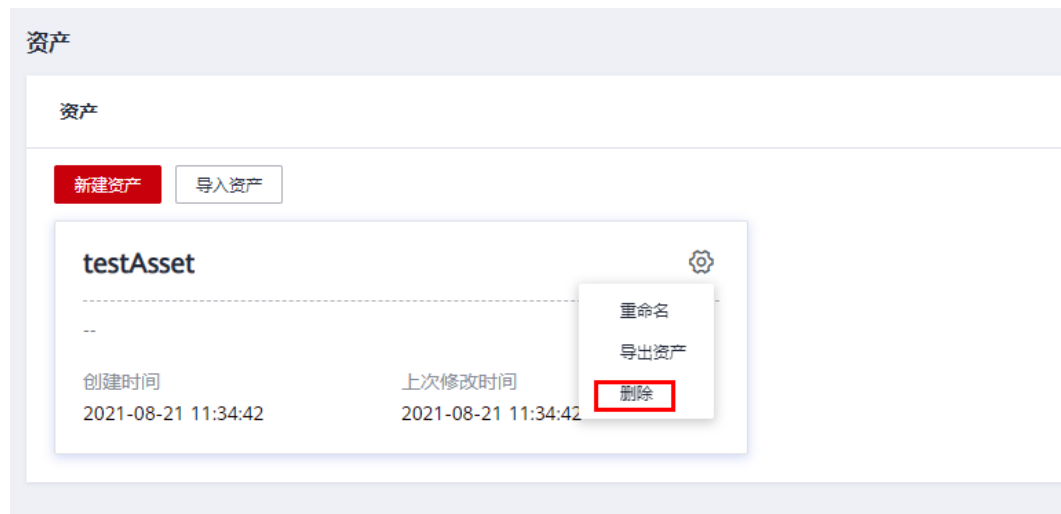


----结束

6.3.2 删除资产

步骤1 在资产管理页，点击设置图标，点击删除，如下图1所示

图 6-18 删除资产



----结束

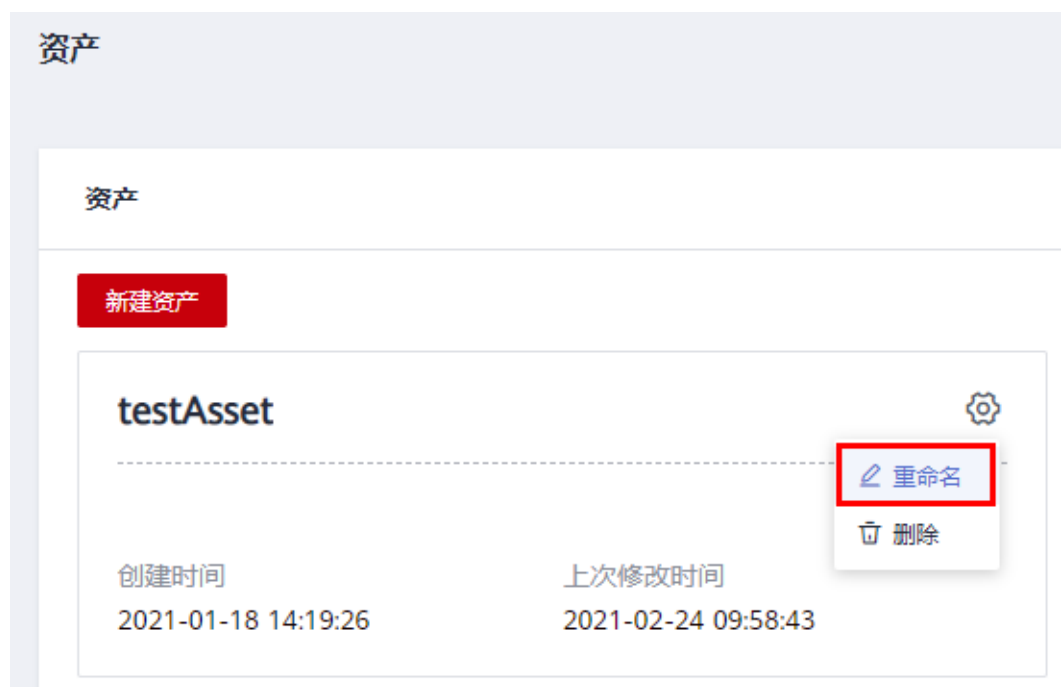
⚠ 注意

删除资产操作，会把该资产本身、以及其下所有子资产一并删除，需谨慎操作。

6.3.3 修改名称

步骤1 在资产管理页，点击设置图标，点击重命名，如下图1所示

图 6-19 修改名称



----结束

6.3.4 工作台

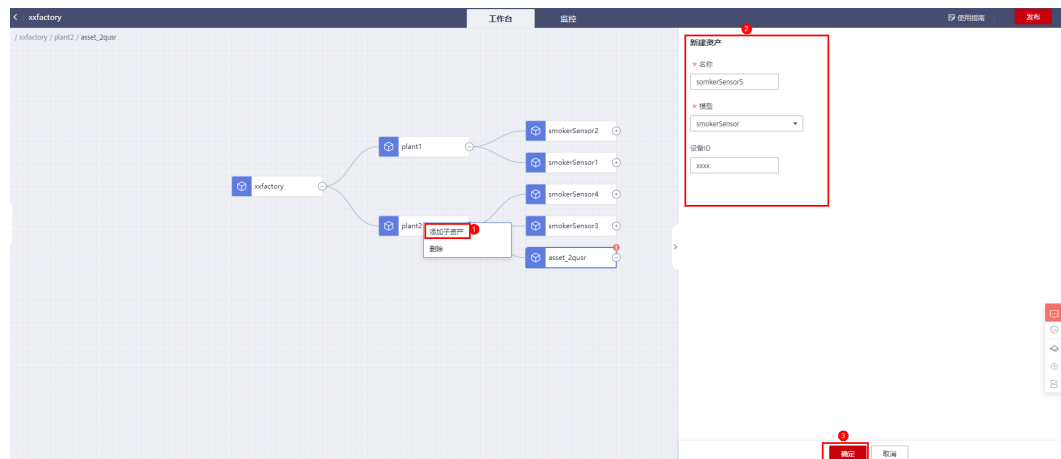
6.3.4.1 创建子资产

右键添加

步骤1 在资产工作台页，选择任一资产，点击鼠标右键，选择添加子资产

步骤2 输入名称、设备ID，选择模型，点击确定

图 6-20 添加子资产



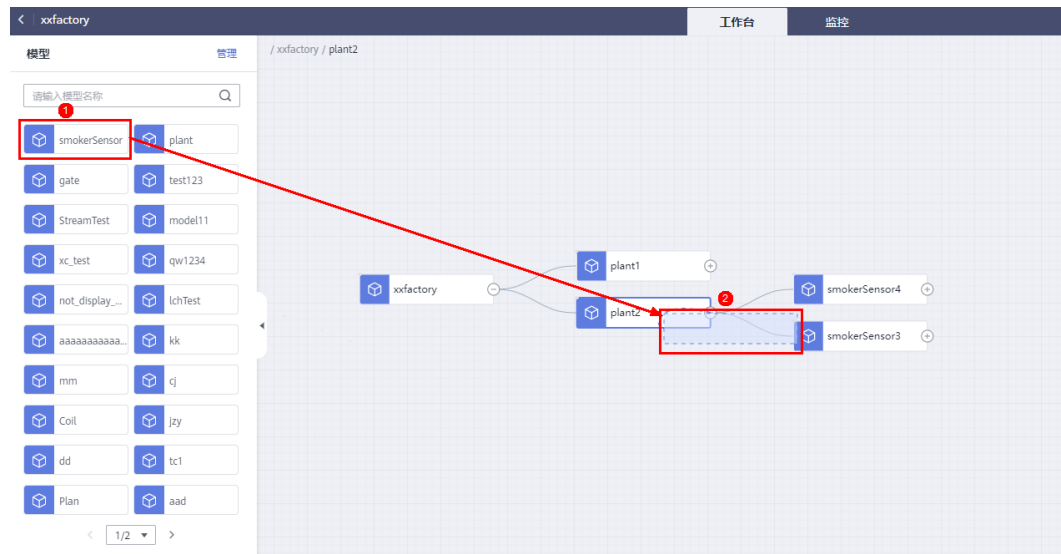
----结束

拖动添加

步骤1 在资产工作台页，选中左侧模型列表中的一个模型进行拖动，入下图2中①所示

步骤2 拖动到目标父资产上释放鼠标，后续流程同手动添加，如下图2中②所示

图 6-21 拖动添加

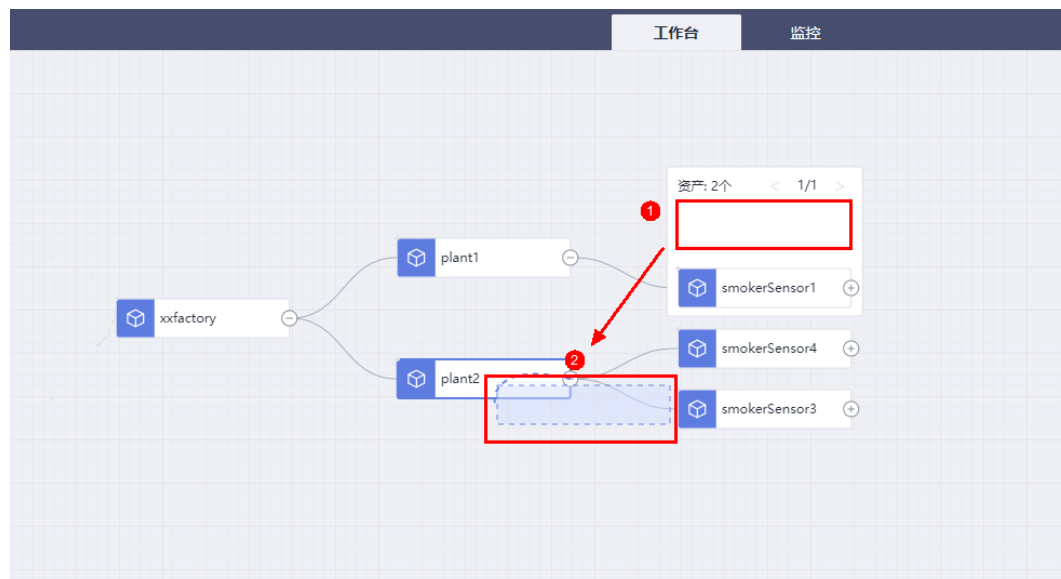


----结束

6.3.4.2 移动子资产

步骤1 选中一个资产，鼠标拖动到目标父资产上释放鼠标

图 6-22 移动子资产

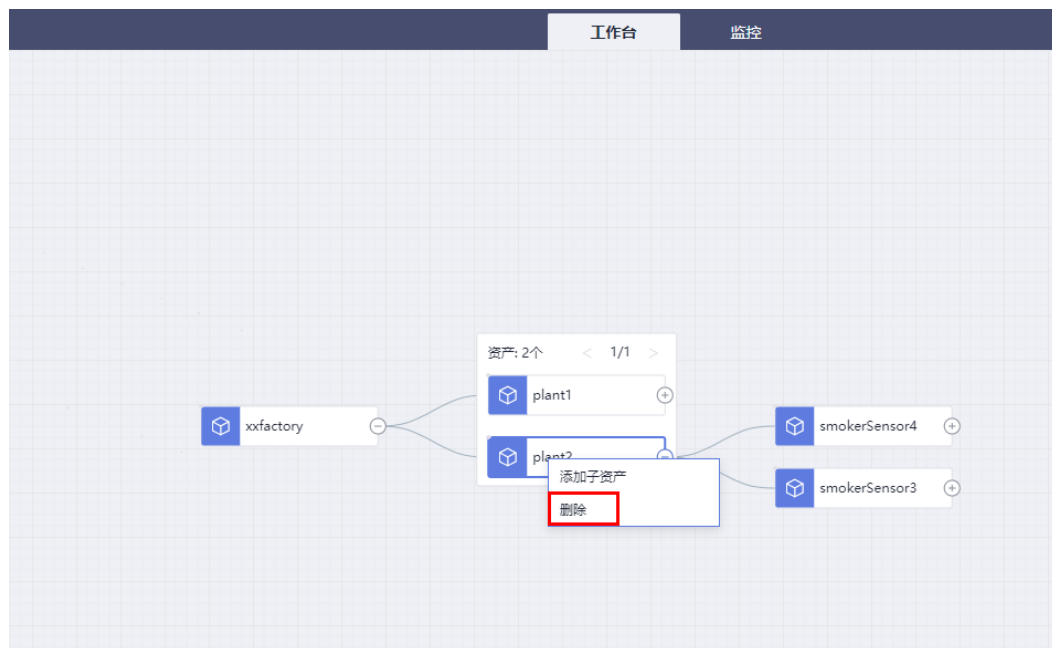


----结束

6.3.4.3 删除子资产

步骤1 选中子资产，点击删除。

图 6-23 删除子资产

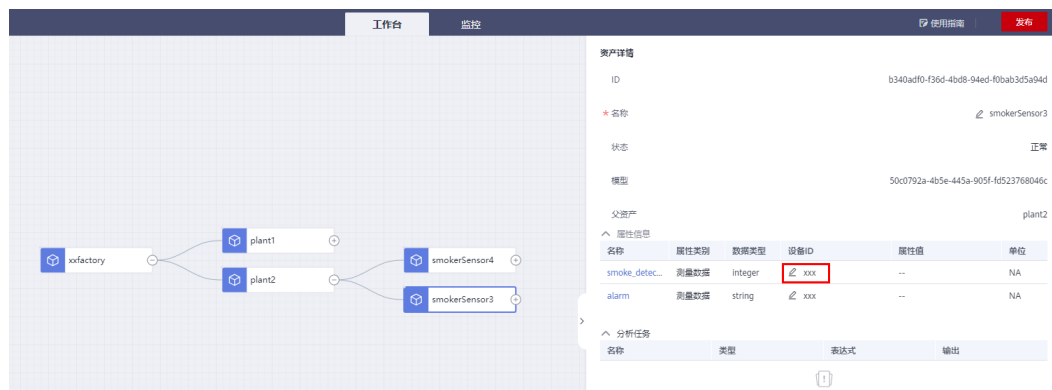


----结束

6.3.4.4 配置设备 ID

步骤1 点击目标属性“设备ID”字段下的修改图标进行修改（只有测量数据类型属性需要配置设备ID），在该资产发布到监控台后，当指定的设备ID上报消息后该属性值即会发生刷新。

图 6-24 配置设备 ID



----结束

6.3.4.5 配置静态值

步骤1 点击目标属性属性值的修改图标进行修改（只有静态配置类型的属性需要配置静态值）。

图 6-25 配置静态值



----结束

6.3.4.6 配置分析任务输入

步骤1 点击目标分析任务，如下图1所示。

图 6-26 目标分析任务



步骤2 点击资产ID的修改图标进行修改，如下图2所示（只有属性引用类型为“引用其他资产属性”的输入参数需要配置资产ID）。

图 6-27 资产 ID



----结束

6.3.4.7 资产状态

未发布的资产有正常（ACTIVE）和异常（INACTIVE）状态，已发布的资产没有资产状态。

只有当一棵资产树中所有资产都处于正常状态时才可以在工作台发布资产。

资产存在以下场景之一则处于异常状态：

- 静态配置类别的属性未配置属性值，如下图1①所示。
- 测量数据类别的属性未配置设备ID，如下图1②所示。
- 分析任务中存在“引用其他资产属性”的输入参数，但未配置引用的资产ID，如下图2所示。

图 6-28 资产详情



图 6-29 分析任务详情



6.3.5 监控

步骤1 当所有资产都处于正常状态时，点击发布按钮发布资产，如下图1所示。

图 6-30 发布



步骤2 在资产监控页面，选中目标资产，在右侧详情面板中点击“刷新”，即可查看资产属性的最新值，如下图2所示。


步骤3 点击图标可进入时序分析页面查看资产属性的时序值（只有属性类别为测量数据和分析任务的属性才有时序值）。

图 6-31 监控页面

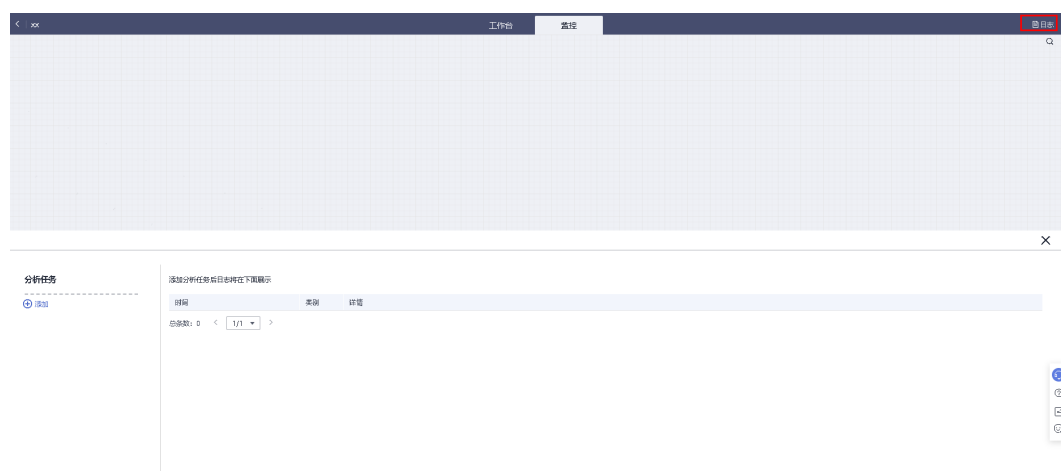


----结束

6.3.6 日志

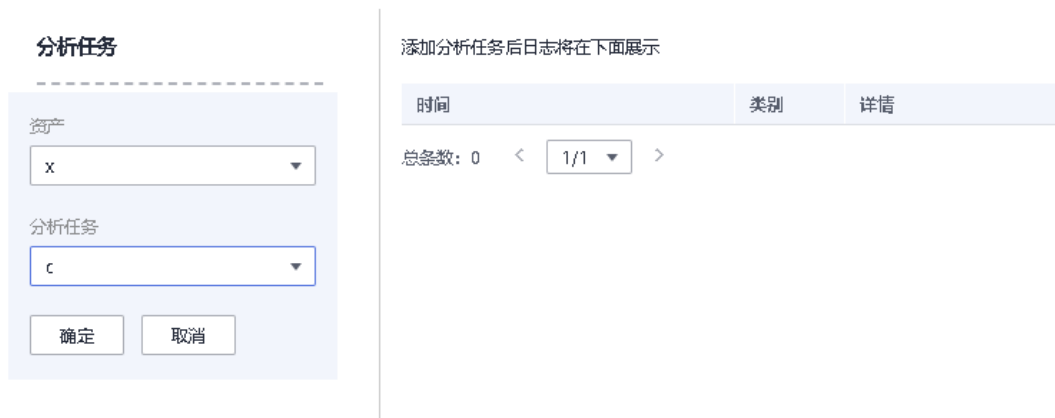
步骤1 在资产监控页面可以观察分析任务执行日志，点击日志按钮打开日志页面，如下图1所示。

图 6-32 日志管理



步骤2 在日志页面点击添加，选择资产和分析任务，再点击确定，如下图2所示。

图 6-33 添加分析任务



步骤3 点击右侧刷新按钮可以查看到最新的执行日志，如下图3所示。

图 6-34 查看执行日志



---结束

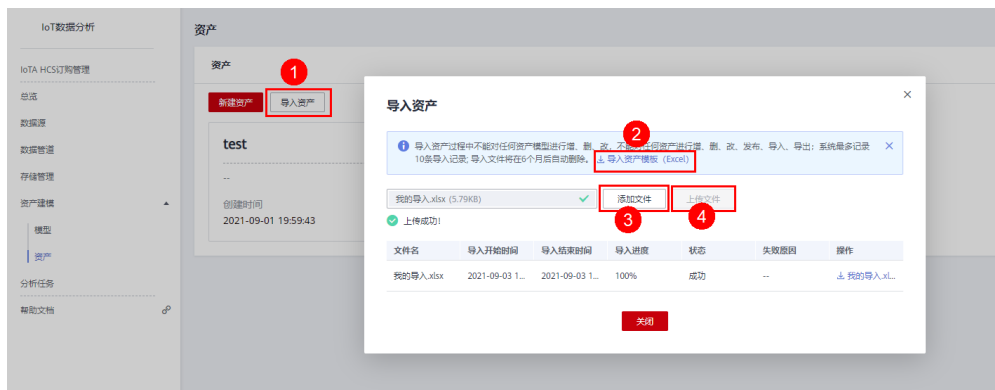
6.3.7 导入资产

步骤1 在资产管理页，点击“导入资产”，如下图1中①所示

步骤2 下载导入资产模板，如下图1中②所示


步骤3 在本地编辑好导入资产文件，点击“添加文件”选择该文件，如下图1中③所示

步骤4 点击“上传文件”导入资产，如下图1中④所示



----结束

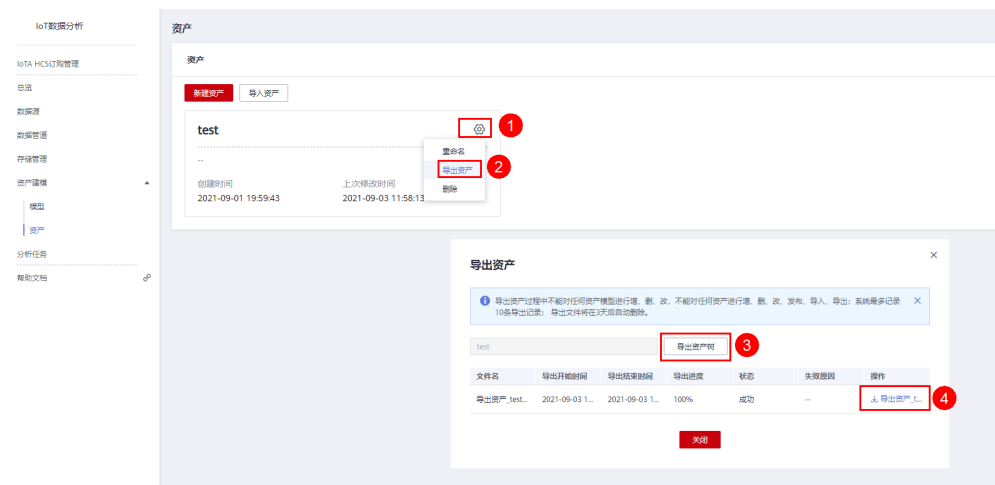
6.3.8 导出资产

步骤1 在资产管理页，选中一个资产点击  按钮，如下图所示图1中①所示

步骤2 单击“导出资产”，如下图所示图1中②所示

步骤3 单击“导出资产树”导出资产，如下图所示图1中③所示

步骤4 等待导出完毕，单击导出的文件进行下载，如下图所示图1中④所示



----结束

6.4 分析任务定义

6.4.1 简介

在IoT数据分析服务资产模型中，可以通过“分析任务”对属性值进行相互运算，并将结果输出到某个属性。

“分析任务”包含三个部分

- 输入参数

- 计算公式
- 输出属性

下面将对“分析任务”进行详细介绍。

6.4.2 分析任务类型

分析任务有三种类型，“转换计算”和“聚合计算”这两种类型的计算公式在资产建模模块进行定义，而“流分析”类分析任务的计算过程通过编辑对应的实时分析作业来定义。

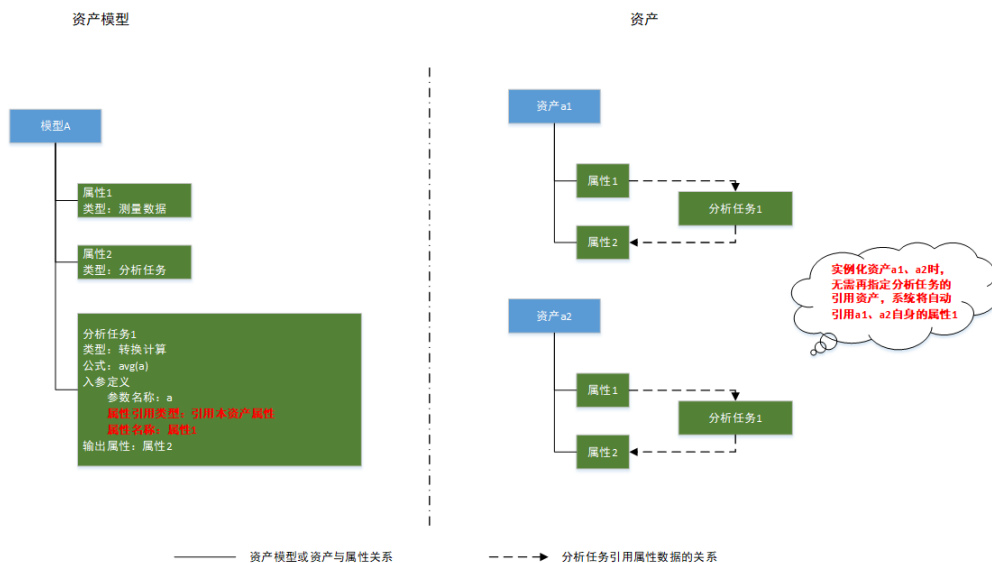
表 6-1 分析任务的类型和功能

类型	功能说明
转换计算	<p>转换计算表达式含义：对单个或多个参数的属性值进行常规数学计算，表达式中不能使用时间聚合。</p> <p>转换计算执行策略：当参数所引用的属性值发生变化时，系统就会执行一次转换计算，当定义了多个参数时，系统取变化参数的值和未变化参数历史5分钟内的最新值进行计算。</p>
聚合计算	<p>聚合计算表达式含义：对单个或多个属性进行周期性的时间聚合计算或空间聚合计算，表达式中可以使用时间聚合算子。</p> <p>聚合计算执行策略：每隔一定周期，系统对整个公式进行一次计算，输入参数值的变化不会触发运算。</p>
流分析	<p>每当输入参数的属性值变化时，系统就会将该输入参数的属性值（不包含其他未变化的参数值）转发到实时分析作业、触发一次实时分析作业计算。</p>

6.4.3 输入参数

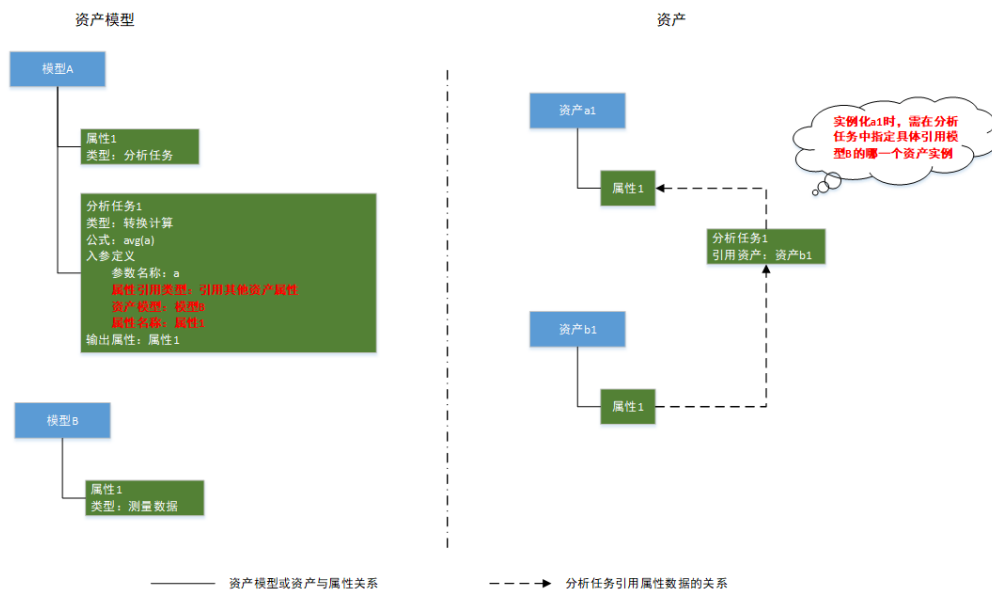
在定义“分析任务”时，首先需要定义分析任务的输入参数，该输入参数是用于定义计算公式中的形式参数，它的实际参数是对应于某个属性、或某些属性的值（即可能是单值或数组），通过3种属性引用类型来约定输入参数如何引用属性的值，详细解释如下：

- 引用本资产属性
利用资产模型对资产进行实例化时，该参数实际上是引用该资产实例自身的属性，此时入参对应的属性是单值，如下图所示。



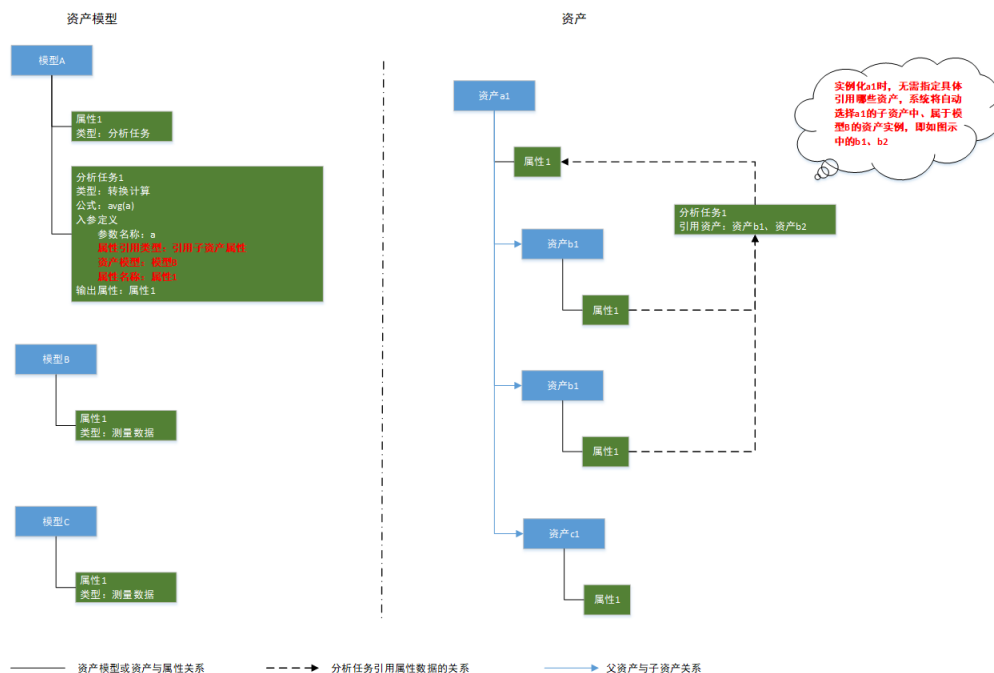
● 引用其他资产属性

利用资产模型对资产进行实例化时，该参数实际上是引用其他资产模型下某一个资产的属性，此时入参对应的属性是单值，如下图所示。



● 引用子资产属性

利用资产模型对资产进行实例化时，该参数实际上是引用该资产实例的下一层子资产中，符合指定资产模型的那一些子资产的属性，此时入参对应的属性是个数组，如下图所示。



注意

参数限制

- 同一个“分析任务”中不能有名称相同的参数。
- 当参数类型为“引用子资产属性”时，要求对应的资产实例下面，必须有至少一个匹配的子资产。
- 聚合计算的输入参数所引用的属性不能直接或间接来源于聚合计算。

6.4.4 计算公式

6.4.4.1 简介

计算公式由输入参数，字面量，操作符，内置函数几种元素组成。

示例

- $a + 1 + 2$
- $a * b * c + 1$
- $\text{Sum}([a, b, c, d])/4$
- $\text{If}(a \geq 0, a, -a)$
- $\text{If}(b == 1, \text{"on"}, \text{"off"})$

6.4.4.2 字面量

当前支持如下字面量

- 整形数字：如-123, 123, 456

- 浮点型数字：如-1.23, 8.88
- 字符串：如"abc", "hello world"
- 数组：如[1,2,3], ["a", "b", 123]

6.4.4.3 操作符

当前支持如下操作符

表 6-2 数学操作符

操作符	描述
+	用于两个数字相加。 当"+"的左操作数是字符串时，会将两个操作数拼接成字符串作为计算结果。例如 "123" + 123的计算结果是字符串"123123"。
-	用于两个数字相减。
*	用于两个数字相乘。
/	用于两个数字相除，如果除数是0，则跳过本次表达式的计算。 当左右操作数均为整数时，则执行整数除法。例如：公式"5/4" 的计算结果是1，而公式"5/4.0"的计算结果是1.25
>=	用于判断左操作数是否大于等于右操作数。
<=	用于判断左操作数是否小于等于右操作数。
==	用于判断左操作数是否等于右操作数。
!=	用于判断左操作数是否不等于右操作数。
&&	逻辑与操作，如a>100 && c<=20
	逻辑或操作，如 a<b c< d

如果在执行运算时发现类型不匹配，会跳过本次表达式的计算。以下是一些常见的不匹配场景

表 6-3 常见类型不匹配场景

公式	原因
"hello" * "world"	* 只能用于数字之间相乘，而"hello"和"world"都是字符串。
123 && a>b	&& 要求两个操作数均为布尔表达式，而123是整数类型。
[1,2,3] + [4,5,6]	* 只能用于数字之间相乘，不能用于数组之间相乘。

6.4.4.4 函数

当前系统内置了如下函数。

表 6-4 函数列表

函数名称	函数分类	函数功能	举例
Sum([...])	数学	求数组的总和值	Sum([1,2,3])，结果为6
Avg([...])	数学	求数组的平均值	Avg([1,2,3])，结果为2
Min([...])	数学	求数组的最小值	Min([1,2,3])，结果为1
Max([...])	数学	求数组的最大值	Max([1,2,3])，结果为3
Len([...])	通用	求数组的长度	Len([1,2,3,4])，结果为4
Get(json, key)	通用	获取Json内容里面key对应的值	Get({'age':30, 'name':"ironman"}, "name")，结果为"ironman"
If(condition, ifTrue, ifFalse)	通用	用于分支判断 当condition为true时，返回ifTrue的值，否则返回ifFalse的值	If(1<2, 'good', 'bad')，结果为'good' If(a>b, 'yes', 'no')，则当a>b时返回'yes'，否则返回'no' If可以嵌套使用。如If(a>b, If(a-b>2, a-b, a+b), If(a<=b, a*b, a/b))
Exit()	通用	退出分析任务，一般和If结合使用	If(a<0,exit(),'ok')
UUID()	通用	生成一个UUID	UUID()
EventTime()	通用	获取触发分析任务的时间，入参变化的时间或者定时任务时间，非运行分析任务的系统时间	EventTime()
GetAssetId(modelName,staticPropertyName,staticPropertyValue)	通用	获取资产ID，根据模型名称、静态属性名称、静态属性值获取资产ID，只能获取分析任务所在资产树以内的资产ID	GetAssetId('modelName','staticPropertyA','staticValue')
GetLastValue(assetId,propertyName,default)	通用	获取资产属性在分析任务触发之前的上一次值。注意，由于上一次的值可能还未入库，不建议应用在非常频繁变化的属性上	GetLastValue('assetId1','propertyA')

函数名称	函数分类	函数功能	举例
GetLastValueEq(assetId,tagPropertyName,tagPropertyValue,targetPropertyName,default)	通用	获取资产属性在分析任务触发之前的上一次值，并且会根据标签属性和标签属性值对同时刻的目标属性进行过滤。注意，由于上一次的值可能还未入库，不建议应用在变化非常频繁的屬性上	GetLastValueEq('assetId1','tagPropertyA','tagPropertyValue','propertyB')

函数名称	函数分类	函数功能	举例
GetValues(assetId,targetPropertyName,tags,default,from,to,limit)	通用	<p>获取资产属性的多个历史值，返回值数组，支持标签过滤，支持返回默认值，支持设置开始和结束时间，支持限定条数。</p> <p>参数：</p> <p>assetId: string类型，目标资产ID，必填</p> <p>targetPropertyName: string类型，目标属性名称，必填</p> <p>tags: object类型，过滤的标签，最多支持3个，若无则填写为{}，多个标签之间是与的关系，必填</p> <p>default: object类型，默认值，必填</p> <p>from: long string类型，开始时间，可以填写为long类型的毫秒，也可以填写为string类型的时间，格式为yyyy-MM-dd'T'HH:mm:ss.SSS'Z'，支持填写为""表示不指定开始时间，必填</p> <p>to: long string类型，结束时间，可以填写为long类型的毫秒，也可以填写为string类型的时间，格式为yyyy-MM-dd'T'HH:mm:ss.SSS'Z'，支持填写为""表示分析任务触发时间，必填</p> <p>limit: integer类型，返回条数，范围为(0,20]，必填</p>	GetValues('assetId1','propertyA',{tagPropertyB:123},-1,EventTime()-3600000,EventTime(),5)
GetByIndex(json,index)	通用	根据数组下标获取数组元素	GetByIndex(["a","b"],1)

函数名称	函数分类	函数功能	举例
StateTime(property, state)	时序聚合	求周期内状态的持续时长，单位为毫秒	StateTime(status, 'on')，结果为周期内status的值为on的总持续时间
TS_Sum(property)	时序聚合	求周期内某个属性的总和值	TS_Sum(qps)，结果为周期内的总qps值。
sin	数学	求弧度的sin值	$\sin(3.14/2) = 0.9999996829318346$
cos	数学	求弧度的cos值	$\cos(3.14) = -0.9999987317275395$
tan	数学	求弧度的tan值	$\tan(3.14/4) = 0.9992039901050427$
asin	数学	求输入值的asin值	$\text{asin}(1) = 1.5707963267948966$
acos	数学	求输入值的acos值	$\text{acos}(-1) = 3.141592653589793$
atan	数学	求输入值的atan值	$\text{atan}(1) = 0.7853981633974483$
sqrt	数学	求输入值的平方根	$\text{sqrt}(2) = 1.4142135623730951$ ， $\text{sqrt}(4) = 2.0$
ceil	数学	求大于等于输入值的最小整数	$\text{ceil}(3.8) = 4.0$ ， $\text{ceil}(-1.2) = -1.0$ ， $\text{ceil}(0) = 0.0$
floor	数学	求小于等于输入值的最大整数	$\text{floor}(3.8)=3$ ， $\text{floor}(-1.2)=-2$ ， $\text{floor}(0) = 0.0$
abs	数学	求输入值的绝对值	$\text{abs}(1.23) = 1.23$ ， $\text{abs}(-1.23) = 1.23$
log	数学	求输入值的自然对数值	$\log(2.718) = 0.999896315728952$
log10	数学	求输入值以10为底的对数值	$\log(100) = 2.0$

说明

- 只有当分析任务类型为”聚合计算“时，计算公式中才能使用“时序聚合”类函数

6.4.4.5 其他

6.4.5 分析任务输出

- 转换计算
单输出：分析任务的结果只能输出到本资产模型的某个属性上，且实例化成资产时固定输出到本资产。

多输出：分析任务的结果可以输出到多个不同资产模型的不同属性上，且实例化成资产时需要指定输出资产ID。

- 聚合计算
同转换计算
- 流计算
计算结果支持多输出，但实例化成资产时只能输出到本资产。

7 实时分析

7.1 作业开发

7.1.1 作业开发概述

作业开发主要包括如下功能：

- [新建作业](#)
- [编排作业](#)
- [保存作业](#)

7.1.2 新建作业

操作场景

实时分析是以作业为管理对象，您首先需要根据实时业务诉求创建实时分析作业。

说明

- IoT数据分析服务实时分析可使用数据管道与资产模型作为数据源，请您先确认已创建数据管道或资产模型，相关操作请参考[数据管道](#)、[资产模型](#)。
- 如果您所要求的实时任务需要发送数据到你申请的DIS云服务，请先进行DIS云服务授权配置，相关操作请参考[数据源--添加DIS数据源](#)。

新建作业

图 7-1 进入实时分析



图 7-2 查看作业

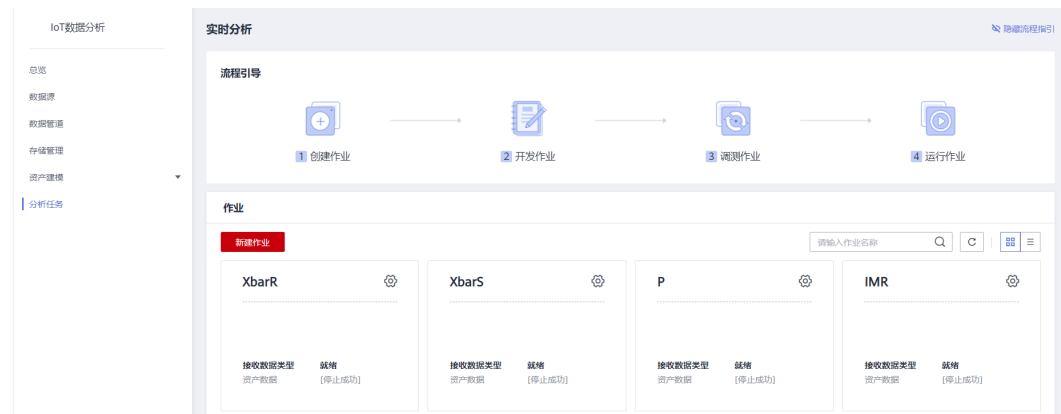


图 7-3 新建作业



参数说明

参数名称	描述	示例
接收数据类型	接收数据类型有“管道数据”、“资产数据”。	管道数据：直接使用数据管道清洗后的数据进行实时分析，当您无需资产建模，直接使用设备数据进行实时分析时，可使用这种接收数据类型。 资产数据：使用资产模型的数据进行实时分析，当您在资产模型中使用“分析任务”时，可使用这种接收数据类型。
作业名称	对作业进行命名	火警实时检测
描述	描述作业的业务	实时检测房间火灾告警

图 7-4 新建作业



7.1.3 编排作业

操作场景

按照业务诉求，以算子编排的方式来开发实时作业。

须知

- 实时作业是由“算子”和“数据流”组成一个DAG（有向无环图）。
- “算子”表示的是对数据的处理，本服务预制了多个算子供您使用，一个实时作业最多支持100个算子（即作业流图上显示的100个处理节点）。
- “数据流”表示在算子之间的传递数据，数据有明确的字段和格式，在编排界面，数据流以算子之间的连线表示。
- 算子分成3类，输入算子、转换算子和输出算子。一个完整的作业必须包含输入算子、转换算子和输出算子，否则不能正常执行。
 - 输入算子：负责实时作业数据的输入，是作业的属性类型。
 - 转换算子：负责对数据进行各种处理。
 - 输出算子：负责把作业处理的结果输出到作业外部。

编排作业

图 7-5 作业编排页面

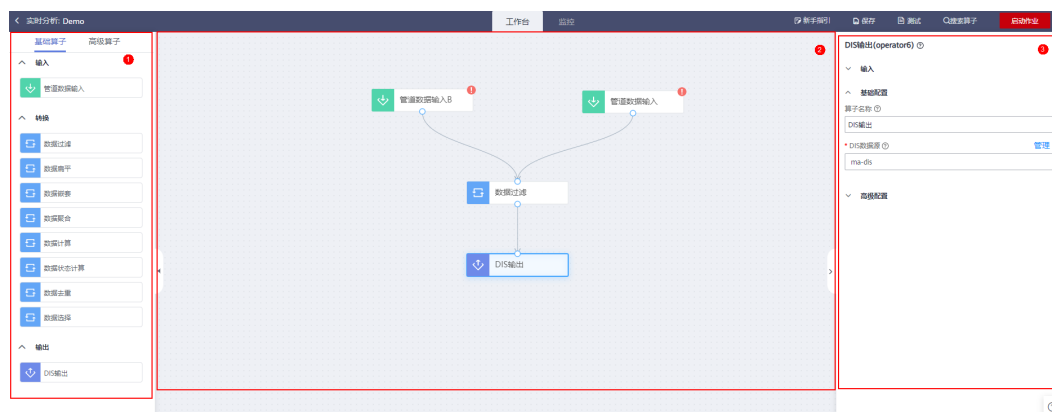


表 7-1 页面区域说明

序号	区域	描述
1	算子货架窗口	算子列表，提供多种输入算子、转换算子、输出算子，以及高级算子。
2	画布窗口	在画布上通过算子和线来编排实时作业。
3	算子参数配置窗口	点击画布上某个算子后，显示此算子的参数。

表 7-2 算子说明

算子类型	算子名称	描述									
输入算子	管道数据输入	<p>使用数据管道清洗过的数据作为实时分析的数据源。 算子配置项如图所示：</p> <p>算子配置(管道数据输入) ?</p> <p>^ 基础配置</p> <p>算子名称 ?</p> <p>管道数据输入</p> <p>* 数据管道 ?</p> <p>RTASinkTest X</p> <p>* 管道输出到 RTA 的数据名称 ?</p> <p>output01 X</p> <p>* 属性 ?</p> <table border="1"> <thead> <tr> <th><input checked="" type="checkbox"/></th> <th>* 属性名称 ?</th> <th>* 属性类型 ?</th> </tr> </thead> <tbody> <tr> <td><input checked="" type="checkbox"/></td> <td>value</td> <td>STRING</td> </tr> <tr> <td><input checked="" type="checkbox"/></td> <td>projectId</td> <td>STRING</td> </tr> </tbody> </table> <ul style="list-style-type: none"> 算子名称：用户指定这个算子的名称。 数据管道：用户选择要从哪个数据管道中获取数据。 注意：此处的管道列表只会列出含有“实时分析输出”算子的管道。 管道输出到RTA的数据名称：用户选择要使用哪个“管道输出数据名称”（“管道输出数据名称”在数据管道作业的“实时分析输出算子”中声明）。 属性：用户选择需要使用哪些属性进行后续的分析任务。 	<input checked="" type="checkbox"/>	* 属性名称 ?	* 属性类型 ?	<input checked="" type="checkbox"/>	value	STRING	<input checked="" type="checkbox"/>	projectId	STRING
<input checked="" type="checkbox"/>	* 属性名称 ?	* 属性类型 ?									
<input checked="" type="checkbox"/>	value	STRING									
<input checked="" type="checkbox"/>	projectId	STRING									

算子类型	算子名称	描述
	资产数据输入	<p>接收来自于资产模型的数据，以便进一步使用实时分析算子对资产数据进行分析，并将分析后的结果返回给资产模型，丰富资产模型内容。</p> <p>注意：使用该算子后，数据输出算子只可选择“资产数据输出”。</p> <p>算子配置项如图所示：</p> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p>算子配置(资产数据输入) ?</p> <p>^ 基础配置</p> <hr/> <p>算子名称 ?</p> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">资产数据输入</div> <p>• 输入参数 ?</p> <p>• 参数名称 ?</p> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">a</div> <p>• 参数类型 ?</p> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">STRING ▼</div> </div> <ul style="list-style-type: none"> 算子名称：用户指定这个算子的名称。 参数名称：指定这个“资产数据”在实时分析作业中使用时的参数名称，在资产建模的分析任务(资产建模->模型->分析任务)中可绑定输入参数与属性，如图所示： <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0; width: fit-content;"> <p>* 输入参数</p> <div style="display: flex; justify-content: space-between;"> <div> <p>只显示类型相同的属性</p> <p>属性名称 ?</p> <div style="border: 1px solid #ccc; padding: 2px;">propA</div> </div> <div> <p>属性引用类型 ?</p> <div style="border: 1px solid #ccc; padding: 2px;">引用本资产属性 ▼</div> </div> </div> </div> 参数类型：指定这个参数的数据类型，包括STRING/INTEGER/DOUBLE/OBJECT四种类型。 <p>注意：每个“资产数据输入”算子只可指定一个参数，若原作业中使用了多个参数，则需要使用多个“资产数据输入”算子。</p>
转换算子	数据过滤	实现了根据条件进行数据过滤。支持多个条件过滤数据，条件间是“与”或“或”的关系。当需要按单条数据本身的字段取值来决定是否过滤数据时使用本算子。
	数据扁平	把数据流中的嵌套Json字段转换为多个独立字段。如果测点类型是Object，那么是一个嵌套结构，需要把结构中的每个字段提取出来，成为数据流中的独立字段，这样数据流的所有字段都是字符串、数字等简单数据类型。

算子类型	算子名称	描述
	数据嵌套	把数据流中的字段打包成嵌套的Json格式字段，在实时作业输出数据时，如果需要输出嵌套的Json格式，那么使用此算子。
	数据聚合	<p>对多条数据进行聚合计算。支持算法为，求和（SUM），求平均（AVERAGE），求最大值（MAX），求最小值（MIN），求最旧值（FIRST），求最新值（LAST），求个数（COUNT），组装数组（ARRAY）。</p> <p>须知 求和（SUM），求平均（AVERAGE），求最大值（MAX），求最小值（MIN）算法支持STRING、INTEGER、LONG、FLOAT、DOUBLE数据类型，不支持其他数据类型。求最旧值（FIRST），求最新值（LAST），求个数（COUNT）支持所有数据类型。配置如果类型不匹配，输出此字段为null。</p> <p>聚合使用数据窗口，支持窗口类型为： 滚动窗口：窗口数据无重叠，需要指定窗口大小。比如每5分钟统计数据条数，那么每5分钟一个窗口，不会有重复统计。 滑动窗口：窗口有数据重叠，需要指定窗口大小，以及窗口移动大小。比如每1分钟统计最近5分钟数据条数，那么1分钟一个窗口，窗口大小为5分钟，重叠了4分钟数据。 会话窗口：窗口无固定大小，以多长时间没有数据作为窗口间隔。</p>
	数据计算	<p>需要生成新的字段或者改变字段的取值时使用。通过表达式配置（表达式配置与”数据管道”相同，详见 表达式说明）新字段的取值计算方式</p> <p>须知 对于新增加字段，在本算子的输出字段配置项中需要手工添加此字段，并且字段类型要和实际表达式计算结果的类型一致，如果不一致，算子会按照输出算子配置强制转换，如果转换失败，输出此字段为null。</p>
	数据状态计算	<p>支持以最新数据取值来进行条件计算。数据可能来自不同时间点，比如温度测点上报了温度，烟感测点上报了烟状态，因为这2个测点是不同数据中字段，并且不是同一时间点，使用其他算子不能进行组合判断，而本算子可以把数据缓存起来，在任意数据有变更时触发条件检测。</p> <p>须知</p> <ul style="list-style-type: none"> • 缓存的字段必须是本算子中条件中右值的字段。 • 仅缓存的字段最新取值。 • 需要指定缓存的索引字段，算子根据索引来查找对应缓存，索引字段支持多个数据字段组合。
	数据去重	<p>为了过滤掉重复数据，使用本算子。支持指定字段是否相同来判断是否是重复数据，比如指定温度字段没有变化则是重复数据，而事件时间字段每条数据取值不同。</p> <p>支持指定去重超时时长，在超时后，收到的第一条数据不按去重处理，然后重新开始去重计时以及去重检查。用于避免长时间数据字段值没有变化导致数据都被丢弃而不能产生数据。</p>

算子类型	算子名称	描述
	数据选择	对数据的字段进行过滤，选择的字段名字才输出，过滤掉不在字段列表中的字段。
输出算子	资产数据输出	<p>将实时分析的计算结果输出到资产模型，用于在资产模型中使用实时分析作业进行分析任务计算的结果。</p> <p>注意：使用该算子后，数据输入算子只可选择“资产数据输入”。</p> <p>算子配置项如图所示：</p>  <ul style="list-style-type: none"> 算子名称：用户指定这个算子的名称。 参数名称：指定分析结果在实时分析任务中输出时的名称，在资产建模的分析任务(资产建模->模型->分析任务)中可绑定输出参数与属性，如图所示：  参数类型：指定这个参数的数据类型，包括STRING/INTEGER/DOUBLE/OBJECT四种类型。
	DIS输出	<p>作业处理后的数据结果输出到您的DIS云服务。</p> <p>说明 请先进行DIS云服务授权配置，相关操作请参考数据源--添加DIS数据源</p>

算子类型	算子名称	描述
高级算子	去噪	<p>去噪算子，对数据流中的数据根据窗口周期进行去噪。</p> <p>主要算子配置项包括：</p> <ul style="list-style-type: none">• 分区字段：指定用于窗口中数据分区的字段列表。当选择多个字段时，按指定字段顺序的组合成字符作为算子并行处理的分区的Key。如果是接收资产数据，那么不填写分区字段，则默认以资产作为分区。如果是接收管道数据，默认只有一个分区，请务必指定字段作为分权键。• 窗口类型：指定使用窗口的类型，TumblingTimeWindows是滚动时间窗口，SlidingTimeWindows是滑动时间窗口，SessionTimeWindows是会话窗口，TumblingCountWindows是滚动计数窗口，SlidingCountWindows是滑动计数窗口。• 时间类型：指定窗口时间如果为EventTime，则窗口时间按照消息中的时间，如果为ProcessingTime，则按系统处理时间。• 窗口大小：对于TumblingTimeWindows和SlidingTimeWindows类型窗口，单位为秒；对于SessionTimeWindows类型窗口，表示无消息的时间间隔，用于分割Session，单位为秒；对于TumblingCountWindows、SlidingCountWindows、TumblingCountTimeWindows和SlidingCountTimeWindows类型的窗口，单位为个。• 滑动大小：窗口滑动长度，对于SlidingTimeWindows类型窗口，单位为秒；对于SlidingCountWindows和SlidingCountTimeWindows类型的窗口，单位为个。• 时间偏移：如果要求窗口时间与时区对齐，填写时区偏移，对于TumblingTimeWindows和SlidingTimeWindows类型窗口有效，单位为秒。• 去噪字段：指定去噪字段进行计算。• 去噪算法：指定去噪的算法类型，MaxCountOfValue/MinCountOfValue是某个取值出现的次数最多/最小则为正常数据。

须知

算子如果配置正确，提示为：



保存成功，作业已处于就绪态，可启动运行。



算子如果配置错误，提示为：



保存成功，但作业存在配置错误，已存为草稿。



7.1.4 保存作业

操作场景

在作业开发过程中或者开发完成，您可以保存作业，这样再次进入作业编排界面时可以继续开发工作。

保存作业

图 7-6 保存作业



须知

保存作业时，如果作业中存在错误，那么系统会提示保存为草稿，如果作业校验正常，那么系统会提示作业就绪，可启动运行。

7.2 作业调试

7.2.1 调试概述

- 为什么实时分析任务没能输出期望的结果？
- 实时分析任务中哪个算子环节出错了？
- 算子处理后的数据结果是什么样的？

以上问题是在开发阶段经常遇到的问题。为了解决以上问题，我们提供了调试实时分析工具。用户可以上报测试数据对实时分析业务流程进行测试，用户可直观查看测试数据在每个算子输入和输出的内容。

工作原理：

启动测试模式后，输入类算子从数据源读取数据，将数据发送到下游的转化类算子，最终专递到输出类算子。调试工具会在每个算子的输入输出端监听数据流，最终将监听收集到的数据在页面进行展示。我们可以根据算子的输入输出数据定位以上我们所遇到的问题。



界面介绍

- [启动按钮](#)

- [任务状态栏](#)
- [输入输出](#)
- [数据监控](#)

[输入输出](#)

[创建任务](#)

[启动测试任务](#)

[下载数据模板](#)

[修改数据模板](#)

[上传数据文本](#)

[查看测试结果](#)

[结束测试任务](#)

[数据模板](#)

- [资产数据输入](#)
- [管道数据输入](#)
- [资产数据输入（旧）](#)
- [数据存储数据源（旧）](#)

表 7-3

算子类别	算子中文名称	算子英文名称	是否支持调试	是否展示输入数据	是否展示输出数据
输入算子	资产数据输入	AssetData Source	是	否	是
		AssetJobSource	是	否	是
	数据存储数据源（旧）	DataStore Source	是	否	是
	资产数据输入（旧）	IoTSource	是	否	是
	管道数据输入	PipelineSource	是	否	是
转化算子	数据聚合	Aggregate	是	是	是
	数据去重	Deduplicate	是	是	是
	去噪	Denoise	是	是	是
	数据过滤	Filter	是	是	是

算子类别	算子中文名称	算子英文名称	是否支持调试	是否展示输入数据	是否展示输出数据
	数据扁平	Flat	是	是	是
	数据计算	Map	是	是	是
	数据嵌套	Nest	是	是	是
	数据选择	SelectFields	是	是	是
	数据状态计算	StateCondition	是	是	是
输出算子	资产输出	AssetDataSink	是	是	否
	DIS输出	DISSink	是	是	否
	资产输出(老)	IoTSink	是	是	否

📖 说明

- 不同的输入算子的测试模板内容格式会不一样，建议在模板的基础上添加测试数据。

7.2.2 界面介绍

- 启动按钮
 - 启动调试任务



- 任务状态栏
 - 查看任务状态。



表 7-4 状态说明

状态名称	描述
就绪	任务完成编辑后，无错误配置，等待启动。
提交中	任务提交后台服务中，此时任务处于初始化中，无法处理上报的测试数据
运行中	任务已完成初始化，等待或正在处理上报测试数据，除当任务异常或人为停止，否则将一直正常运行。
停止中	当任务异常或人为停止任务时，任务将停止处理测试数据。
停止	任务正常停止，可编辑。

- 输入输出




- 上报测试数据的控制页面。用户可在数据源类型选择对应数据源；下载对应数据源的测试数据模板文本并修改数据内容；最后上传测试数据文本。

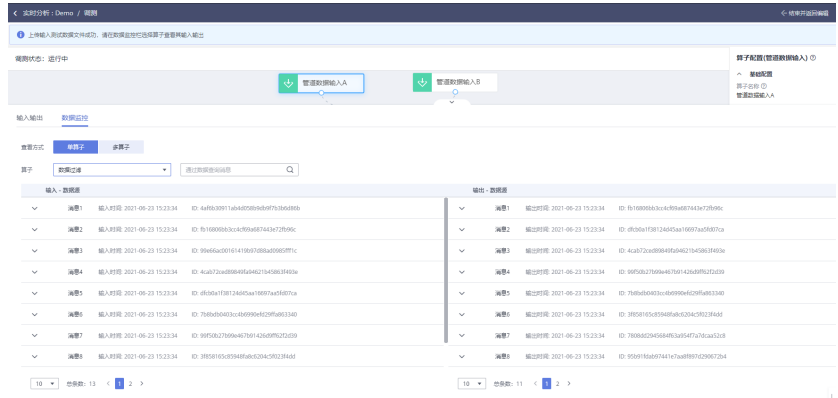


- 数据监控


- 数据监控也是查看数据处理结果的查询页面。页面查看方式包括单算子查询和多算子查询。


- 单算子查询


- 下拉“算子”下拉框选择查看“数据过滤”。
- 点击搜索输入框  按钮搜索查询，搜索框支持以消息id或data内容匹配查询。
-  和  分别显示“数据过滤”算子结果。

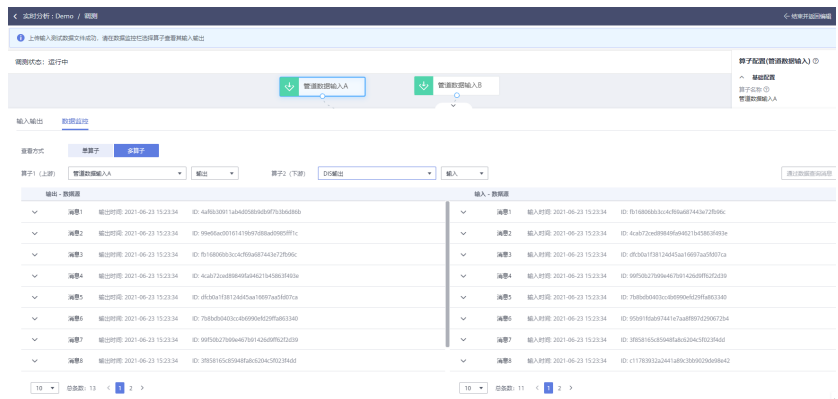


■ 多算子查询

- 下拉“算子1（上游）”下拉框选择查看“管道数据输入A”的“输出”；下拉“算子2（下游）”下拉框选择查看“DIS输出”的“输入”。
- 点击搜索输入框  按钮搜索查询，搜索框支持以消息id或data内容匹配查询。

○  显示“管道数据输入A”的“输出”结果。

○  显示“DIS输出”的“输入”结果。



7.2.3 操作指导

使用测试功能主要包含启动测试任务，上报数据，查看测试结果，停止测试任务。其中上报数据，查看测试结果可重复执行，不必反复启停任务。

步骤1 创建任务

1. 创建并编辑任务。
2. 保存任务。



3. 确保任务为“就绪”态。

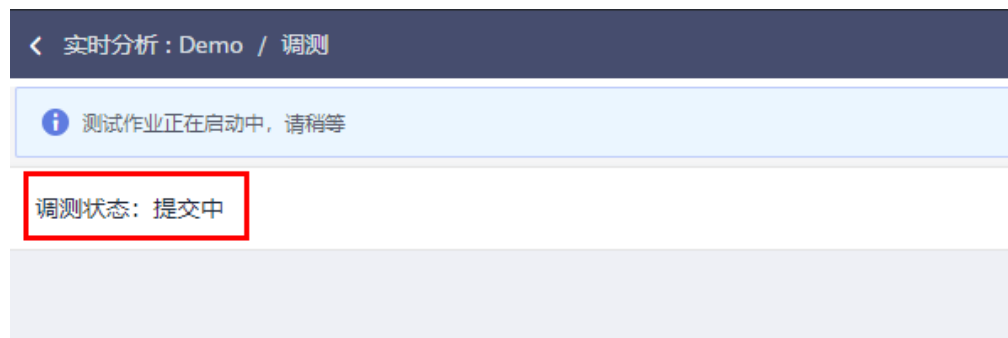


步骤2 启动测试任务

1. 点击测试按钮，等待测试状态为“运行中”。目前部分算子暂不支持测试功能(见[算子支持表](#))，测试按钮将不可点击。



2. 自动跳转测试页面，当前状态为提交中，需要耐心等待。状态包含：“就绪”、“提交中”、“运行中”、“停止中”、“停止”。

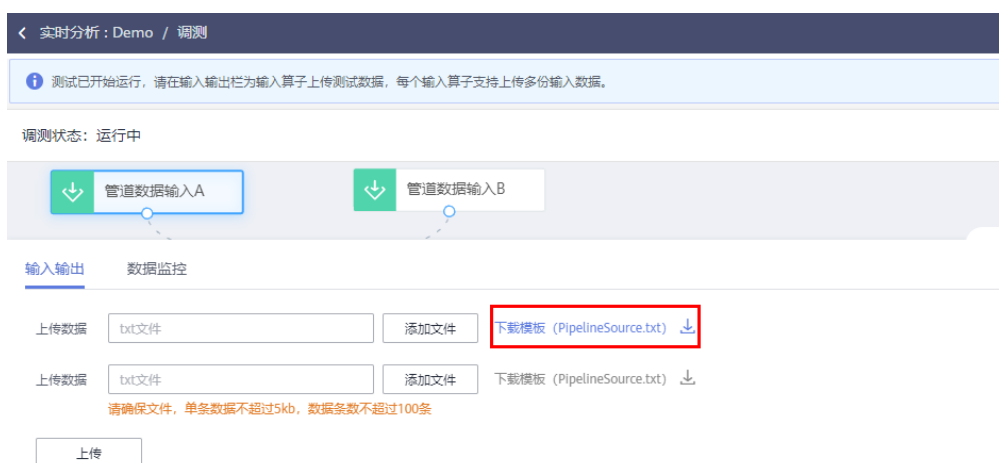


3. 当调测状态为“运行中”时，完成启动测试任务操作。



步骤3 下载数据模板

1. 点击下载测试模板。



说明

- 不同的输入算子的测试模板内容格式会不一样，建议在模板的基础上添加测试数据。

步骤4 修改数据模板

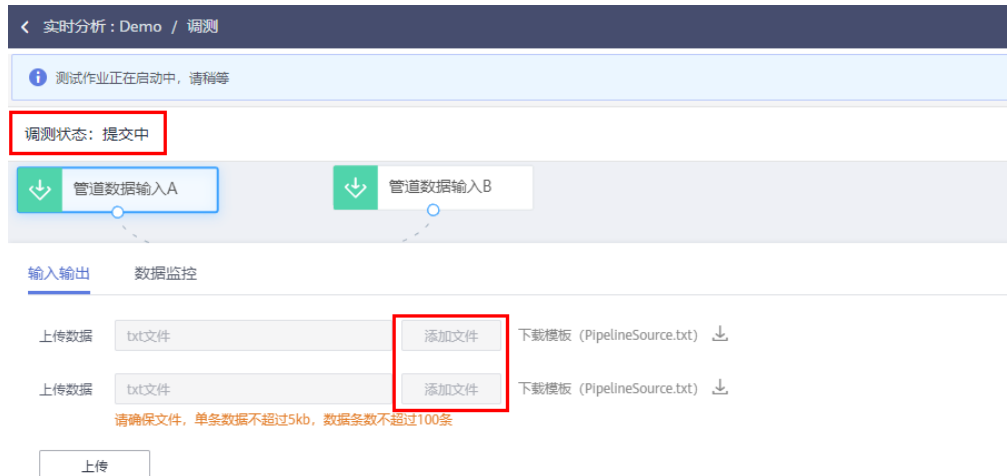
1. 删除多余不需要的数据，写入业务数据内容。参考[数据模板](#)章节。

说明

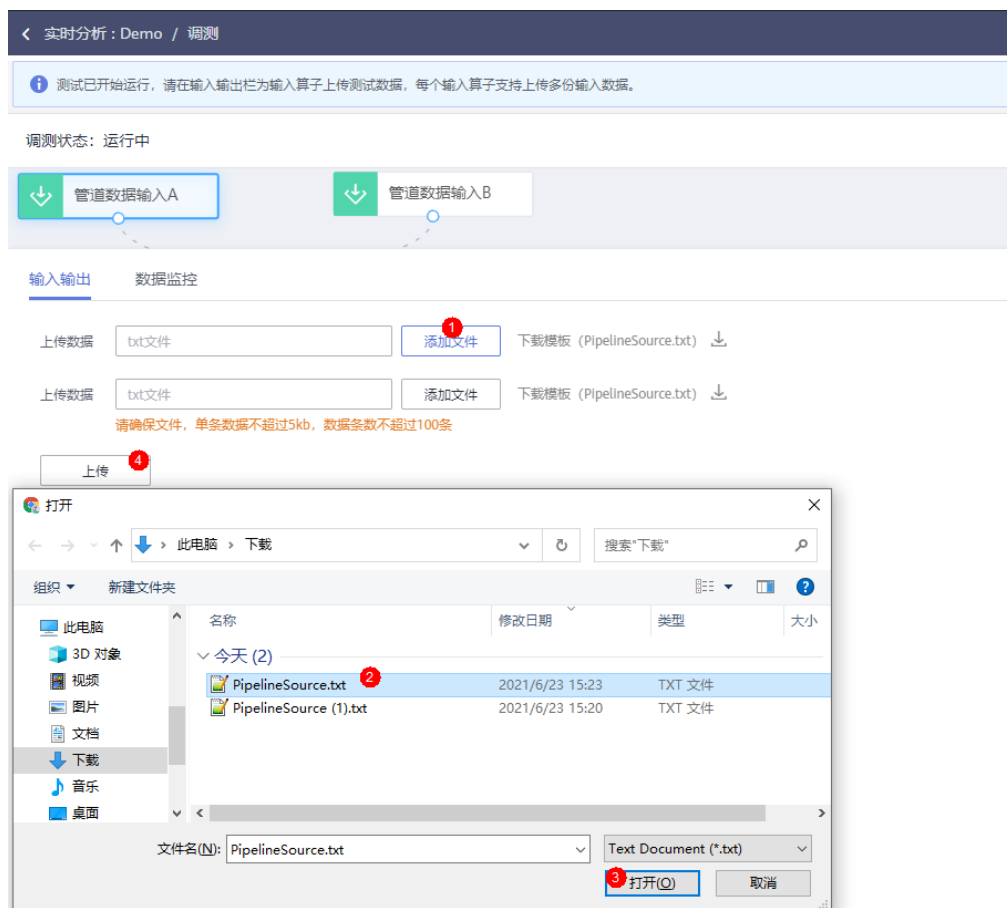
- 测试数据文本支持UTF-8编码。
- 测试数据必须是合法Json格式，且不换行。
- 测试数据文本可写多条数据，每条数据以换行符分隔，即一条数据必须写在一行文本里。
- 每条数据大小不能超过5kb。
- 一个文本里不能超过100条测试数据。
- 测试数据文本中必填项不能缺损、错误，建议参考测试模板内容。
- 注释内容不会上报。

步骤5 上传数据文本

1. 当调试状态为“就绪”或“提交中”时，暂不能添加上传测试数据。



2. 当调测状态为“运行中”时，才能添加上传测试数据。

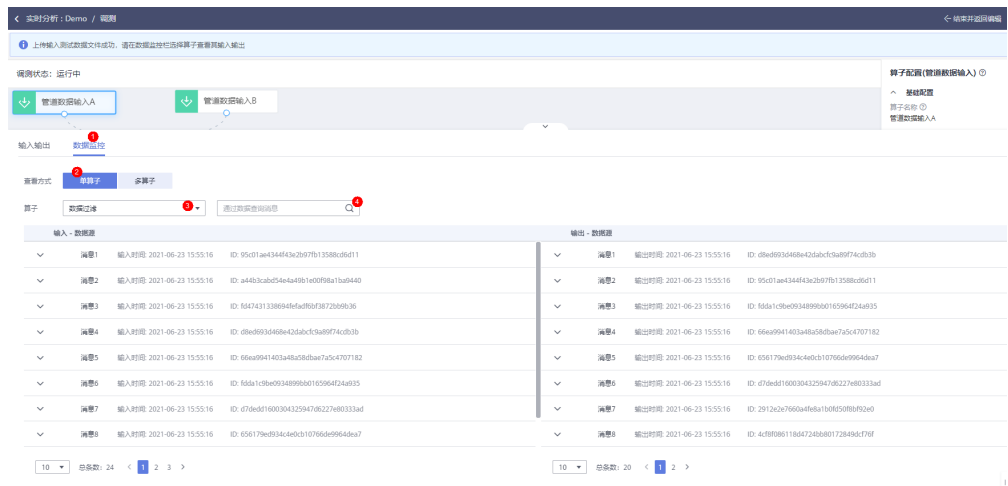


3. 上传成功



步骤6 查看测试结果

1. 切换到数据监控页签
2. 选择查看模式
3. 选择目标算子
4. 点击查询按钮



步骤7 结束测试任务

1. 点击右上角“结束并返回编辑”按钮



步骤8 如果需要重复测试，可重复步骤5、步骤6。

----结束

说明

- 调试实时分析需要占用计算资源。
- 一小时内算子没有新增测试数据，测试数据将会被老化。重启测试任务后，测试数据也不会保留。
- 同一个任务只能处于调测状态或正常运行态。

7.2.4 数据模板

- 资产数据输入

数据模板	字段名	字段说明	是否必填
<pre>{ "type": "FunctionCall", "message": { "projectId": "080dd91ecc802xxxxx01be5f9 d68d", "jobId": "J_79519506820943xxx", "context": { "assetId": "assetId01", "analysisId": "analysisId01" }, "input": { "integer": "", "double": "", "string": "", "object": { "a": "", "b": "" } }, "timestamp": 1624440752632 }, "msgVersion": "1.0" }</pre>	type	数据类型，必须模板默认值，用户不能修改	是
	message	消息体，必须模板默认值，用户不能修改	是
	projectId	租户id，必须模板默认值，用户不能修改	是
	jobId	当前实时分析任务id，必须模板默认值，用户不能修改	是
	context	资产信息，必须模板默认值，用户不能修改	是
	input	输入数据，必须包含算子配置的“参数名称”。例如input中包含“integer”字段。  数据支持integer、double、string、object（json）类型	是
timestamp	时间戳字段，Long类型，用户可自定义数值。	是	

数据模板	字段名	字段说明	是否必填
	msgVersion	消息版本，必须模板默认值，用户不能修改	是

- 管道数据输入

数据模板	字段名	字段说明	是否必填
<pre>{ "header": { "projectId": "080dd91ecc8026bxxxxx01be5f9d68d", "dataSourceId": "", "pipelineId": "J_79519506820943xxx", "rtaOutputName": "output01" }, "body": { "eventTime": "1624442214553", "data": { "integer": "", "double": "", "string": "", "object": { "a": "", "b": "" } } } }</pre>	header	消息头，必须模板默认值，用户不能修改。	是
	projectId	租户id，必须模板默认值，用户不能修改。	是
	dataSourceId	数据源id，必须模板默认值，用户不能修改。	是
	pipelineId	管道任务id，必须模板默认值，用户不能修改。	是
	rtaOutputName	RTA 的数据名称，必须模板默认值，用户不能修改。	是
	body	输入数据，必须包含算子配置的“参数名称”。例如input中包含“integer”字段。	是
	eventTime	时间戳字段，Long类型，用户可自定义数值。	是
data	管道输出字段集合，数据支持integer、double、string、object (json) 类型。用户可自定义data内部字段名和值。	是	

- 资产数据输入 (旧)

数据模板	字段名	字段说明	是否必填
<pre>{ "type": "MeasPointValue", "message": { "projectId": "092b845deb00dxxxxxx0181b e20fa1", "viewId": "trace_test", "viewName": "", "assetId": "", "assetName": "", "assetCategory": "", "measurePointId": "", "measurePointName": "", "measurePointCategory": "", "valueType": "", "value": "", "eventTime": "20210623T181127+0800" }, "msgVersion": "1.0" }</pre>	type	数据类型，必须模板默认值，用户不能修改	是
	message	消息体，必须模板默认值，用户不能修改	是
	projectId	租户id，必须模板默认值，用户不能修改	是
	viewId	视图id，必须模板默认值，用户不能修改	是
	viewName	视图名称	否
	assetId	资产id	否
	assetName	资产名称	否
	assetCategory	资产目录	否
	measurePointId	测点id	否
	measurePointName	测点名称	否
	measurePointCategory	测点目录	否
	valueType	数据类型，数据支持 integer、double、string、object (json) 类型，用户自行定义。	是
	value	数据值，数据支持 integer、double、string、object (json) 类型，用户自行定义。	是
eventTime	时间字段，用户可自定义数值。	是	
msgVersion	消息版本，必须模板默认值，用户不能修改。	是	

- 数据存储数据源（旧）

数据模板	字段名	字段说明	是否必填
<pre>{ "header": { "projectId": "092b845deb00d45cxxxxx1be 20fa1", "dataSouceId": "", "pipelineId": "", "dataStoreId": "0264bbad- xxxx-4f5d-xxxx-5d87c1bfaa90" }, "body": { "deviceId": "", "eventTime": "1624503993013", "data": { "integer": "", "double": "", "string": "", "object": { "a": "", "b": "" } } } }</pre>	header	消息头，必须模板默认值，用户不能修改。	是
	projectId	租户id，必须模板默认值，用户不能修改。	是
	dataSouceId	数据源id，必须模板默认值，用户不能修改。	是
	pipelineId	管道任务id，必须模板默认值，用户不能修改。	是
	dataStoreId	数据存储id，必须模板默认值，用户不能修改。	是
	body	输入数据，必须包含算子配置的“参数名称”。例如input中包含“integer”字段。	是
	deviceId	设备id，用户可自定义。	否
	eventTime	时间戳字段，Long类型，用户可自定义数值。	是
data	管道输出字段集合，数据支持integer、double、string、object (json) 类型。用户可自定义data内部字段名和值。	是	

7.3 作业运行

7.3.1 作业运行概述

作业运行主要包括如下功能：

- [启动作业](#)
- [查询作业状态](#)
- [查看作业操作记录](#)
- [停止作业](#)

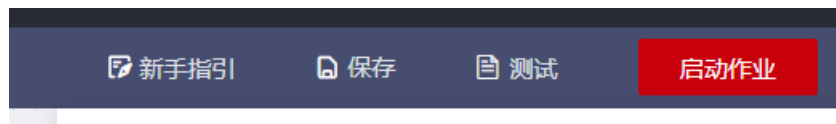
7.3.2 启动作业

操作场景

启动作业。

启动作业

启动运行



资源配置

资源配置

* RTU数 ?

* 并行数 ?

启动

取消

表 7-5 参数说明

参数名称	描述	示例
RTU	作业运行RTU个数，当使用独享资源运行时，每个作业至少使用2个RTU，最多可使用1024个RTU。	2
并发数	作业运行的并发数据，RTU越多时，支持更高的并发数。并发数不得超过 $(RTU个数 - 1) \times 4$ 。	2

说明

RTU为实时分析作业的资源单位，实时分析功能按照RTU个数和时长来计费。

7.3.3 查询作业状态

操作场景

查询作业状态，以确定运行作业是否正常。

查询作业状态

图 7-7 查询作业状态



作业运行状态有：草稿、就绪、调试中、提交中、运行中、停止中、提交成功、提交失败、停止成功。

7.3.4 查看作业操作记录

操作场景

查看此作业的操作记录。

查看作业操作记录

查看作业操作记录

操作时间	操作内容	状态
2021-07-19 18:10:46	启动作业	成功
2021-07-19 18:10:39	更新实时作业	成功
2021-07-19 17:35:43	停止作业	成功
2021-07-19 17:35:25	启动作业	成功
2021-07-19 17:35:17	更新实时作业	成功
2021-07-19 17:35:08	更新实时作业	成功
2021-07-19 17:35:04	停止作业	成功
2021-07-19 10:09:37	启动作业	成功
2021-07-19 10:09:30	停止作业	成功
2021-07-14 17:26:41	启动作业	成功

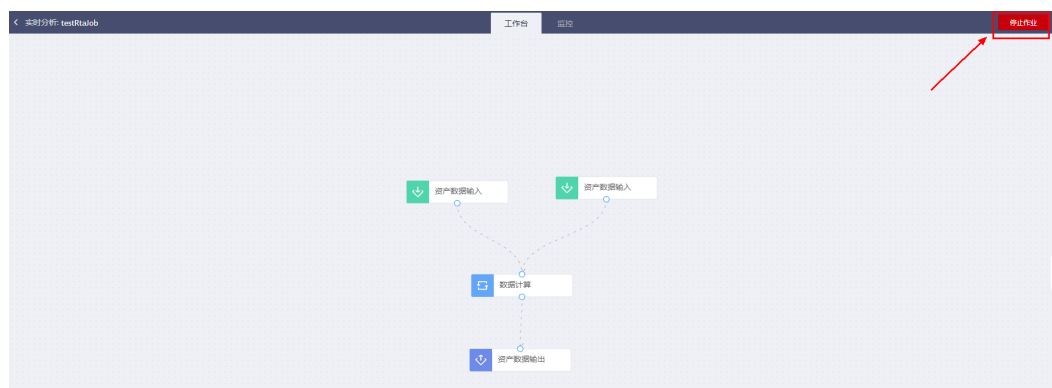
7.3.5 停止作业

操作场景

停止作业。

停止作业

图 7-8 点击工作台右上角的“停止运行”按钮停止作业



8 时序分析

8.1 时序分析简介

8.1.1 概念简介

物联网数据的数据体量大、价值密度小、数据时效性高、质量低。针对物联网数据的这些特点，在海量历史数据中提供近实时的高效探索分析能力，及时发现物联网数据中的隐含规律则显得尤为重要。IoT数据分析服务的时序分析，提供基于模型驱动的海量历史数据时序分析洞察能力，让物联网数据的规律清晰可见。

8.1.2 功能简介

时序分析主页面分区和主要功能描述如下面图表所示。

时序分析页面



表 8-1 页面分区说明

序号	区域	描述
1	时序探索导航栏	基于租户创建的资产模型提供资产数据的时序探索能力；基于租户在存储管理中每一个数据存储中配置的属性列表、提供了设备数据时序探索能力。
2	时间选择区	针对已选的资产数据或设备数据，设置时序分析的起止时间窗，提供日历和时间轴两种设置方式。
3	预览列表区	针对已选的资产数据或设备数据，展示属性的详细信息，并支持对所选属性做勾选、分析、移除等操作。
4	图表分析区	针对已选的资产数据或设备数据，提供基于时间的多种聚合计算能力，支持折线图、热力图、散点图等多种图表分析方式。

8.2 资产时序探索

8.2.1 选择根资产

前提条件

您已经在IoT数据分析服务的“资产建模”模块完成资产模型的构建，如未完成请先前往创建。

选择根资产

在模型导航栏中选“资产时序探索”、在“选择根资产”下拉框中选择您在需要的根资产。

图 8-1 选择根资产

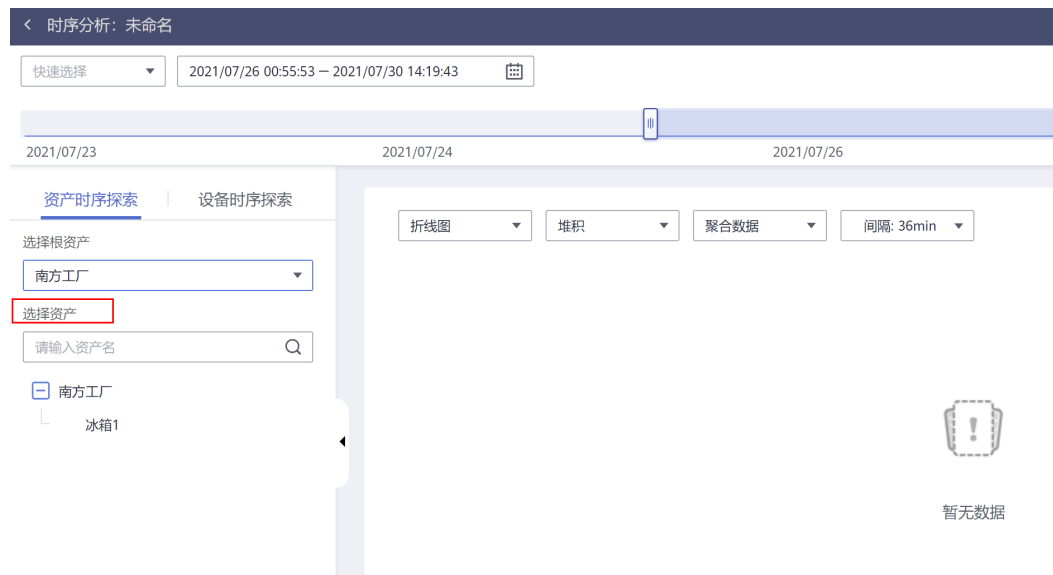


8.2.2 选择资产

选择资产

在“选择根资产”步骤完成后，对应的资产树会显示在模型导航栏的“选择资产”区域，逐层展开查找并选择您关心的资产，您也可以在搜索框内指定资产名称直接查找。

图 8-2 选择资产



资产局部路径置顶

为方便资产的选择查找，在模型导航栏的“选择资产”区域，针对资产树的非叶子节点，提供资产的局部路径置顶功能，置顶后只展示资产树的所选路径子节点的资产，对应操作和置顶后的效果如下图所示。

图 8-3 资产局部路径置顶

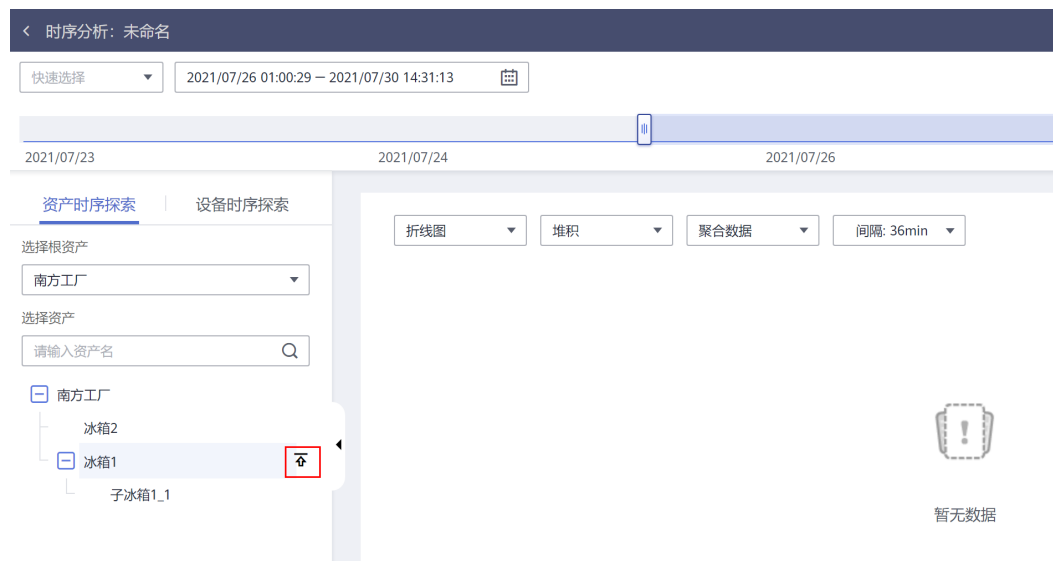


图 8-4 资产树置顶后局部显示

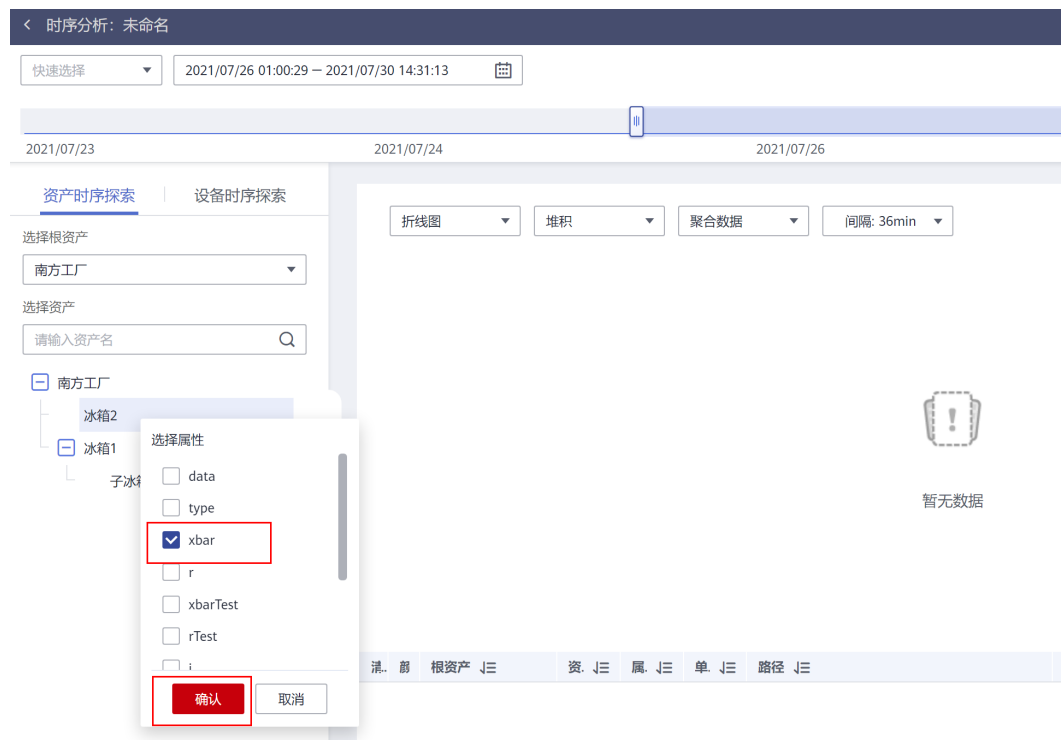


8.2.3 选择属性

选择属性

在模型导航栏的“选择资产”区域，选择您关注的资产后，单击该资产出现属性弹框，查看并勾选需要做时序分析的属性，点击“确定”完成。

图 8-5 选择属性

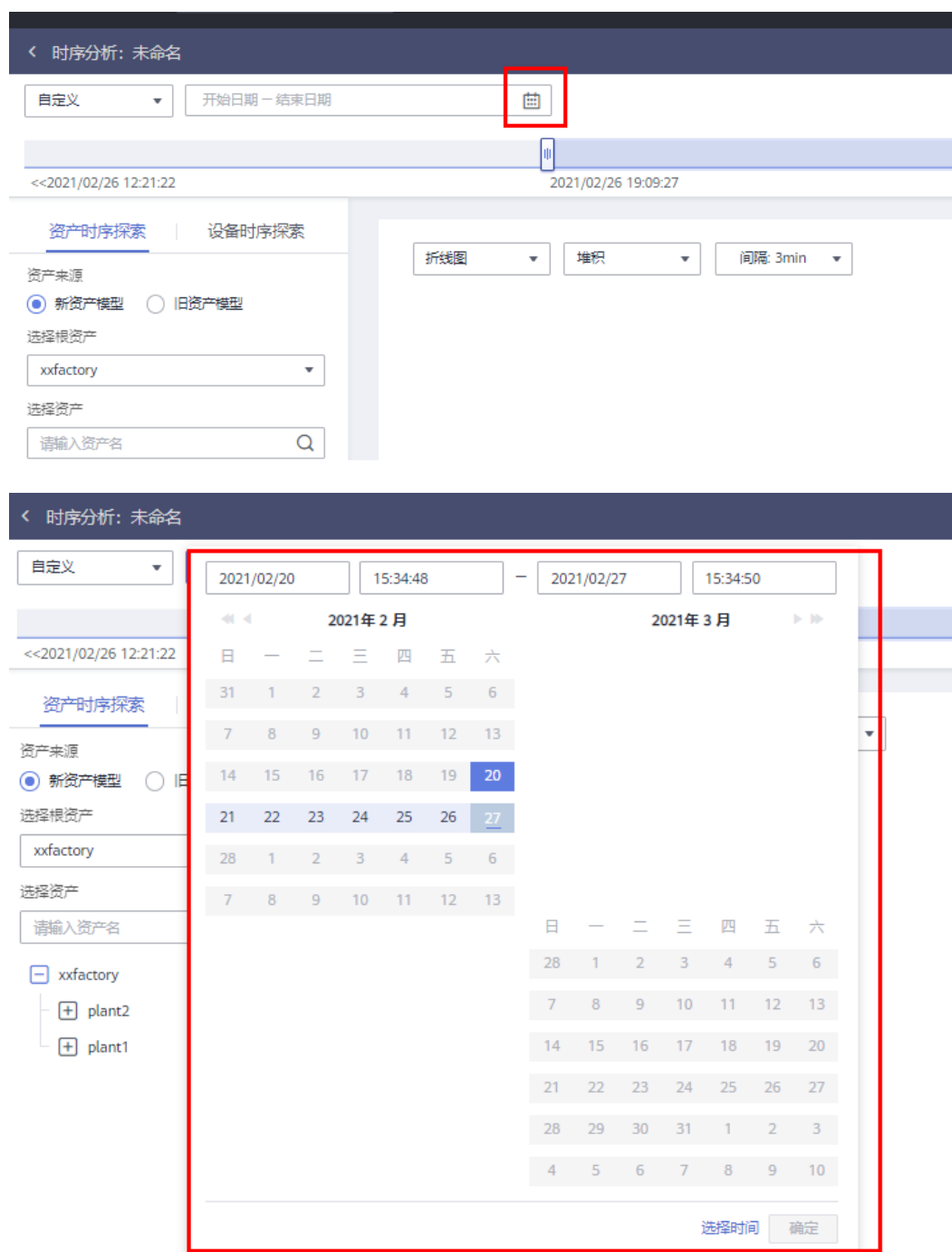


8.3 设置时间窗

日历选择时间

在时间选择区，可以通过日历控件设置时间窗。点击日历图标按钮，在弹出的对话框中设置起始和结束时间，点击“确认”完成。

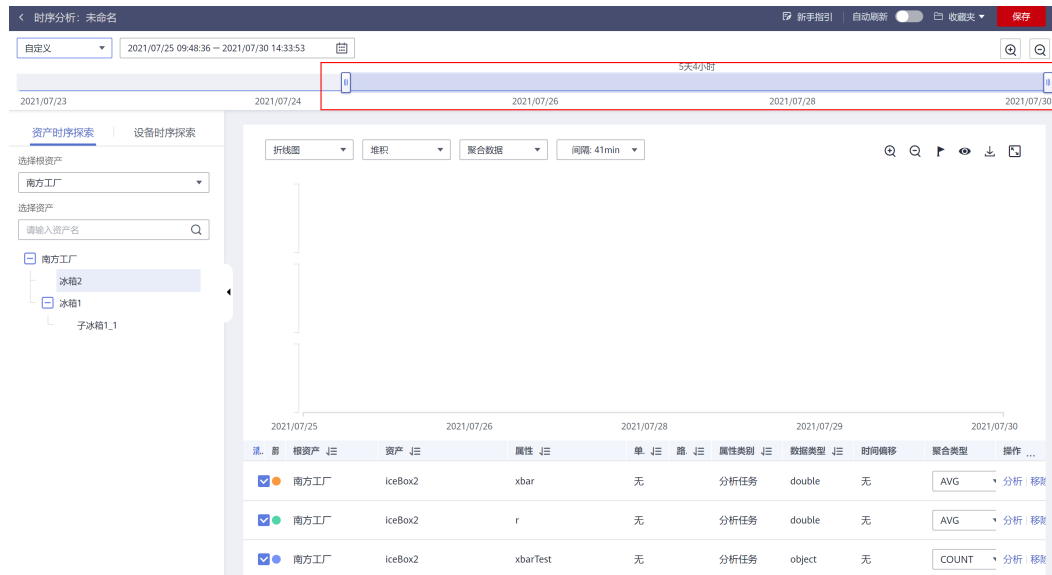
图 8-6 日历选择时间



时间轴选择时间

在时间选择区，可以通过时间轴拖动设置时间窗。时间轴的蓝色部分为选中的时间窗，起止时间显示与日历控件同步。

图 8-7 时间轴选择时间



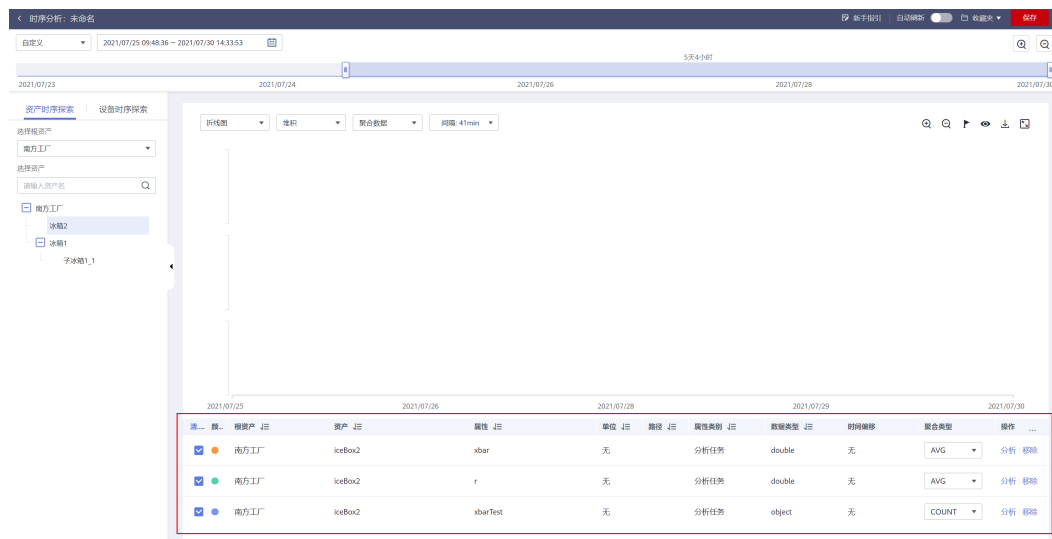
8.4 时序洞察

8.4.1 预览列表展示

预览列表展示属性详情信息

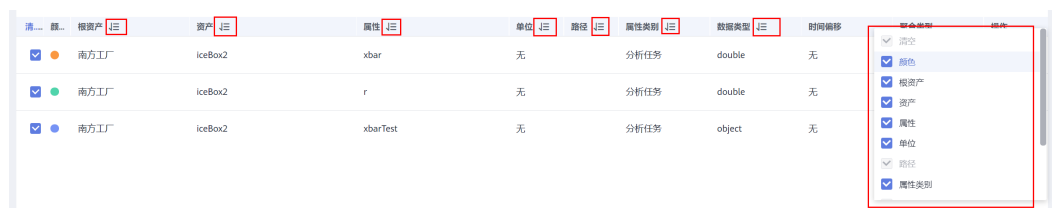
在时序探索导航栏勾选了资产和属性后，会对应在预览列表区展示所选属性的详情信息，包含属性所属的资产、属性名称、单位、绘图颜色等。

图 8-8 预览列表展示属性详情



预览列表区支持按列进行升降序排序和列编辑。

图 8-9 预览列表支持按列排序和列编辑

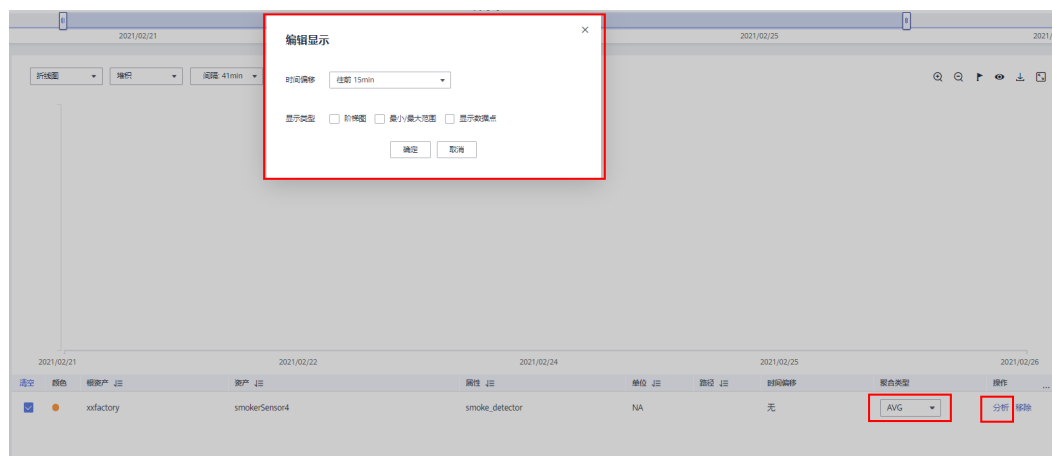


预览列表控制属性分析操作

在预览列表区，支持对所选属性做勾选、分析、移除操作，以控制该属性在图表分析区的展示，其中分析操作包含如下内容：

- 聚合类型：对所选属性的时序数据按采样间隔进行聚合计算，当前支持的聚类操作包括AVG、MIN、MAX、SUM、COUNT共计5种；
- 时间偏移：对所选属性按时间做偏移，会在预览列表新增一条记录，常用于同一测点不同时间的对比分析；
- 显示类型：支持阶梯图，最小/最大范围，显示数据点，仅对折线图有效。

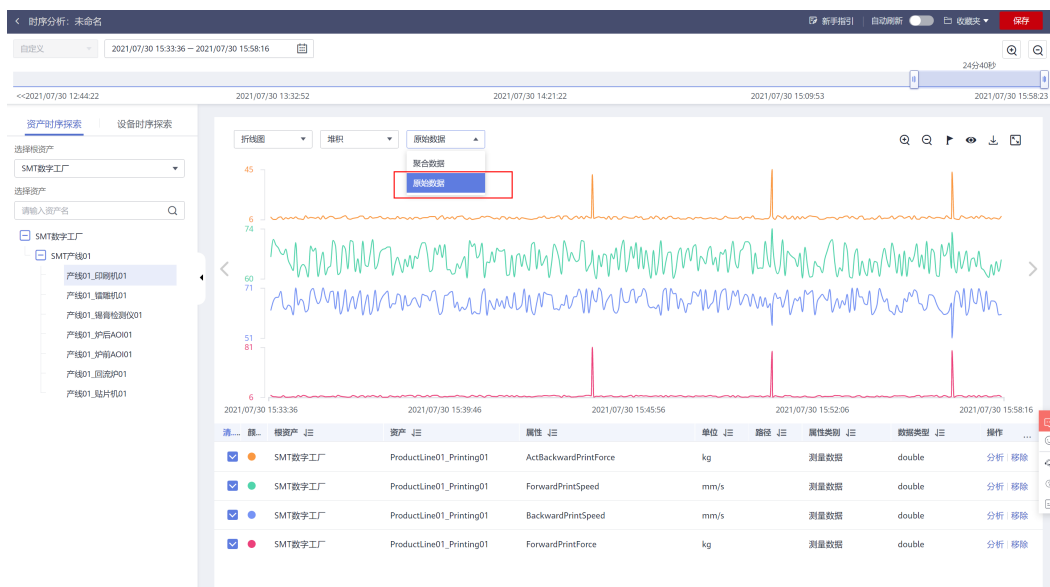
图 8-10 预览列表分析操作



8.4.2 原始数据查看

在IoT数据分析服务时序分析功能中，核心功能之一就是图形化方式查看资产数据、或者设备数据的历史曲线。以查看资产数据历史曲线为例，在时序探索导航栏中选择“资产时序探索”页签，然后选择您所关注的资产属性，在图标分析区中选择“原始数据”，即可展示出资产属性数据的历史曲线，如下图所示。

图 8-11 选择原始数据、查看历史数据曲线

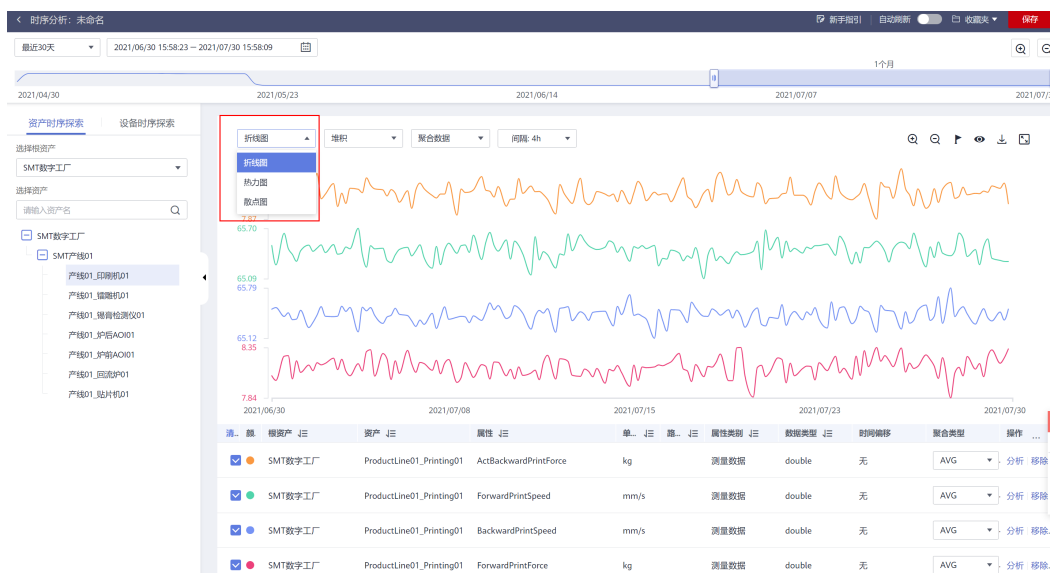


8.4.3 图表探索分析

选择图表类型

在图表分析区，支持对所选属性按折线图、热力图、散点图3种方式做时序分析。

图 8-12 选择图表类型



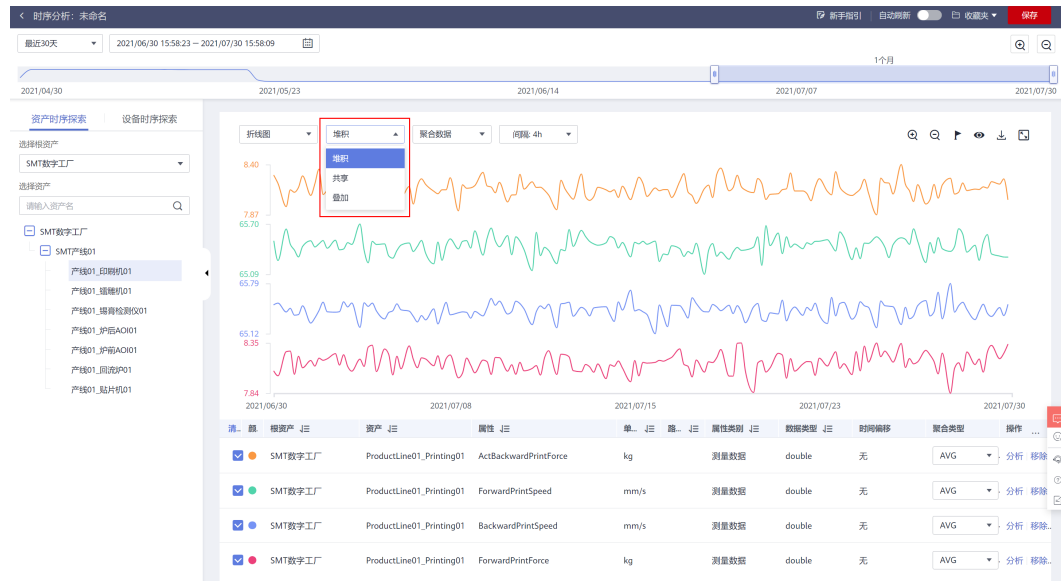
设置折线图 Y 轴显示

在图表分析区，折线图可以设置Y轴的显示方式，包含如下3种方式：

- 堆积：Y轴不重叠；
- 共享：Y轴重叠，Y轴单位长度代表的刻度等长；

- 叠加：Y轴重叠，Y轴单位长度代表的刻度不等长。

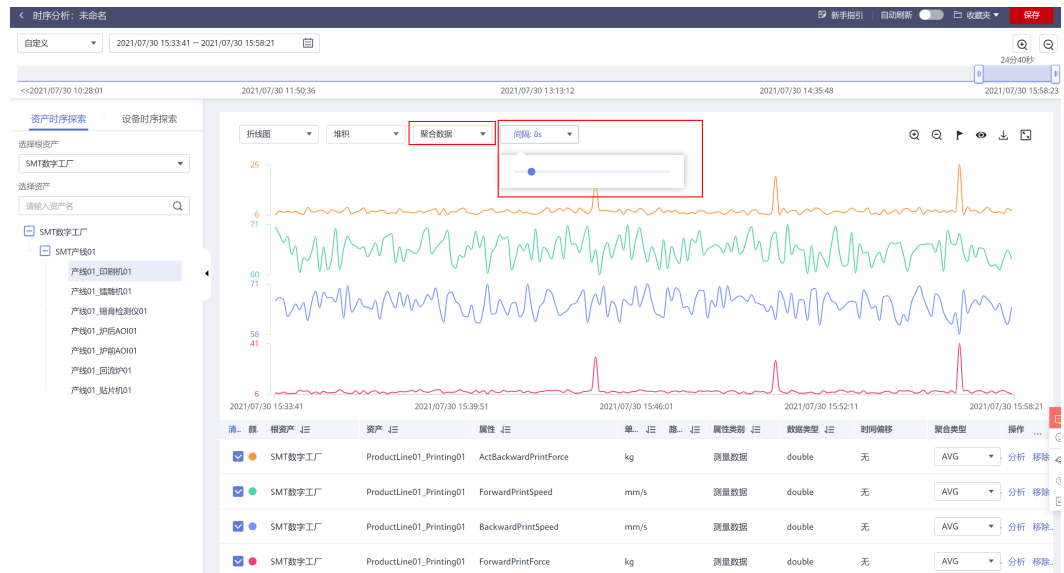
图 8-13 设置折线图 Y 轴显示



设置采样间隔

在图表分析区，选择“聚合数据”后，可根据实际的分析场景设置不同的采样间隔时间，设置完成后，时序分析会返回所选属性按采样间隔时间做聚合计算后的取值结果，并做图表呈现。

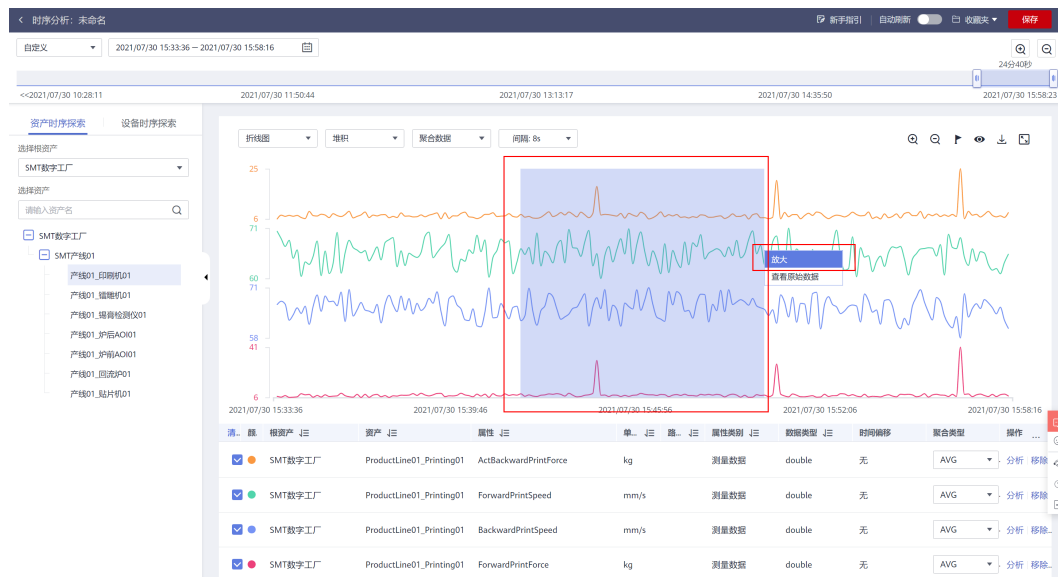
图 8-14 设置采样间隔



图表缩放

在图表分析区，当您需要时序数据做上卷下钻分析时，可以通过放大、缩小按钮或者在图表上选中指定区域放大来实现。

图 8-15 图表缩放

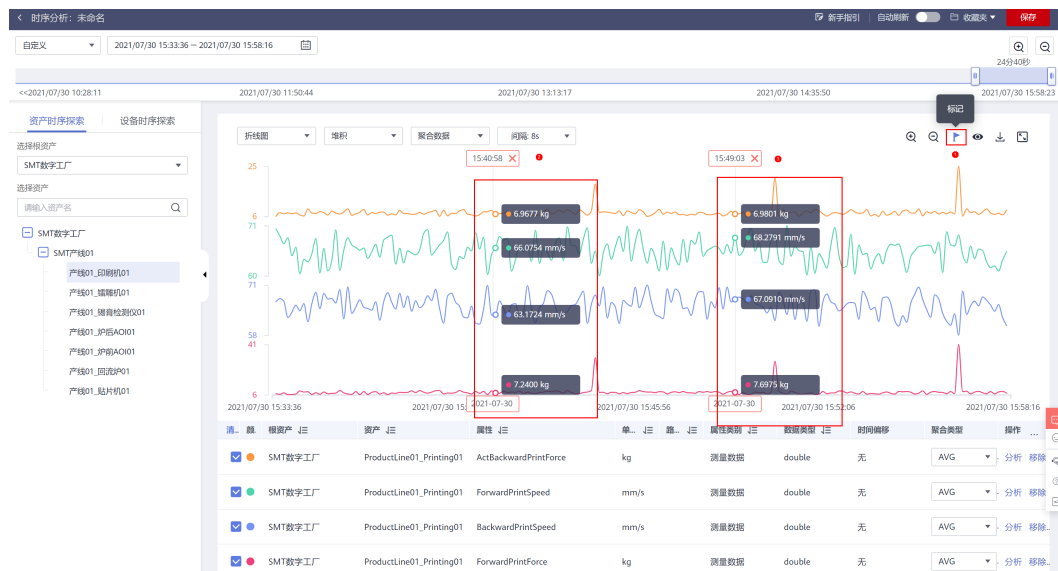


图表标记

在图表分析区，支持您对指定时序数据点做标记，以方便您做探索分析，当前仅折线图支持标记功能。

标记功能需要在点击标记图标后开启，开启后在图表区域单击即可完成标记动作，若需关闭标记功能，则再次点击标记图标即可。

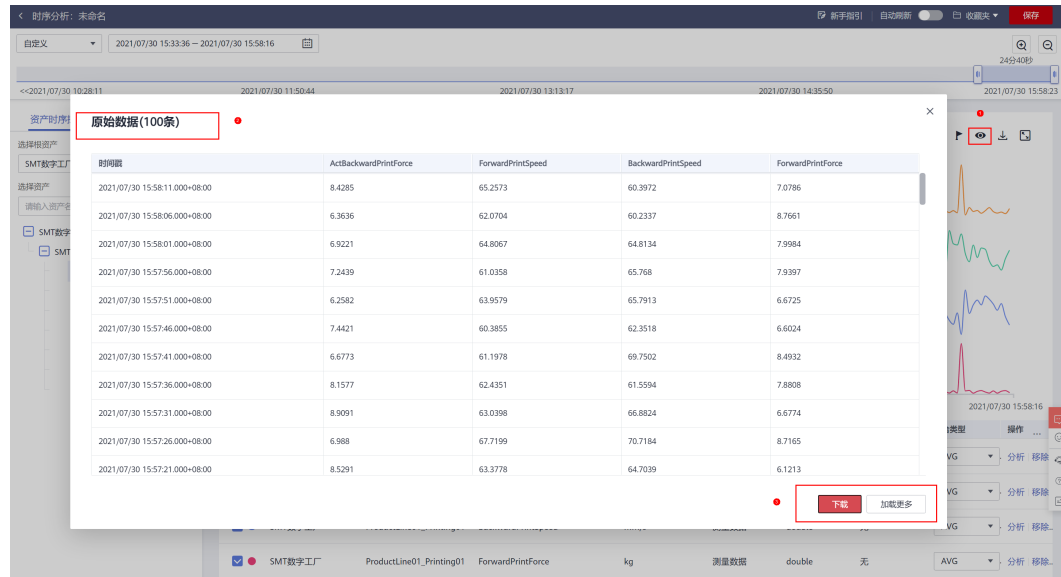
图 8-16 图表标记



浏览原始数据

在图表分析区，支持浏览并下载所选属性在指定时间窗的原始时序数据，并支持对原始数据做下载操作，当前最多支持浏览和下载2000条，下载格式为txt。

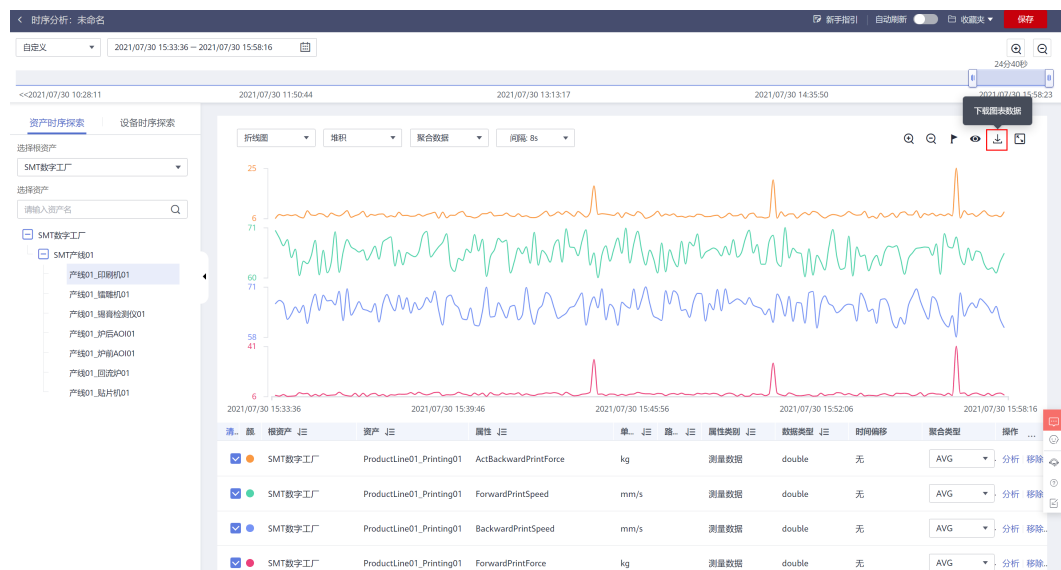
图 8-17 浏览原始数据



下载图表数据

在图表分析区，支持下载属性的时序聚合数据，下载格式为txt。

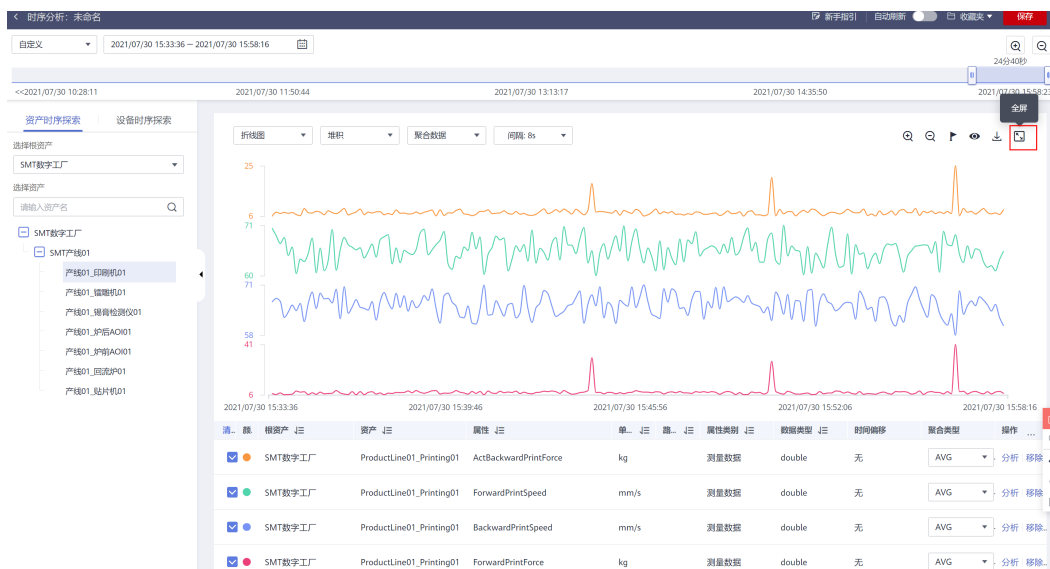
图 8-18 下载图表数据



图表全屏显示

图表分析区支持全屏显示，最大限度利用屏幕空间，提升探索分析的易用性。

图 8-19 图表全屏显示

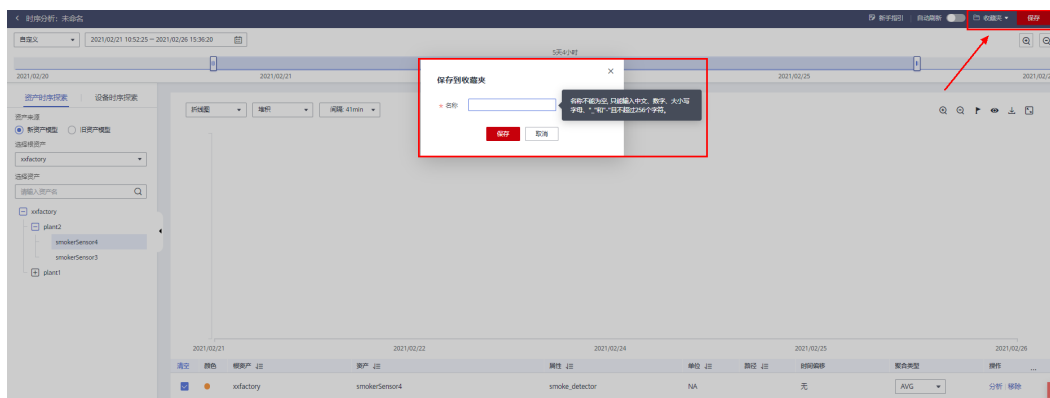


8.5 收藏夹

保存收藏夹

时序分析的结果支持保存到收藏夹，方便下次快捷打开，不同的收藏夹不能重名。

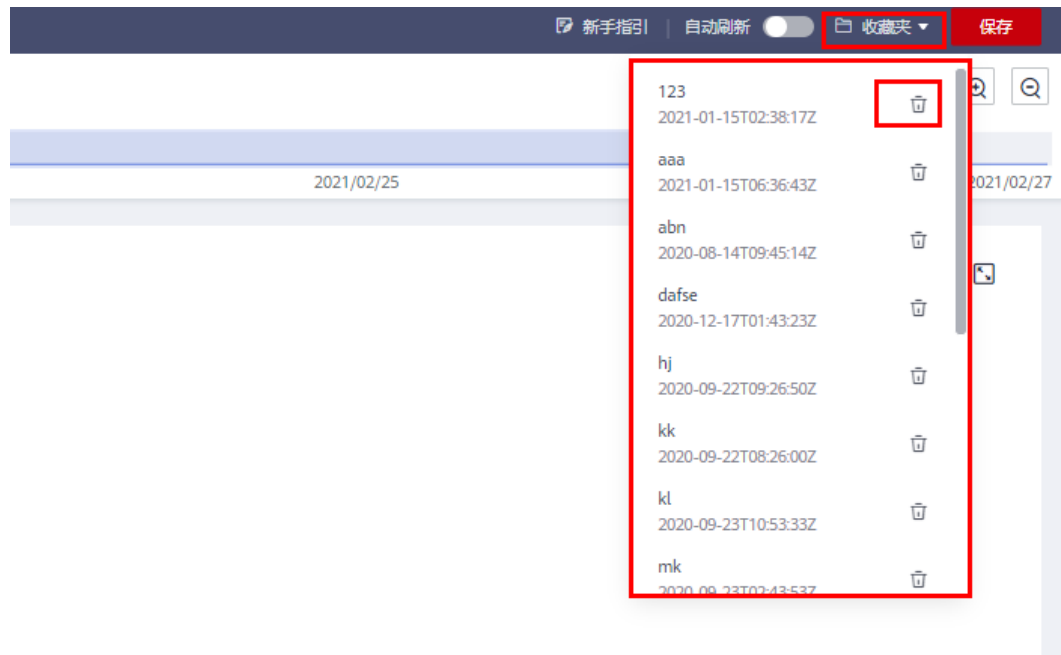
图 8-20 保存收藏夹



查看和删除收藏夹

已经保存的收藏夹可直接点击查看、或删除。

图 8-21 查看或删除收藏夹



8.6 自动刷新

自动刷新

当您需要对所选属性做最新固定长度时间窗的时序洞察分析时，可以用到自动刷新功能，开启自动刷新开关后每5秒自动刷新一次，并返回最新时序分析结果。

图 8-22 自动刷新



9 审计

9.1 支持云审计的关键操作

操作场景

平台提供了云审计服务。通过云审计服务，您可以记录与云服务器相关的操作事件，便于日后的查询、审计和回溯。

前提条件

已开通云审计服务。

支持审计的关键操作列表

表 9-1 数据源支持审计的关键操作列表

操作名称	资源类型	事件名称
创建数据源	datasource	createDataSource
修改数据源	datasource	updateDataSource
删除数据源	datasource	deleteDataSource
验证数据源	datasource	validateDataSource

表 9-2 资产模型支持审计的关键操作列表

操作名称	资源类型	事件名称
创建资产	asset	createAsset
修改资产	asset	updateAsset
删除资产	asset	deleteAsset

操作名称	资源类型	事件名称
创建测点	measurePoint	createMeasurePoint
修改测点	measurePoint	updateMeasurePoint
删除测点	measurePoint	deleteMeasurePoint
创建视图	view	createView
修改视图	view	updateView
删除视图	view	deleteView
刷新默认视图	view	refreshDefaultView
发布视图	view	publishView
另存视图	view	saveAsView
创建资产模板	assetTemplate	createAssetTemplate
修改资产模板	assetTemplate	updateAssetTemplate
删除资产模板	assetTemplate	deleteAssetTemplate
创建资产分类	assetCategory	createAssetCategory
修改资产分类	assetCategory	updateAssetCategory
删除资产分类	assetCategory	deleteAssetCategory
创建测点分类	measureCategory	createMeasureCategory
修改测点分类	measureCategory	updateMeasureCategory
删除测点分类	measureCategory	deleteMeasureCategory

表 9-3 时序分析支持审计的关键操作列表

操作名称	资源类型	事件名称
更新温数据存储配置	warmStorage	updateWarmStorage
保存时序分析收藏夹	tsiFavorites	saveTsiFavorites
删除时序分析收藏夹	tsiFavorites	deleteTsiFavorites

表 9-4 实时分析支持审计的关键操作列表

操作名称	资源类型	事件名称
创建作业	realTimeJob	createJob
更新作业	realTimeJob	updateJob

操作名称	资源类型	事件名称
删除作业	realTimeJob	deleteJob
启动作业	realTimeJob	startJob
停止作业	realTimeJob	stopJob
启动测试	realTimeJob	startDebug
停止测试	realTimeJob	stopDebug
上传测试文件	realTimeJob	uploadInputs

9.2 如何查看审计日志

操作场景

在您开启了云审计服务后，系统开始记录云服务器的相关操作。云审计服务管理控制台保存最近7天的操作记录。

本节介绍如何在云审计服务管理控制台查看最近7天的操作记录。

操作步骤


1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“服务列表”，选择“管理与部署 > 云审计服务”，进入云审计服务信息页面。
4. 单击左侧导航树的“事件列表”，进入事件列表信息页面。
5. 事件列表支持通过筛选来查询对应的操作事件。当前事件列表支持四个维度的组合查询，详细信息如下：
 - 事件来源、资源类型和筛选类型。
在下拉框中选择查询条件。
其中筛选类型选择事件名称时，还需选择某个具体的事件名称。
选择资源ID时，还需选择或者手动输入某个具体的资源ID。
选择资源名称时，还需选择或手动输入某个具体的资源名称。
 - 操作用户：在下拉框中选择某一具体的操作用户，此操作用户指用户级别，而非租户级别。
 - 事件级别：可选项为“所有事件级别”、“normal”、“warning”、“incident”，只可选择其中一项。
 - 时间范围：可选择查询最近七天内任意时间段的操作事件。
6. 展开需要查看的事件，查看详细信息。

图 9-1 展开记录

事件名称	资源类型	事件来源	资源ID	资源名称	事件级别	操作用户	事件记录时间	操作
login	user	IAM	26e96eda1		normal		2017/06/29 10:22:32 GMT+0...	查看详情
事件ID:	ce90cce3-5c71-11e7-910d-57ac1cd228ee		源IP地址:					
备注信息:	--		事件产生时间:	2017/06/29 10:20:52 GMT+08:00				

7. 在需要查看的记录右侧，单击“查看事件”，弹窗中显示了该操作事件结构的详细信息。
更多关于云审计的信息，请参见[《云审计服务用户指南》](#)。