

解决方案实践

端边云 Serverless 大数据湖解决方案

文档版本 1.0
发布日期 2023-08-03



版权所有 © 华为技术有限公司 2024。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。

安全声明

漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

目录

1 方案概述	1
2 资源和成本规划	3
3 实施步骤	5
3.1 准备工作.....	5
3.2 快速部署.....	13
3.3 开始使用.....	20
3.4 快速卸载.....	24
4 附录	26
5 修订记录	27

1 方案概述

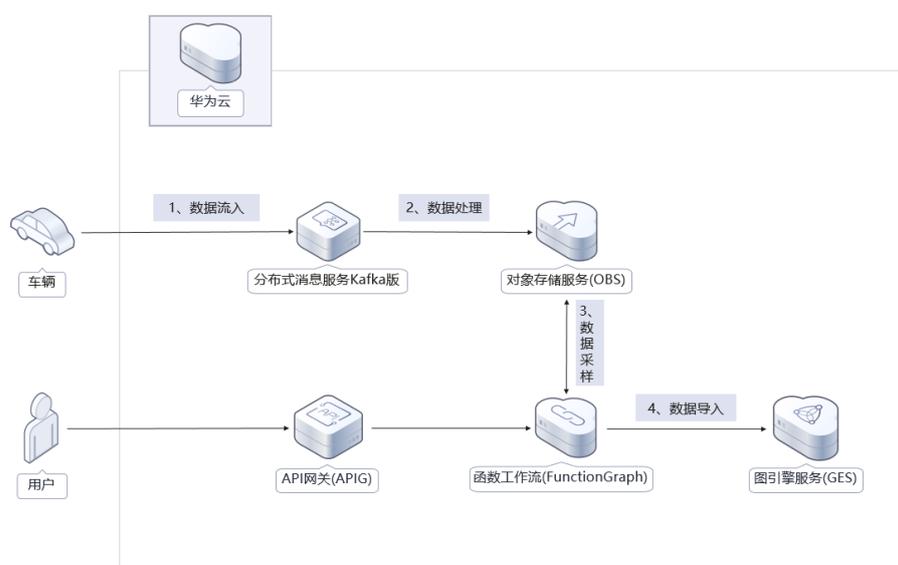
应用场景

该解决方案基于华为云ServerLess技术架构，通过联合伙伴EMQ IoT、华为云**函数工作流 FunctionGraph**、**对象存储服务 OBS**、**分布式消息服务Kafka版**、**图引擎服务 GES**等服务的组合，实现车联网数据的端侧灵活采集/传输、云端接入/分析/转换，在云端形成Serverless架构数据湖，并以图谱方式进行车辆故障分析。适用于各种车辆故障排查场景。

方案架构

该解决方案基于FunctionGraph、OBS、KAFKA、GES等服务的组合，辅助运维人员定位车辆故障原因。

图 1-1 方案架构图



该解决方案将会部署如下资源：

- 创建分布式消息服务 Kafka版专享版实例，并创建 Topic，用于写入车辆数据，为用户提供便捷高效的消息队列。
- 创建一个对象存储服务 OBS桶，用于存储车辆数据。
- 使用函数工作流 FunctionGraph创建3个函数，用于数据转换、导入数据到GES。
- 创建一个API网关 APIG API，将车辆数据生成GES故障图封装为一个API用户调用。
- 创建一个图引擎服务 GES图，用于生成时序图谱，查询关联事件，定位故障点位。

方案优势

- 可视化分析
基于GES图数据引擎，可视化车辆时序状态，加速问题排查和定位。
- 降本增效
基于FunctionGraph按需自动化加工数据，降低人工处理成本，FunctionGraph及OBS等Serverless服务按需计费，费用低。
- 一键部署
一键轻松部署，即可完成基于GES的故障图分析的应用构建。

约束与限制

- 该解决方案部署前，需注册华为账号并开通华为云，完成实名认证，且账号不能处于欠费或冻结状态。
- 卸载解决方案前，请先确保OBS桶中无数据，否则解决方案将卸载失败。

2 资源和成本规划

该解决方案主要部署如下资源，以下费用仅供参考，具体请参考华为云官网[价格详情](#)，实际收费以账单为准。

表 2-1 成本预估（仅供参考）

华为云服务	计费说明	每月花费（调用100万次）
函数工作流 FunctionGraph	<ul style="list-style-type: none">区域：华北-北京四产品：函数请求次数： 0-100万次：0元/100万次 100万次以上：1.33元/100万次计量时间： 0-400,000 GB/秒：0元/GB-秒 400,000 GB/秒以上：0.00011108元/GB-秒	0元
对象存储服务 OBS	<ul style="list-style-type: none">区域：华北-北京四存储空间：数据存储（多AZ存储）默认存储类别：标准存储桶策略：私有请求费用：0.0100元/万次存储空间：0.1390元/GB/月流量费用：<ul style="list-style-type: none">公网流出流量 / 00:00-08:00（闲时） 0.2500元/GB公网流出流量 / 08:00-24:00（忙时） 0.5000元/GB 费用包括存储空间、请求费用、流量费用两部分，具体请参考 OBS计费详情 。	费用包括存储空间、请求费用、流量费用两部分，详细请参考每月账单。

华为云服务	计费说明	每月花费（调用100万次）
分布式消息服务Kafka版	<ul style="list-style-type: none">● 区域：华北-北京四● 按需计费：4.62元/小时● 规格：kafka.2u4g.cluster.small● 代理数：3个● 存储空间：超高IO 300GB● 购买时长：720小时● 购买量：1	3326.40 元
API网关（共享版）	<ul style="list-style-type: none">● 区域：华北-北京四● 计费项：<ul style="list-style-type: none">● API调用次数：月累计超过次数：0~1000万次（含）0.06 元/万次● 流量：（公网流出流量）0.8 元/GB	费用包含调用次数、数据流量两部分，详细请参考每月账单。
图引擎服务GES	<ul style="list-style-type: none">● 区域：华北-北京四● 按需计费：1元/小时● 规格：一万边● 购买时长：720小时● 购买量：1	720.00 元
合计	-	4046.40+OBS费用+APIG费用

3 实施步骤

- 3.1 准备工作
- 3.2 快速部署
- 3.3 开始使用
- 3.4 快速卸载

3.1 准备工作

创建 rf_admin_trust 委托（可选）

步骤1 进入华为云官网，打开[控制台管理](#)界面，鼠标移动至个人账号处，打开“统一身份认证”菜单。

图 3-1 控制台管理界面



图 3-2 统一身份认证菜单



步骤2 进入“委托”菜单，搜索“rf_admin_trust”委托。

图 3-3 委托列表



- 如果委托存在，则不用执行接下来的创建委托的步骤
- 如果委托不存在时执行接下来的步骤创建委托

步骤3 单击步骤2界面中的“创建委托”按钮，在委托名称中输入“rf_admin_trust”，委托类型选择“云服务”，选择“RFS”，单击“下一步”。

图 3-4 创建委托

委托 / 创建委托

* 委托名称

* 委托类型 普通帐号
将帐号内资源的操作权限委托给其他华为云帐号。
 云服务
将帐号内资源的操作权限委托给华为云服务。

* 云服务

* 持续时间

描述

0/255

步骤4 在搜索框中输入“Tenant Administrator”权限，并勾选搜索结果。

图 3-5 选择策略

委托“rf_admin_trust”将策略与资源策略

策略名称: Tenant Administrator

名称	类型
Tenant Administrator	系统角色

步骤5 选择“所有资源”，并单击下一步完成配置。

图 3-6 设置授权范围

根据当前选择的策略，策略赋予以下授权范围方案，更便于您最小化授权，可进行选择。了解如何保留的应用场景选择最佳的授权范围方案

选择授权范围方案

所有资源
授权后，IAM用户可以按照权限使用帐号中所有资源，包括企业项目、区域项目和全局服务资源。

[展开其他方案](#)

步骤6 “委托”列表中出现“rf_admin_trust”委托则创建成功。

图 3-7 委托列表



----结束

创建 IAM Agency Management FullAccess 策略（可选）

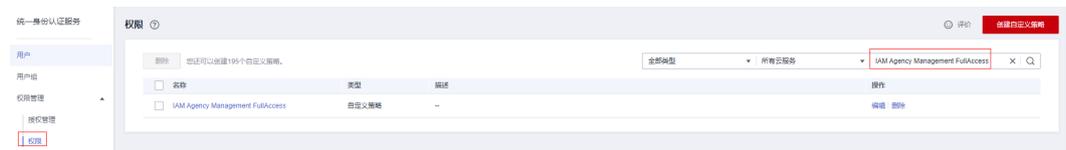
步骤1 打开“统一身份认证”菜单

图 3-8 统一身份认证菜单



步骤2 进入“权限管理”->“权限”菜单，在搜索框输入“IAM Agency Management FullAccess”当前账号是否存在IAM委托管理权限

图 3-9 权限列表



- 如果搜索结果不为空，则当前账号已经存在IAM委托管理权限，不需要重复创建
- 如果过搜索结果为空，则继续创建“IAM Agency Management FullAccess”权限

步骤3 单击“创建自定义策略”按钮

图 3-10 创建自定义策略



步骤4 输入策略名称为“IAM Agency Management FullAccess”，选择“JSON视图”，在策略内容中输入如下JSON代码，单击确认按钮

图 3-11 创建自定义策略

* 策略名称

策略配置方式 可视化视图 JSON视图

* 策略内容

```
1 {
2   "Version": "1.1",
3   "Statement": [
4     {
5       "Effect": "Allow",
6       "Action": [
7         "iam:agencies:updateAgency",
8         "iam:permissions:listRolesForAgencyOnDomain",
9         "iam:permissions:revokeRoleFromAgencyOnDomain",
10        "iam:permissions:listRolesForAgency",
11        "iam:permissions:checkRoleForAgencyOnProject",
12        "iam:roles:listRoles",
13        "iam:agencies:deleteAgency",
14        "iam:permissions:checkRoleForAgency",
15        "iam:permissions:listRolesForAgencyOnProject",
16        "iam:permissions:checkRoleForAgencyOnDomain",
17        "iam:agencies:listAgencies",
18        "iam:permissions:grantRoleToAgencyOnDomain",
19        "iam:permissions:revokeRoleFromAgencyOnProject",
20        "iam:agencies:getAgency",
21        "iam:agencies:createAgency",
22        "iam:permissions:grantRoleToAgency",
23        "iam:permissions:grantRoleToAgencyOnProject",
24        "iam:permissions:revokeRoleFromAgency"
25      ]
26     }
27   ]
28 }
```

策略描述

作用范围 全局级服务

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Action": [
        "iam:agencies:createAgency",
        "iam:agencies:listAgencies",

```

```
"iam:agencies:getAgency",  
"iam:agencies:deleteAgency",  
"iam:agencies:updateAgency",  
"iam:permissions:revokeRoleFromAgencyOnProject",  
"iam:permissions:revokeRoleFromAgencyOnDomain",  
"iam:permissions:revokeRoleFromAgency",  
"iam:permissions:grantRoleToAgencyOnDomain",  
"iam:permissions:grantRoleToAgencyOnProject",  
"iam:permissions:grantRoleToAgency",  
"iam:permissions:listRolesForAgencyOnDomain",  
"iam:permissions:listRolesForAgencyOnProject",  
"iam:permissions:checkRoleForAgencyOnDomain",  
"iam:permissions:checkRoleForAgencyOnProject",  
"iam:permissions:listRolesForAgency",  
"iam:permissions:checkRoleForAgency",  
"iam:roles:listRoles"  
  ],  
  "Effect": "Allow"  
}  
]
```

步骤5 界面无报错，则成功创建IAM Agency Management FullAccess权限

----结束

给 rf_admin_trust 委托添加 IAM Agency Management FullAccess 策略（可选）

步骤1 打开“统一身份认证”菜单

图 3-12 统一身份认证菜单



步骤2 进入“委托”菜单，选择rf_admin_trust委托

图 3-13 委托列表



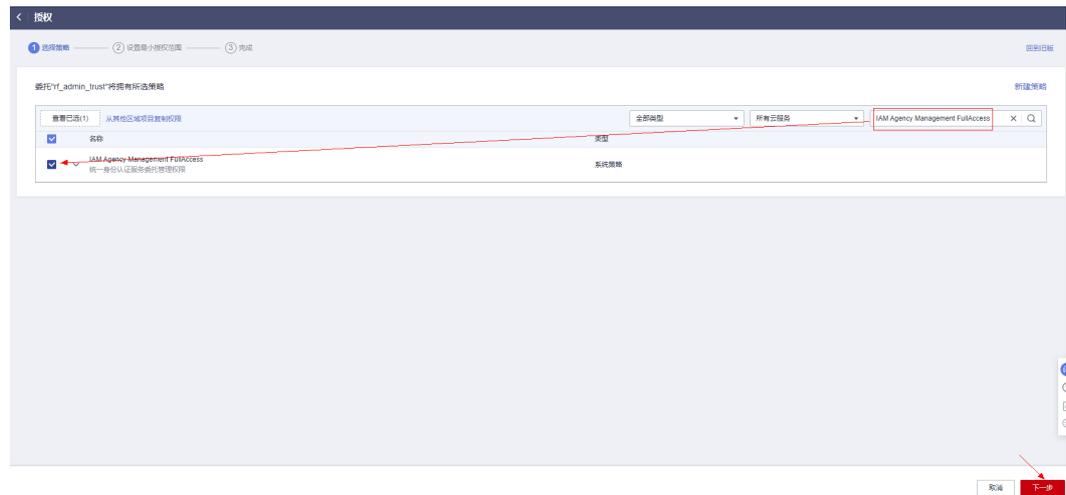
步骤3 进入“授权记录”菜单，单击“授权”按钮

图 3-14 授权记录



步骤4 在搜索框输入IAM Agency Management FullAccess，勾选过滤出来的记录，单击下一步，并确认完成权限的配置

图 3-15 配置 IAM Agency Management FullAccess 策略



步骤5 配置好后的情况：rf_admin_trust委托拥有Tenant Administrator和IAM Agency Management FullAccess权限

图 3-16 授权记录列表



----结束

获取子网 ID

本章节主要帮助用户在快速部署该解决方案之前，获取部分依赖的资源，以供一键部署时使用。

步骤1 登录[华为云官网控制台](#)，单击[虚拟私有云VPC](#)，选择该方案资源需要部署的VPC，单击该VPC下的[子网](#)，获取网络ID。

图 3-17 子网列表



图 3-18 子网网络 ID



----结束

3.2 快速部署

本章节主要帮助用户快速部署“端边云车联网Serverless数据湖”解决方案。

表 3-1 参数填写说明

参数名称	类型	是否必填	参数解释	默认值
subnet_id	String	必填	子网ID, 该模板使用已有子网, 该方案所涉及到的其他云服务资源将会部署在该子网下, 查询子网ID请参考 获取子网ID 。	空

参数名称	类型	是否必填	参数解释	默认值
secgroup_name	String	必填	安全组名称，该模板新建安全组，安全组规则请参考 安全组规则修改（可选） 或者参考 Kafka帮助文档 修改安全组。取值范围：1-64个字符，支持数字、字母、中文、_(下划线)、-(中划线)、.(点)。	device-edge-cloud-serverless-datalake-demo
obs_bucket_name	String	必填	对象存储服务OBS桶名称，用于存放应用数据，全局唯一。取值范围：长度1~59个字符，以小写字母或数字开头，仅支持小写字母、数字、中划线(-)、点(.)。	空
kafka_name	String	必填	分布式消息服务Kafka实例名称，取值范围：以字母开头，长度为4-64个字符，仅支持字母、数字、连字符(-)和下划线(_)。	device-edge-cloud-serverless-datalake-demo
kafka_flavor_id	String	必填	Kafka实例规格，可选值： c6.2u4g.cluster、 c6.4u8g.cluster、 c6.8u16g.cluster、 c6.12u24g.cluster、 c6.16u32g.cluster。	c6.2u4g.cluster

参数名称	类型	是否必填	参数解释	默认值
kafka_storage_spec	String	必填	Kafka实例存储空间磁盘类型 ，创建完成后不可修改。请根据业务IO要求选择。可选值： dms.physical.storage.high.v2(高I/O磁盘类型)、 dms.physical.storage.ultra.v2(超高I/O磁盘类型)。	dms.physical.storage.ultra.v2
kafka_storage_space	Number	必填	Kafka实例存储空间容量，存储空间不支持扩容。存储空间主要用于存储消息（包含副本，Kafka默认使用3副本），除了存储消息外还需要预留部分空间用于存储日志和元数据。例如：业务消息体积预估100GB，则存储空间最少应为 100GB * 副本数 + 预留磁盘大小 100GB。取值需为300的整数倍，范围： 300-900,000。	300
kafka_manager_password	String	必填	Kafka实例Manager用户密码，创建完成后，可参考部署指南登录Kafka控制台修改密码。取值范围：8-32个字符，密码至少必须包含大写字母、小写字母、数字和特殊字符`~!@#\$%^&*()-_+= [];:\",<.>?和空格中的三种，并且不能以-开头。不能与用户名或倒叙的用户名相同。	空

参数名称	类型	是否必填	参数解释	默认值
functiongraph_name	String	必填	函数工作流functiongraph函数名称，取值范围：长度为2-47个字符，可包含字母、数字、_（下划线）和-（中划线），以字母开头，以字母或数字结尾。	device-edge-cloud-serverless-datalake-demo
api_name	String	必填	API名称，该方案使用共享版API网关，对外提供API服务。取值范围：3-244个字符，中英文字符开头，只能由中英文字符、数字、下划线组成（_），中文字符必须是UTF-8或Unicode格式。	device_edge_cloud_serverless_datalake_demo
ges_graph_name	String	必填	图形名称，取值范围：长度为4-64个字符，可包含字母、数字、_（下划线），以字母开头。	device-edge-cloud-serverless-datalake-demo
graph_size_type	String	必填	图形大小类型索引，0：表示1万边、1：表示100万条边、2：表示1000万条边、3：表示1亿条边、4：表示10亿条边、5：表示100亿条边、6：表示数据库版本、401：表示10亿条增强边。	0

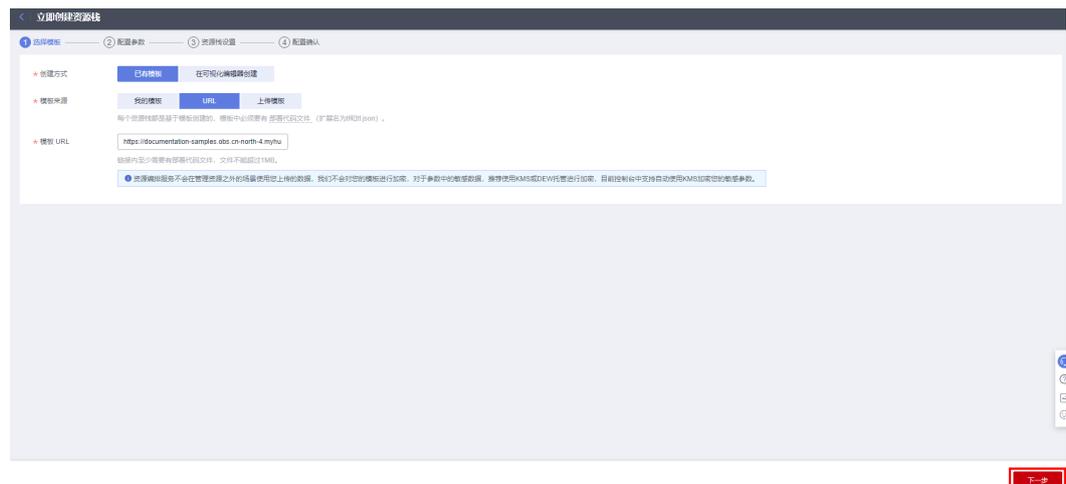
步骤1 登录[华为云解决方案实践](#)，选择“端边云车联网Serverless数据湖解决方案”，单击“一键部署”，跳转至解决方案创建资源栈界面。

图 3-19 解决方案实施库



步骤2 在选择模板界面中，单击“下一步”。

图 3-20 选择模板



步骤3 在配置参数界面中，参考表 3-1 完成自定义参数填写，单击“下一步”。

图 3-26 执行计划确认



步骤8 待“事件”中出现“Apply required resource success”，表示该解决方案已经部署完成。

图 3-27 部署完成



----结束

3.3 开始使用

安全组规则修改（可选）

须知

- Kafka默认使用9092端口，该方案默认针对子网网段开放，请参考[修改安全组规则](#)，配置IP地址白名单，以便能正常访问服务。

安全组实际是网络流量访问策略，包括网络流量入方向规则和出方向规则，通过这些规则为安全组内具有相同保护需求并且相互信任的云服务器、云容器、云数据库等实例提供安全保护。

如果您的实例关联的安全组策略无法满足使用需求，比如需要添加、修改、删除某个TCP端口，请参考以下内容进行修改。

- 添加安全组规则：根据业务使用需求需要开放某个TCP端口，请参考[添加安全组规则](#)添加入方向规则，打开指定的TCP端口。

- 修改安全组规则：安全组规则设置不当会造成严重的安全隐患。您可以参考[修改安全组规则](#)，来修改安全组中不合理的规则，保证云服务器等实例的网络安全。
- 删除安全组规则：当安全组规则入方向、出方向源地址/目的地址有变化时，或者不需要开放某个端口时，您可以参考[删除安全组规则](#)进行安全组规则删除。

步骤1 查看Kafka。在控制台单击“服务列表”，选择“[分布式消息服务Kafka版](#)”，在“Kafka专享版”页面找到该方案所创建的Kafka实例。单击进入实例管理控制台。

图 3-28 Kafka 实例

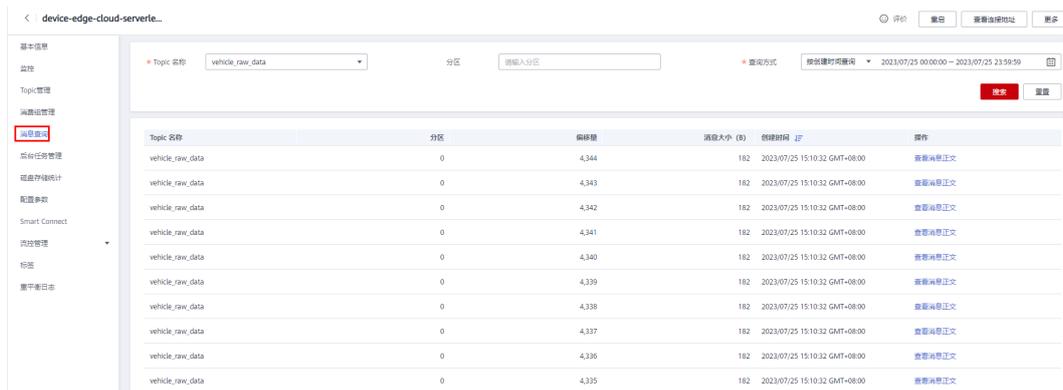


步骤2 查看Topic。单击Topic管理，向Topic: vehicle_raw_data 中写入消息，请参考[Kafka 开发指南-python文档](#)，此Topic默认老化时间72小时，到期Topic会自动消失，如果逾期请手动创建以vehicle_raw_data命名的Topic。

图 3-29 Topic



图 3-30 查看消息



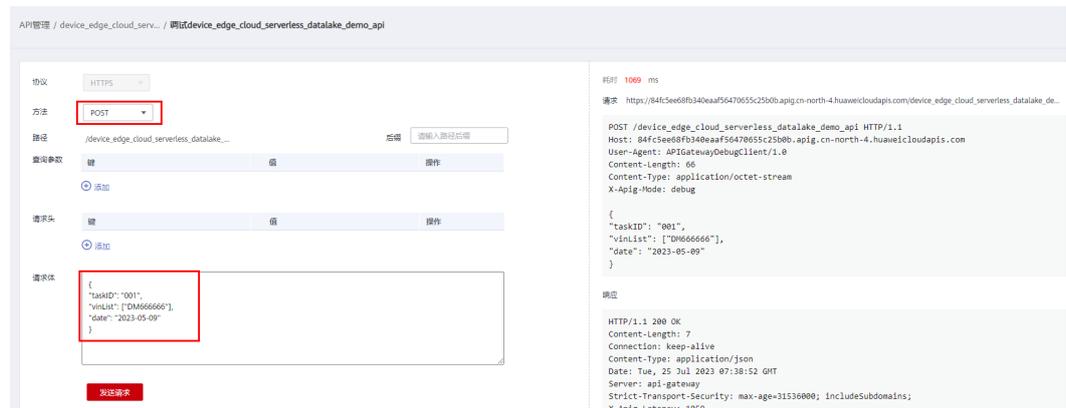
步骤3 进入API调试界面。进入华为云[共享版API网关控制台 API管理界面](#)，查找新建的API依次单击“更多”→“调试”。

图 3-31 API 管理界面



步骤4 API调用。方法选用“POST”，在请求体中填入请求体body内容，单击发送请求。

图 3-32 API 调用



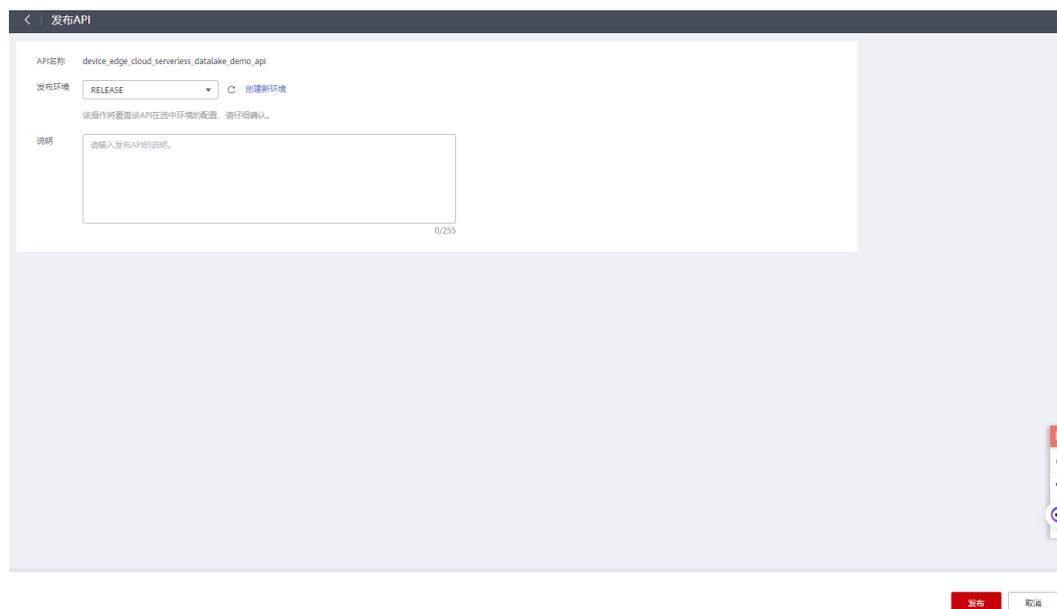
参数说明：
taskID: 为任务ID, 即本次转换的任务ID;
vinList: 为本次任务指定转换的车架号列表, 例如值为["DM666666"], 将DM666666车架号的原始数据转换为csv格式上传至GES;
date: 为本次任务指定转换日期的车辆数据, 例如值为2023-05-09, 把DM666666车架号2023-05-09原始数据转换为csv格式上传至GES;
hours: 为本次任务指定日期下更细精度的小时数, 不传则解析全天的原始数据, 例如值为["01","02","03","04","05","06"], 把DM666666车架号2023-05-09 01-06小时的原始数据转换为csv格式上传至GES。

步骤5 (可选) 发布API。进入华为云[共享版API网关控制台 API管理界面](#), 查找新建的API单击“发布”, 并在API发布页面中, 单击“发布”发布API。

图 3-33 进入发布 API 页面



图 3-34 发布 API



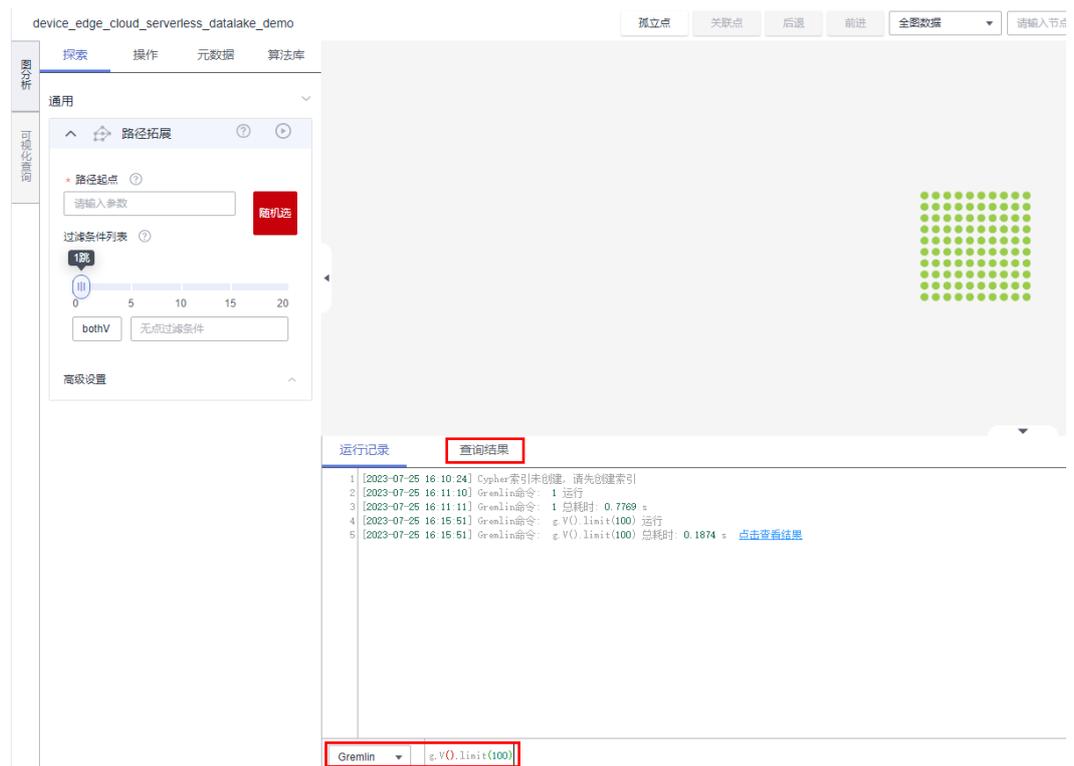
步骤6 登录华为云[图引擎服务 GES](#)，单击“图管理”，选择方案本方案创建的图，单击“操作”列的“访问”。

图 3-35 图管理



步骤7 在图引擎编辑器页面，在Gremlin查询区输入查询语句（请参考[Gremlin查询文档](#)获取），按“回车”键执行操作。可以通过查询语句获取点、边、属性等信息，并可以对数据进行修改、局部遍历和属性过滤等操作。

图 3-36 查询图数据



----结束

3.4 快速卸载

📖 说明

- 该解决方案涉及到对象存储服务 OBS桶，如果OBS桶中有数据的话会导致资源栈删除失败。请根据需要迁移备份后清空OBS桶中的数据，再卸载该解决方案。
- 如果按开始使用**步骤5 发布API**，发布了API，在删除资源栈前，需要将发布的API下线，如果未下线会导致资源栈删除失败。

步骤1 解决方案部署成功后，单击该方案堆栈后的“删除”。

图 3-37 一键卸载



步骤2 在弹出的删除堆栈确定框中，输入Delete，单击“确定”，即可卸载解决方案。

图 3-38 删除堆栈确定



----结束

4 附录

名词解释

- **分布式消息服务Kafka版**：分布式消息服务Kafka版是一个高吞吐、高可用的消息中间件服务，适用于构建实时数据管道、流式数据处理、第三方解耦、流量削峰去谷等场景，具有大规模、高可靠、高并发访问、可扩展且完全托管的特点，是分布式应用上云必不可少的重要组件。
- **函数工作流 FunctionGraph**：是一项基于事件驱动的函数托管计算服务。使用 FunctionGraph 函数，只需编写业务函数代码并设置运行的条件，无需配置和管理服务器等基础设施，函数以弹性、免运维、高可靠的方式运行。此外，按函数实际执行资源计费，不执行不产生费用。
- **对象存储服务 OBS**：一个基于对象的海量存储服务，为客户提供海量、安全、高可靠、低成本的数据存储能力。
- **API网关 APIG** 是为企业和开发者提供的高性能、高可用、高安全的云原生网关服务，融合安全、负载均衡、流量入口治理、微服务流量治理、运维等多项能力，帮助企业轻松实现API安全开放、API高并发调用和入口流量、微服务流量精细化治理，简单、快速、低成本、低风险地实现内部系统集成和业务能力开放变现。
- **图引擎服务 GES**，是针对以“关系”为基础的“图”结构数据，进行查询、分析的服务。广泛应用于社交应用、企业关系分析、风控、推荐、反欺诈等具有丰富关系数据的场景。

5 修订记录

表 5-1 修订记录

发布日期	修订记录
2023-07-30	第一次正式发布。