

数据工坊(DWR)

快速入门

文档版本 01
发布日期 2022-02-14



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。

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 准备工作	1
2 处理数据	2
2.1 入门流程.....	2
2.2 示例：图片质量变换.....	2
3 发布算子	7
4 修订记录	12

1 准备工作

注册华为账号并开通华为云，再完成实名认证

如果您已有一个华为账号，请跳到下一个任务。如果您还没有华为账号，请参考以下步骤创建。

1. 根据提示信息完成注册。
注册成功后，系统会自动跳转至您的个人信息界面。
2. 关于注册华为账号并开通华为云，请参见[注册华为账号并开通华为云](#)。
3. 请参考[实名认证](#)完成个人或企业账号实名认证。

为账户充值

您需要确保账户有足够金额。

- 关于DWR收费标准，请参见[计费说明](#)。
- 关于充值，请参见[如何给华为云账户充值](#)。

创建 IAM 用户

IAM用户由账号管理员创建并管理，可以确保账号及资源的安全性，有关IAM的详细介绍请参见[IAM用户](#)。此处介绍如何创建一个具有DWR访问权限的IAM用户。如果您不需要使用IAM用户，可以略过此部分内容。

1. 在华为云官网，单击页面右上角的“控制台”，使用注册账号登录控制台。
2. 在“控制台”页面，鼠标移动至右上方的用户名，在下拉列表中选择“统一身份认证”。
3. 创建用户组。
详细步骤请参见[创建用户组](#)，创建成功后，为该用户组配置Tenant Administrator权限。
4. 创建用户。
详细步骤请参见[创建用户](#)，创建用户时“所属用户组”需选择[步骤3](#)创建的用户组。

2 处理数据

2.1 入门流程

当需要使用DWR服务对OBS内存储的数据自动进行多项复杂任务（如解析、转码、截图等）处理时，建议参考[表2-1](#)逐步操作。

表 2-1 DWR 使用流程简介

操作步骤	操作说明
步骤一：准备工作	注册华为云账号，并开通数据工坊DWR服务。
（可选）步骤二：浏览算子市场	在算子市场查找需要的算子，了解已有算子的功能。
步骤三：创建 workflow	创建工作流对算子进行编排。
步骤四：启动 workflow	当前DWR支持同步和异步两种方式启动 workflow。 <ul style="list-style-type: none">• 创建事件触发器异步启动• 使用API同步启动• 使用API异步启动
步骤五：查看运行记录	查看历史运行记录，了解 workflow 的执行情况。

2.2 示例：图片质量变换

场景介绍

对于需要对大量图片文件进行压缩，快速发布上线的场景，可以使用DWR可视化的质量变换 workflow 来完成任务，帮助用户提升图片处理效率，节省内容发布时间。

实现流程

该实例通过创建触发器的方式异步启动 workflow，实现流程为：

1. 在OBS服务中创建桶
2. 在DWR服务中创建工作流
3. 在DWR服务中创建事件触发器
4. 在OBS服务中上传文件触发 workflow 执行

使用限制

当前支持使用的区域为：北京四、上海一、上海二、广州、新加坡。

操作步骤

步骤1 在OBS服务中创建桶

在OBS服务控制台创建两个桶，一个用于上传待处理的数据对象，一个用于存储处理后的数据对象。桶的详细介绍参见[创建桶](#)。

数据输入桶：“piccomp”

数据处理后输出桶：“piccomp-output”

说明

这里创建两个桶是为了防止无限循环。因为处理后的数据如果放在源桶，又跟事件触发器执行的条件匹配，会继续触发 workflow 执行。详细介绍参见[了解更多无限循环原因](#)。

步骤2 在DWR服务中创建工作流

1. 登录DWR控制台，在“工作流”页面单击“创建工作流”，进入“工作流编排”页面。
2. 将“质量变换”算子拖拽至编排区域进行串联编排，并填写相关属性配置信息，如[图2-1](#)所示。
 - quality_q代表图片按照原图的q%进行压缩，取值范围为[1, 100]，可根据实际情况配置。
 - quality_bucket为图片处理后的输出桶，这里设置为“piccomp-output”。详细介绍参见[质量变换参数说明](#)。
3. 保存并填写工作流名称“piccomp”。

图 2-1 编排 workflow

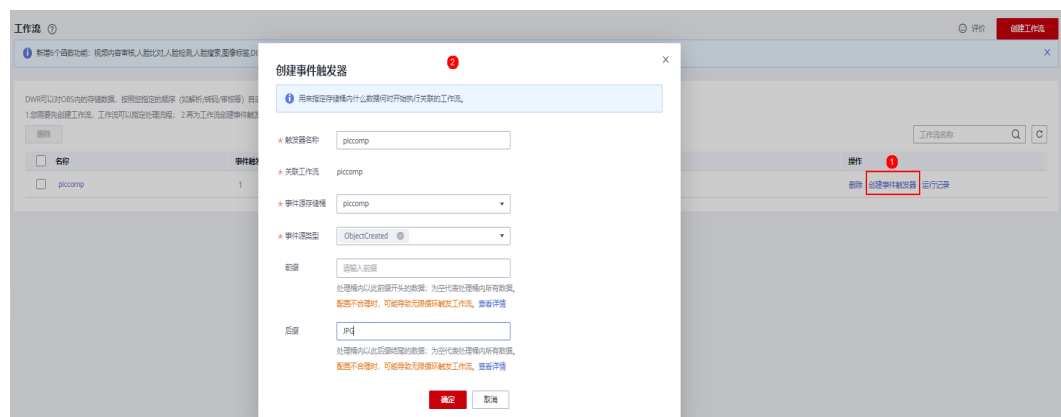


步骤3 在DWR服务中创建事件触发器，如图2-2所示。

在工作流列表中，单击工作流“piccomp”操作列的“创建事件触发器”进行触发器的创建，如图2-2所示。

事件源存储桶选择“piccomp”。

图 2-2 创建事件触发器



步骤4 上传文件触发 workflows 执行

登录OBS服务控制台，进入桶“piccomp”对象列表页面上上传图片，如图2-3所示。
上传图片至piccomp桶后OBS会自动生成事件触发 workflows 运行，将图片压缩。

说明

触发器执行存在一定的延时，大概5分钟生效。

图 2-3 上传文件



步骤5 查看图片压缩效果

压缩后的图片存放在piccomp-output桶的output文件夹中。如图2-4所示，图片大小均有变小，实现了对图片的压缩。

图 2-4 压缩图片



步骤6 （可选）查看运行记录

1. 在DWR控制台左侧导航栏选择“工作流”，进入“工作流”页面
2. 在工作流列表中，单击工作流“piccomp”操作列的“运行记录”可以查看函数流的执行记录。详细介绍参见[函数流执行历史管理](#)。

图 2-5 查看运行记录



----结束

3 发布算子

操作场景

所有开发者均可以创建用于数据处理的算子，审核通过后发布至DWR算子市场，开放给所有华为云用户使用。

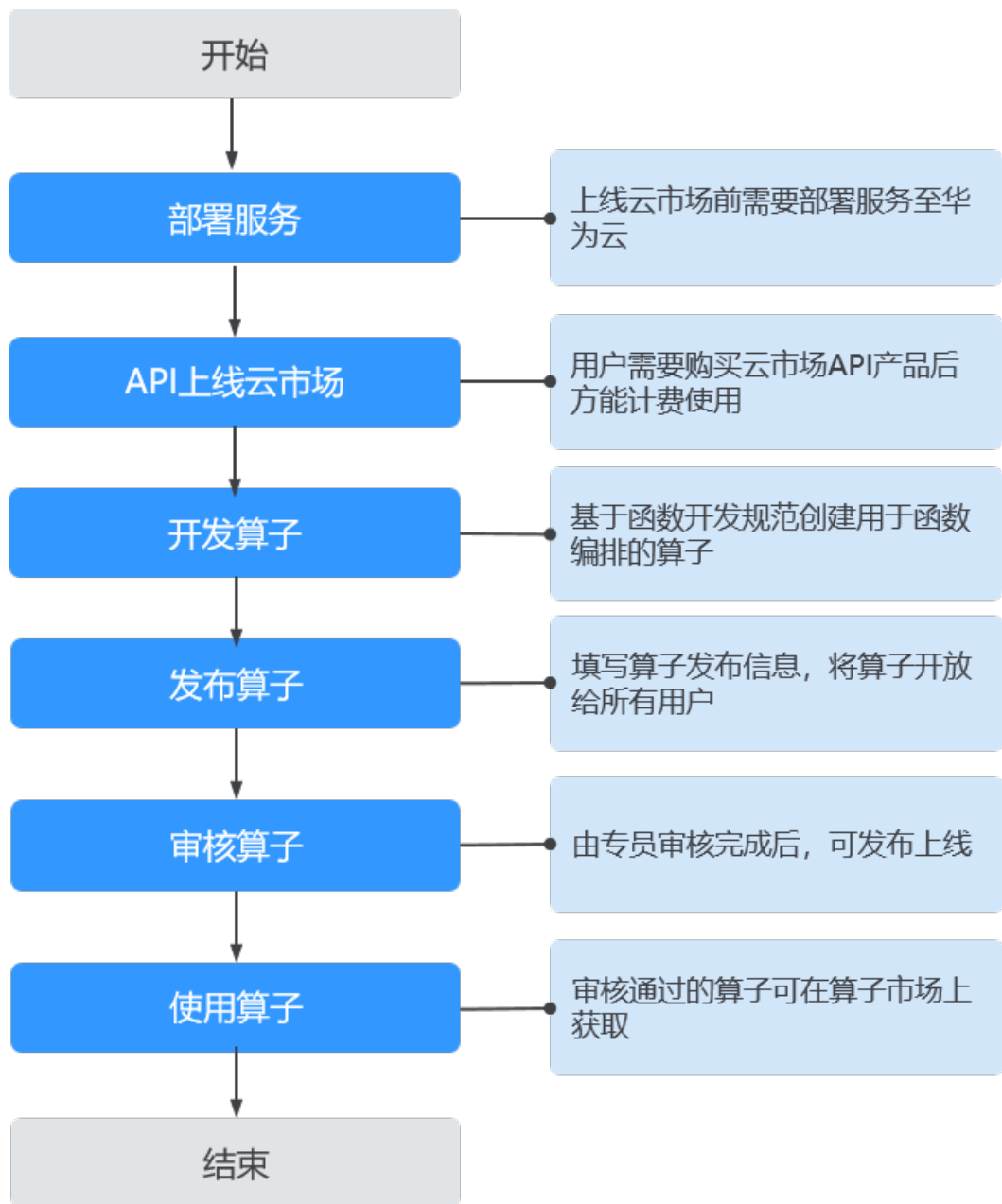
当算子被华为云用户使用之后，算子发布者将会获得相应的收益，详情参见[计费说明](#)。

算子从API上线、开发、到发布的整个流程如[图3-1](#)所示，本节主要介绍如何使用DWR发布已开发好的算子到算子市场。

说明

算子开发完成后，您可以参考[抽帧截图（自定义算子）](#)来验证算子。

图 3-1 算子发布流程



约束与限制

当前发布算子功能仅支持IAM主账号。

前提条件

算子已开发完成, 开发规范参考[自定义函数开发规范](#), 建议使用Go、Python开发(算子的冷启动效果更好)。

算子已在华为云云市场上架, 上架指导参考[发布API类商品操作指导](#)。

操作步骤


- 步骤1** 登录管理控制台。在左侧导航栏上方，单击 ，选择“存储 > 数据工坊DWR”。进入DWR页面。
- 步骤2** 在左侧导航栏选中“发布算子”，进入“发布算子”页面。
- 步骤3** 单击界面右上角的“发布公共算子”。
- 步骤4** 配置算子基本信息。

表 3-1 配置基本信息

参数	说明
算子名称	不能与本用户已有的算子重名。
算子提供方	-
算子描述	-
API链接	填写华为云云市场已上架的算子链接。
算子分类	根据算子市场上提供的分类进行选择。
算子logo	支持主流图片格式

图 3-2 算子基本信息

基本信息

中文信息 英文信息 (必填)

* 算子名称
Ⓞ 不能与本用户已有的算子重名

* 算子提供方

* 算子描述
0/256

* API链接

* 算子分类

* 算子Logo
+

- 步骤5** 请参考模板上传所需文件。

上传文件包括：算子代码包、中文和英文帮助文档、中文和英文开源声明、中文和英文服务协议文档、中文和英文测试报告。

图 3-3 上传文件

算子文件
请参考模板，分别上传以下算子文件。 [下载所有模板文件](#)

- * 算子代码包
- * 算子执行入口
格式为[文件名].[执行函数名]，不超过128个字符
- * 帮助文档
- * 服务协议
- * 开源声明
- * 测试报告

步骤6 配置算子参数。

参数	说明
Inputs参数（可选）	提供算子入参中动态参数对应的参数值，其Key值需要与动态参数中保持一致。
动态参数（可选）	提供算子入参的参数列表，可配置静态参数或动态参数，动态参数需要从Inputs参数中根据定义的Key取值。
权限版本	华为云统一身份认证（IAM）的权限版本，1.0版本以服务为粒度，提供有限的服务相关角色用于授权。1.1版本支持细粒度授权，可以精确到具体服务的操作、资源以及请求条件等。

步骤7 单击右下角的“提交审核”。

审核通过后，算子将发布至算子市场。您可以过滤第三方的算子提供方，查看您发布的算子。

----结束

算子参数配置示例

Inputs参数

Input结构体参数说明参见[创建工作流API](#)。

说明

regex参数设置的正则表达式请使用[regexexploit](#)工具校验。

```
[
{
  //算子所在工作流输入列表
  "parameter_name": "bucket",
  "parameter_value": "",
  "value_type": "",
  "default": "",
  "type": "string",
  "label": "Body",
  "constraints": {
```

```
    "regex": ".*"//正则表达式请使用regex工具校验
  },
  "invisible": false,
  "description": "doc destination bucket name"
}
]
```

动态参数

```
{
  "bucket": {
    "get_input": "$.inputs.bucket"//该值需要跟inputs参数中的parameter_name取值保持一致
  }
}
```

权限版本

obs授权参见[对象相关授权项](#)和[桶相关授权](#)。

```
[
  { //1.1版本支持细粒度授权，可以精确到具体服务的操作、资源以及请求条件等
    "version": "1.1",
    "statement": [
      {
        //对IAM用户组授予OBS指定资源的指定操作权限
        "action": [
          "obs:bucket:HeadBucket",
          "obs:bucket:ListBucketMultipartUploads",
          "obs:object:AbortMultipartUpload",
          "obs:object:PutObject",
          "obs:bucket:GetBucketAcl",
          "obs:object:GetObject"
        ]
      }
    ]
  }
]
```

4 修订记录

发布日期	修订记录
2022-02-14	第一次正式发布。