

Airflow使用指南

1 商品说明

Apache Airflow是一个用于编排、调度和监控工作流的平台。它允许用户以编程的方式定义复杂的工作流，这些工作流可以包含多个任务，并且这些任务之间可以有依赖关系。工作流在 Airflow 中被定义为有向无环图 (DAG, Directed Acyclic Graph)。DAG 由一系列的任务 (Task) 组成，任务之间通过定义好的依赖关系连接起来。

本商品通过鲲鹏服务器+ EulerOS2.0进行安装部署。

2 商品购买

您可以在云商店搜索“Airflow”购买(推荐)，其中，地域、规格、推荐配置使用默认，购买方式根据您的需求选择按需/按月/按年，短期使用推荐按需，长期使用推荐按月/按年，确认配置后点击“立即购买”。

下面以按需方式举例：

【自定义开通】和【模板配置开通】两种购买方式任选一种即可！

【自定义开通】方式：

工作流任务调度平台

现代数据编排平台。它能够以低代码创建高性能敏捷 workflows。

总计 **¥0.00** 原价 ¥0.00 费用明细

该镜像为固定定价，具体价格以ECS实例规格选配时的询价为准。

地域: 华北-北京四, 华东-上海一, 华南-广州, 华北-乌兰察布一, **西南-贵阳一**, 华北-北京一, 华北-北京二, 华南-深圳, 华北-乌兰察布-汽车一, 华南-广州-友好用户环境, 亚太-马尼拉 **展开**

规格: x86-v1.0, **EulerOS2.0**, Ubuntu24.04

推荐配置: 4核16G云主机_40G硬盘 (kc1.xlarge.4), 8核16G云主机_40G硬盘 (kc1.2xlarge.2), 8核32G云主机_40G硬盘 (kc1.2xlarge.4) **展开**

自定义云主机 > 支持同时选配弹性IP、公网带宽和数据盘

购买方式: 按需, 按月, 按年

立即购买

购买工作流任务调度平台

服务选型 2 配置

基本信息: 部署位置: 华为云, 区域: 华南-广州

模板配置:

模板名称	模板简介	创建时间
mm	mm-v1	2025/04/30 15:41:00 GMT+08:00

协议及授权: 请仔细阅读并同意以下协议和授权

- 《华为云高可用通用用户协议》
- 《工作流任务调度平台服务协议支持条款》

模板或已开通 **自定义开通**

购买弹性云服务器

基础配置: 计费模式: 包年包月, 按量计费, 竞价计费; 区域: 华南-广州; 可用区: 可用区5

实例: 规格类型: 业务场景选型; CPU架构: x86; 实例规格: kc2.xlarge.2 (4vCPUs, 8GB); 实例规格: kc1.xlarge.2 (4vCPUs, 8GB)

配置费用: **¥1,066 小时**

立即购买

操作系统

镜像 ①

公共镜像 私有镜像 共享镜像 **市场镜像**

3_1_EulerOS2.0... 任务调度平台(v1) [查看详情](#)

使用在云市场选用的第三方镜像需遵守第三方服务协议，请确保已阅读并同意《华为云市场通用商品用户协议》、《华为云市场通用商品支持条款》。

存储与备份

系统盘 ①

磁盘类型 **通用型SSD** 系统盘大小(GB) **200**

IOPS上限4,200, IOPS突发上限6,000 [加购](#) [来选购](#) [高级设置](#)

增加一块数据盘

您还可以挂载2块VBD磁盘或2块SCSI磁盘。

开启备份

备份可以帮您您在服务器故障时恢复数据，为了您的数据安全，建议开启备份。

网络

虚拟私有云 ①

vpc-default-amb(192.168.0.0/16) [新建虚拟私有云](#)

主网卡

subnet-default-amb(192.168.0.0/20) [自动分配IP地址](#) 可用私有IP数: 4,096个

新增网卡

您还可以增加2块网卡

购买量 **1** 由 **配置费用 ¥1.066 小时**

实际用量以账单为准, [了解详情](#)

立即购买

配置概要

生成Open API

基础配置

计费模式: 按需计费
区域/可用区: 华南-广州 | 随机分配

实例
规格: 鲲鹏通用计算增强型 | kc2.xlarge.2 | 4vCPU | 8GB

操作系统
镜像: 3_1_EulerOS2.0... 任务调度平台(v1)

存储与备份

系统盘: 通用型SSD, 200GB

网络

虚拟私有云: vpc-default-amb(192.168.0.0/16)
主网卡: subnet-default-amb(192.168.0.0/20)
源IP地址: 开启

安全组

Sys-WebServer

公网访问

弹性公网IP: 全流量BGP | 按带宽计费 | 5 Mbit/s | 随买随放

云服务器管理

云服务器名称: ecs-ab06
登录凭证: [密钥](#) | [设置密码](#)
标签: --

我已阅读并同意《弹性云服务器服务协议》 [\(点击查看声明\)](#)

我已阅读并同意《任务调度平台服务协议》 [\(点击查看支持条款\)](#)

[华为云市场通用商品用户协议](#)

对外提供网站服务，需完成ICP备案, [了解详情](#)

安全组

选择安全组 ①

Sys-WebServer(1534561-494b-4d8d-ab88-eb125892f6) [新建安全组](#)

请确保所选安全组已开放22端口 (Linux SSH登录)、3389端口 (Windows远程登录) 和 ICMP 协议 (Ping)。[配置安全组规则](#)

建议将安全组入方向规则中高危端口的源地址设置为已知IP地址，安全风险防护地址组，避免因网络入侵出现业务中断、数据泄露或数据勒索等严重后果。 [查看高危端口](#)

离开安全组规则

公网访问

弹性公网IP ①

现在购买 使用已有 暂不购买

续费 **自动续费BGP** [\(续费BGP已过期\)](#)

不低于99.95%可用性保障

公网带宽 ①

按带宽计费 [\(高带宽大流量场景的优选\)](#) **按流量计费** [\(低成本流量突发场景的优选\) **加入共享带宽** \[\\(多业务流量聚合分布场景\\)\]\(#\)](#)

指定带宽上限，按使用时间计费，与使用的流量无关。

带宽大小 (Mbit/s)

带宽范围: 1-2,000 Mbit/s 开启DDoS基础防护 **免费**

释放行为

随实例释放

对于设置了随实例释放的弹性公网IP，将在删除云服务器时同时执行删除。

购买量 **1** 由 **配置费用 ¥1.066 小时**

实际用量以账单为准, [了解详情](#)

立即购买

配置概要

生成Open API

基础配置

计费模式: 按需计费
区域/可用区: 华南-广州 | 随机分配

实例
规格: 鲲鹏通用计算增强型 | kc2.xlarge.2 | 4vCPU | 8GB

操作系统
镜像: 3_1_EulerOS2.0... 任务调度平台(v1)

存储与备份

系统盘: 通用型SSD, 200GB

网络

虚拟私有云: vpc-default-amb(192.168.0.0/16)
主网卡: subnet-default-amb(192.168.0.0/20)
源IP地址: 开启

安全组

Sys-WebServer

公网访问

弹性公网IP: 全流量BGP | 按带宽计费 | 5 Mbit/s | 随买随放

云服务器管理

云服务器名称: ecs-ab06
登录凭证: [密钥](#) | [设置密码](#)
标签: --

我已阅读并同意《弹性云服务器服务协议》 [\(点击查看声明\)](#)

我已阅读并同意《任务调度平台服务协议》 [\(点击查看支持条款\)](#)

[华为云市场通用商品用户协议](#)

对外提供网站服务，需完成ICP备案, [了解详情](#)

云服务器管理

云服务器名称 允许更名

购买多台云服务器时，支持批量勾选或取消勾选后再勾选。

登录凭证 ①

密钥 **密码**

请牢记密码，忘记密码可登录ECS控制台找回密码。

用户名 密码 确认密码

标签 ①

如果您需要给同一标签的多种云服务器，则所有服务器可在标签输入框下勾选同一标签。建议在IAM中 [创建标签策略](#)

[添加新标签](#)

您还可以添加19个标签。

高级配置

云服务器组: -- 云服务器组描述: -- 固件: --

购买

使用时长

设置定时释放时间

购买量 **1** 由 **配置费用 ¥1.066 小时**

您最多可以创建196台云服务器，申请更多云服务器请单击 [申请扩容](#)

购买量 **1** 由 **配置费用 ¥1.066 小时**

实际用量以账单为准, [了解详情](#)

立即购买

配置概要

生成Open API

基础配置

计费模式: 按需计费
区域/可用区: 华南-广州 | 随机分配

实例
规格: 鲲鹏通用计算增强型 | kc2.xlarge.2 | 4vCPU | 8GB

操作系统
镜像: 3_1_EulerOS2.0... 任务调度平台(v1)

存储与备份

系统盘: 通用型SSD, 200GB

网络

虚拟私有云: vpc-default-amb(192.168.0.0/16)
主网卡: subnet-default-amb(192.168.0.0/20)
源IP地址: 开启

安全组

Sys-WebServer

公网访问

弹性公网IP: 全流量BGP | 按带宽计费 | 5 Mbit/s | 随买随放

云服务器管理

云服务器名称: ecs-ab06
登录凭证: [密钥](#) | [设置密码](#)
标签: --

我已阅读并同意《弹性云服务器服务协议》 [\(点击查看声明\)](#)

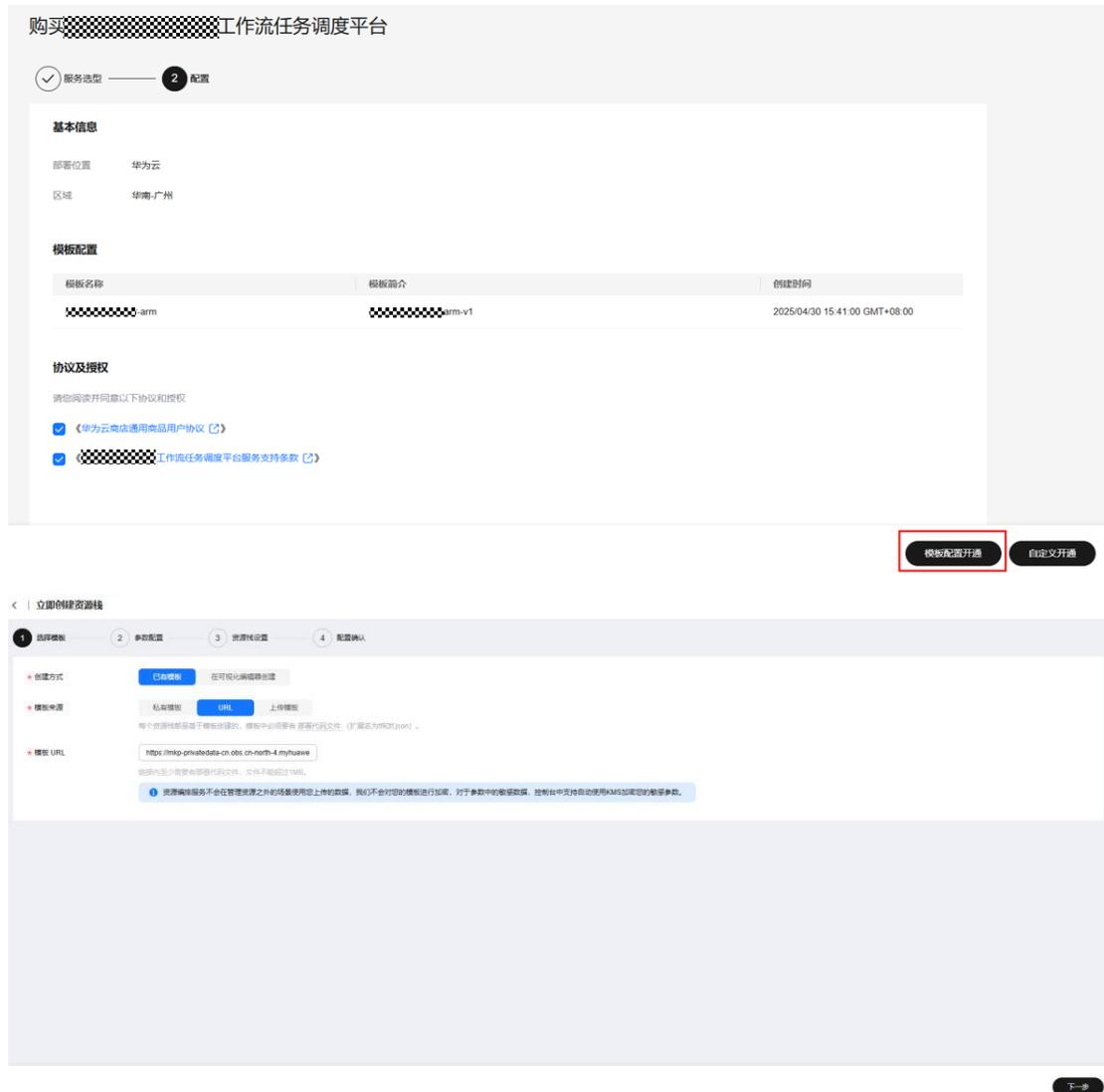
我已阅读并同意《任务调度平台服务协议》 [\(点击查看支持条款\)](#)

[华为云市场通用商品用户协议](#)

对外提供网站服务，需完成ICP备案, [了解详情](#)

【模板配置开通】方式:

使用RFS模板直接部署



必填项填写后，点击 下一步

选择模板 2 参数配置 3 资源栈设置 4 配置确认

配置参数

请输入关键字来筛选参数名称

参数名称	值	类型	描述
ECS实例密码	<input type="password"/>	字符串	ECS实例的登录凭据密码。密码策略要求：密码要求长度为8到24位。密码至少必须包含大写字母、小写字母、数字和特殊字符 (@!#\$%^&*~+ ()_!~)
系统盘大小	40	number	设置系统盘大小 (至少40G, 默认40G)
数据盘大小	50	number	设置注册的数据盘大小, 若不需要数据盘, 可设置值为0, 可按照实际情况配置, 默认为50。
版本	请选择	字符串	选择版本
vpc IPv4网段	192.168.0.0/16	字符串	取值范围: 10.0.0.0/8 to 10.255.255.0/24, 172.16.0.0/12 to 172.31.255.0/24, or 192.168.0.0/16 to 192.168.255.0/24
子网 IPv4网段	192.168.10.0/24	字符串	必须是CIDR格式, 且在VPC的CIDR范围内, 子网掩码不能大于24。
子网的网关	192.168.10.1	字符串	子网的网关, 必须是子网段内的合法IP地址
计费类型(不包含应用防火墙)	请选择	字符串	prePaid-预付费, 即包年包月; postPaid-后付费, 即按量付费
订购周期类型(不包含应用防火墙)	month	字符串	当chargingMode为prePaid时生效且必须填写, 取值范围: month-月, year-年
订购周期数(不包含应用防火墙)	1	字符串	当chargingMode为prePaid时生效且必须填写, 取值范围: periodType=month (周期类型为月) 时, 取值为1, 0; periodType=year (周期类型为年) 时...

上一步 下一步

选择模板 3 参数配置 4 资源栈设置 配置确认

IAM权限委托

huaweicloud 请选择委托

选择权限委托可明确IAM用户在资源栈上可执行的所有操作 (如创建、更新、删除资源栈等)。如果委托权限不足, 可能导致资源部署失败, 生成执行计划操作失败。前往IAM创建委托

用户不配置权限委托系统采用当前登录用户的权限进行部署, 权限委托有授予用户最小化权限对云资源部署的操作, 可有效避免授权或者参数错误导致部署失败或安全事故, 推荐用户配置。

回滚设置 开启回滚表示操作失败时, 会自动回滚至系统所保留的上一个成功的资源栈状态, 创建后可在资源栈详情页修改。

删除保护 防止资源栈被意外删除, 创建后可在资源栈详情页修改。

上一步 下一步

创建直接计划后, 点击 确定

< | 立即创建资源栈

选择模板 3 参数配置 4 资源栈设置 配置确认

配置参数

参数名称	值	类型	描述
ECS实例密码	*****	字符串	密码长度为8到24位。密码至少必须包含大写字母、小写字母、数字和特殊字符 (@!#\$%^&*~+ ()_!~) 中的三种!
系统盘大小	40	number	设置系统盘大小 (至少40G, 默认40G)
数据盘大小	50	number	设置注册的数据盘大小, 若不需要数据盘, 可设置值为0, 可按照实际情况配置, 默认为50。
版本	v1.13.0	字符串	选择版本
vpc IPv4网段	192.168.0.0/16	字符串	取值范围: 10.0.0.0/8 to 10.255.255.0/24, 172.16.0.0/12 to 172.31.255.0/24, or 192.168.0.0/16 to 192.168.255.0/24
子网 IPv4网段	192.168.10.0/24	字符串	必须是CIDR格式, 且在VPC的CIDR范围内, 子网掩码不能大于24。
子网的网关	192.168.10.1	字符串	子网的网关, 必须是子网段内的合法IP地址
计费类型(不包含应用防火墙)	postPaid	字符串	prePaid-预付费, 即包年包月; postPaid-后付费, 即按量付费
订购周期类型(不包含应用防火墙)	month	字符串	当chargingMode为prePaid时生效且必须填写, 取值范围: month-月, year-年
订购周期数(不包含应用防火墙)	1	字符串	当chargingMode为prePaid时生效且必须填写, 取值范围: periodType=month (周期类型为月) 时, 取值为1, 0; periodType=year (周期类型为年) 时, 取值为1, 3

资源栈设置

IAM权限委托 -- 回滚 未开启 删除保护 未开启

操作指南: 创建执行计划 (免费) 后可获取预部署

上一步 创建执行计划 直接部署资源栈

创建执行计划

通过执行计划, 可以预定的资源变更信息。

执行计划名称: executorPlan_20250324_1057_4eda

描述: 请输入对执行计划的描述

确定 取消

点击 部署

基本信息 资源 输出 事件 模板 执行计划

执行计划名称ID	状态	操作	创建时间	描述	操作
executionPlan_20250324_1957_4da	创建成功, 待部署	查看部署详情	2025/03/24 10:58:08 GMT+08:00	-	重新部署

如下图“Apply required resource success.”即为资源创建完成

基本信息 资源 输出 事件 模板 执行计划

事件时间	事件类型	事件描述	资源名称/类型	关联资源ID
2025/03/24 11:00:06 GMT+08:00	日志	Apply required resource success.	-	-
2025/03/24 11:00:01 GMT+08:00	生成成功	module.ecs.huaweicloud_compute_instance.ecs[0]: Creation complete after 57s [id=aa08d928-dc2e-4095-b388-4b333ec8b44]	ecs ECS	aa08d928-dc2e-4095-b388-4b333ec8b44
2025/03/24 11:00:01 GMT+08:00	完成概要	Apply complete! Resources: 0 added, 0 changed, 0 destroyed.	-	-
2025/03/24 10:59:54 GMT+08:00	正在生成	module.ecs.huaweicloud_compute_instance.ecs[0]: Still creating... [50s elapsed]	ecs ECS	-
2025/03/24 10:59:44 GMT+08:00	正在生成	module.ecs.huaweicloud_compute_instance.ecs[0]: Still creating... [40s elapsed]	ecs ECS	-
2025/03/24 10:59:34 GMT+08:00	正在生成	module.ecs.huaweicloud_compute_instance.ecs[0]: Still creating... [30s elapsed]	ecs ECS	-
2025/03/24 10:59:24 GMT+08:00	正在生成	module.ecs.huaweicloud_compute_instance.ecs[0]: Still creating... [20s elapsed]	ecs ECS	-
2025/03/24 10:59:14 GMT+08:00	正在生成	module.ecs.huaweicloud_compute_instance.ecs[0]: Still creating... [10s elapsed]	ecs ECS	-
2025/03/24 10:59:04 GMT+08:00	生成成功	module.vpc.huaweicloud_vpc_subnet.subnet: Creation complete after 9s [id=4b3ec9ef-a475-4a7d-9e7c-00000316c3763]	subnet Subnet	4b3ec9ef-a475-4a7d-9e7c-00000316c3763
2025/03/24 10:59:04 GMT+08:00	正在生成	module.ecs.huaweicloud_compute_instance.ecs[0]: Creating...	ecs ECS	-

3 商品使用

3.1 准备工作

在使用ECS控制台配置前，需要您提前配置好安全组规则。

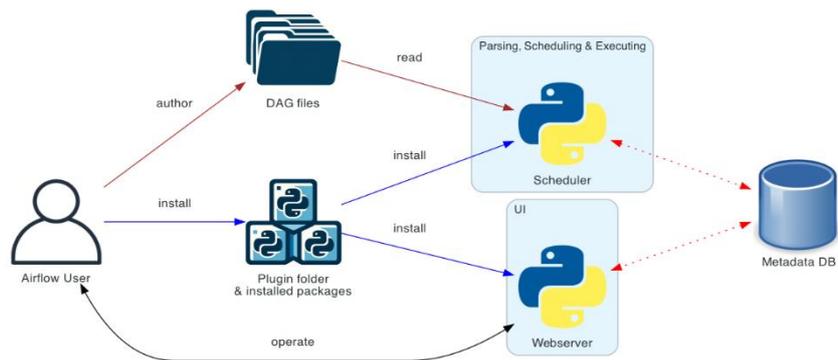
安全组规则的配置如下：

- 入方向规则放通端口见下面列表，源地址内必须包含您的客户端公网ip，否则无法访问。
- 入方向规则放通CloudShell连接实例使用的端口22，以便在控制台登录调试。
- 出方向规则建议一键放通。

端口说明：

组件	默认端口	说明
Airflow	8080	Airflow 的 web ui 默认端口, 用于创建和管理 DAGs
后端数据库	依据数据库厂商	如: GaussDB 8000
SSH Client	22	SSH 远程访问主机端口

3.2 Airflow 的使用



架构图示例:

1、修改/etc/hosts 中自己的 ip

192.168.0.24 hadoop1 --> *.*.*.* hadoop1

2、命令行进入 conda 虚拟环境

conda activate python39

3、启动 web 程序

```
airflow webserver --port 8080 -D
```

4、启动 scheduler 调度程序

```
airflow scheduler -D
```

5、浏览 web ui

```
ip:8080
```

6、用户及密码

```
airflow/123456
```

3.3 参考文档

- [Airflow官网](#)