

## 目录

目录	1
运维中心概述	2
名词解释	2
任务运维	2
组件实例	3
生产实例	3
测试实例	4
事件补发	4
作业临时重跑	4
诊断分析	4
运营配置	4
常见问题	5

# 运维中心概述

运维中心为大数据云平台中其它有作业调度执行需求的组件(数据集成、流计算、数据管理等)及第三方应用提供统一的作业调度、执行平台。运维中心支持多租户、支持细粒度的执行资源控制、支持多种依赖配置(流依赖、作业依赖、事件依赖,同时还可以在依赖上进行偏移)。通过基于事件消息驱动的方式,保证作业执行更加高效。

## 名词解释

- **作业流:** 指一个由作业节点组成的图。每个作业节点按照配置完成一定的处理逻辑。作业节点之间要么无关联,要么通过有向边进行依赖关联,但关联时不能形成环路。一个画布中的全部作业节点及其依赖(如果有的话)称为一个作业流。一般来说,在作业流调度模型中,作业流为调度单元,而其中的作业节点为最小粒度的执行单元。
- **作业:** 作业流中的一个节点,即由用户定义的完成一定工作的逻辑单元。在任务调度模型中,作业(或任务)是最小执行单元。
- **实例化:** 生成作业实例的动作称为实例化。实例化主要是设置作业执行时需要的业务时间、批次。
- **翻牌:** 一个作业流执行结束后,根据策略生成该作业流下后续一到多个批次作业流实例及作业流/作业依赖的动作称为翻牌。
- **事件:** 可能触发系统对作业流/作业进行实例化、状态改变、派发/执行等操作的消息。
- **批次号:** 作业流/作业实例的编号,对同一个作业流/作业,此编号唯一。默认情况下批次号从1开始递增,手动触发时可能要求输入批次号,例如测试时要求输入批次号。
- **触发:** 每个作业/作业流运行都有一个前提条件。这个前提条件可以是一个外部事件,也可以是一个流内作业或者跨流作业。当满足作业/作业流的前提条件时,作业就会运行,这个满足前提条件后运行的过程叫做触发。
- **外部依赖:** 当一个作业/作业流的前提条件为一个外部事件时,这个外部条件可以是一个外部文件。此时成这个作业依赖一个外部依赖。
- **流内依赖:** 同一个作业流内作业间的相互依赖称为流内依赖。流内依赖在作业流内体现为两个作业间的连线。
- **跨流依赖:** 一个作业/作业流依赖其他作业流内的某个作业。称为跨流依赖某个作业。一个作业/作业流依赖其他作业流,称为跨流依赖某个作业流。
- **批次偏移:** 当存在跨流依赖或者外部依赖时,被依赖的作业/作业流或者外部事件可能存在多批次,依赖多批次中某个特定批次时,称为此时的依赖存在批次偏移。
- **时间偏移:** 当存在跨流依赖或者外部依赖时,被依赖的作业/作业流或者外部事件不是当前作业流实例运行的业务日期而历史的某个外部事件或者历史的作业/作业流实例时,称为此时的依赖存在时间偏移。

## 任务运维

在任务运维页面可以查看离线计算作业流/作业、流计算作业信息、数据采集任务等,并可以进行作业流/作业流的测试、立即执行、预定义暂停、发布、删除、查看实例等操作。

### 1. 作业流管理

在作业流管理页面可以查看作业流信息,并可以进行作业流的测试、立即执行、预定已暂停、发布、删除操作。如果作业流太多,可以通过条件搜索你想管理的作业流。

#### (1) 上线启动

作业发布后此时作业状态为【发布】态,需要进行【上线启动】,作业流才能真正的发布到线上,进行周期性运行,上线启动的过程中主要是进行资源组校验。

#### (2) 立即执行

对状态为【上线】的作业流可以进行【立即执行】操作,输入需要运行的业务日期和批次号,作业流按照业务日期和批次号立即执行。

#### (3) 修数据

点击作业流上的【修数据】,选择一个回退日期,作业流从回退日期运行到当前时间,进行历史数据修数调整。

#### (4) 预定义暂停

【预定义暂停】是让作业流在指定的业务日期内停止调度。只有当作业流处于“已发布”状态才可以使用“项目管理”和“运

维人员”角色子账户进行“预定义暂停”操作。点击**预定义暂停**，在弹出对话框中打开**【预定义暂停】**按钮，输入需要暂停的业务日期时间段，点击**确认**。

#### (5) 作业设置

在作业流上可以进行调度作业的启用/禁用操作。

### 2. 作业管理

在作业管理页面可以查看作业信息，并可以查看作业实例。如果作业太多，可以通过条件搜索你想管理的作业。

#### (1) 优先级调整

每个作业任务都有一个优先级，数字越高优先级越低，当特定任务需要优先运行时可以对作业进行优先级调整。

#### (2) 资源调整

可以对每个作业进行资源调整，此处的调整包括，执行资源的选择和逻辑资源组修改。

#### (3) 查看实例

点击作业上查看实例跳转到运行实例列表。

## 组件实例

运维中心支持离线计算、流计算、数据采集、图计算、数据挖掘、数据服务、数据管理等多个组件的实例运行管理。以离线计算为例说明实例管理功能。

### 生产实例

(1) 离线计算-作业流实例 用户可以在此页面查看作业流实例运行情况，并对作业流实例进行相关操作。除**【查看DAG】**操作外，其他操作均需先通过复选框选中待操作作业流对象后才能进行操作。

- 作业流删除

从作业流实例及此作业流下所有作业实例数据库中删除，删除后页面将不可见，也不可恢复。仅处于**【完成】**、**【失败】**状态的作业流才可以。

- 启用翻牌

启用翻牌，作业流执行完成以后自动根据条件实例化下一业务日期的作业流。

- 暂停翻牌

停止翻牌，作业流执行完成后不再实例化下一业务日期的作业流，但当前已经实例化的作业流会继续执行完。仅处于**【运行中】**、**【未就绪】**状态的作业流才可以**【暂停翻牌】**。

- 终止

终止作业流实例。正在执行的作业被强制结束，未执行的作业被取消。仅处于**【运行中】**的作业流才可以被终止。

- 查看DAG

点击**查看DAG**，跳转到作业流实例的运行DAG页面。

页面以流图的形式展示了作业流中各个作业关联关系，并用颜色标注了作业状态。

- 依赖项检查

点击作业流实例**依赖项检查**，跳转到作业流实例的依赖项页面。

(2) 离线计算-作业实例 用户可以在此页面查看作业实例运行情况，并对作业实例进行相关操作。除**【查看日志】**操作外，其他操作均需先通过复选框选中待操作作业流对象后才能进行操作。

- 作业重跑

作业处于**【终止】**、**【成功】**或**【失败】**时，可以选中作业，点击此按钮让作业实例再次运行。

- 作业挂起

作业处于**【未就绪】**、**【就绪】**、**【等待】**、**【已派发】**时，可以选中作业，点击此按钮让作业实例在条件全部具备的条件下

也不被执行。页面中作业实例挂起状态变为【挂起】。

- 作业解挂

解除作业【挂起】状态，页面中作业实例挂起状态变为默认的【非挂起】状态。

- 作业强制成功

把作业状态强制置为【成功】，此过程中作业不会被执行。非【运行中】状态的都可以执行【作业强制成功】操作，此功能主要用于触发依赖于此作业的下游作业执行。

- 作业优先级调整

作业派发过程中会根据作业优先级进行派发，调整作业优先级可以使作业被优先派发执行。作业优先级一共有5级，数字越大优先级越高，其中第5级设置只对项目管理员开放，以方便临时加一些需紧急执行的作业。作业处于【未就绪】、【就绪】、【等待】或【取消】状态时可以调整作业优先级。

- 作业终止

终止作业可让正在运行的作业强制结束，让还未执行的作业不再执行。

- 日志

点击作业列表右侧日志可以查看作业运行日志。 跳转到作业实例日志页面。

- 依赖项检查

点击作业列表右侧**依赖项检查**，查看作业实例运行的依赖项。作业实例的依赖项包括【作业流】、【作业依赖】、【外部依赖】三种依赖。

## 测试实例

“测试实例”与“生产实例”的功能、页面、操作完全一样，区别仅在于“测试实例”里的任务在测试环境运行，且“开发人员”角色可以进行所有操作。

## 事件补发

当作业/作业流因为某些原因需要进行外部依赖的手动补发时，可以在【事件补发】页面进行统一的补发操作。补发分为输入和输出两个方向的补发操作。

- 当用户选择【输入】方向补发时，可补发的对象包括作业流和作业。
- 当选择补发方向为【输出】时，可补发的对象包括外部事件、作业流、作业。

## 作业临时重跑

点击**作业临时重跑**，选择项目空间，在目录中找到要操作的作业拖拽至右侧，勾选要重跑的作业，选择执行次数和业务日期，填写批次号，点击运行。

## 诊断分析

诊断分析功能主要确定作业/作业流中在运行是有哪些异常以及使用血缘分析来确定异常的事件影响范围有多大。

### (1) 异常诊断

进程出现异常时用户可通过异常诊断页面对问题进行排查 点击**诊断分析** > **异常诊断**。

选择项目和作业流，点击**全面诊断**，列出目前针对作业流进行的诊断项。

当诊断某项诊断有异常时，会在诊断结果中标显示，可以点击**有异常**，进行详细异常原因查看。

### (2) 血缘分析

点击**血缘分析**，选择作业维度或作业流维度，以作业为例，依次选择项目 > 作业流 > 作业名称，输入父、子节点数查看上、下依赖关系。双击屏幕上的某个节点可以转换到以该节点为中心进行查看。

## 运营配置

(1) 逻辑资源组 进入【调度配置】，选择**逻辑资源组配置**查看列表，显示逻辑资源组队列以及当前逻辑资源组逻辑资源总量、使用量和剩余量。

点击**添加逻辑资源组**，创建新的逻辑资源组。设置资源组名称、总量和是否生效等。

对于列表中的资源组可以点击进行**修改**。

## (2) 周期日历设置

进入**【周期日历设置】**查看列表，显示目前系统支持的所有的调度周期。类型为**【公共设置】**的是系统默认设置周期日历，用户自己添加的类型为**【自定义】**。

点击**周期日历设置**按钮浏览上传本地文件进行添加，新建一个调度周期日历。

已经添加的日历可以通过对应按钮进行**【修改】**、**【查看】**、**【下载】**操作。

- 点击**修改**跳出修改弹窗。用户重新上传日历文件，进行周期日历的修改。
- 点击**查看**按钮，显示当前周期日志中的日期设置。
- 点击**下载**按钮，将当前周期日历已表格形式下载到本地。

## 常见问题

### 1. 调度运维有哪些优势？

调度运维通过可视化的任务编排，对不同类型的任务进行统一的调度，监控运维。实现一个多任务，高并发，体验好的调度平台，支持客户在不同场景下，不同需求的调度需求，在数据平台建设中起到基础和保障的作用。

### 2. 什么是调度的批次？

调度任务在一个自然日内运行的次数称为调度的批次。

### 3. 作业流停止翻牌时为什么会失败的情况？

仅处于**【运行中】**、**【未就绪】**状态的作业流才可以进行**暂停翻牌**操作，如果作业流里作业少且执行快，作业流处于**【运行中】**、**【未就绪】**时间很短，页面显示的状态与实际不符。这种情况可以在作业流实例化后快速操作避免操作失败。