南京金宁汇科技有限公司		
文档编号:	JNHT-JJGLPT -DES-UserManual	

一站式大数据开发与治 理平台 用户手册

V0.0.1

编写:廖宁	审核: 李延辉	批准:
日期: 2021年11月11日	日期: 2021 年 11 月 18 日	日期:年月日

变更履历

版本	文件内容描述	编写日期	编写	审核	批准
V0. 0. 1	工作空间、数据源、数 据同步、同步任务调度 操作流程	2021年11月11日	廖宁	李延辉	

目录

1	概述		1
2	系统介	↑绍	1
	2. 1	产品/项目中的用户角色	1
3	基本功	b能	2
	3. 1	登录	2
	3. 2	首页	3
	3. 3	工作空间	4
		3.3.1 空间设置	4
		3.3.1.1 工作空间新增	4
		3.3.1.2 工作空间列表操作	8
		3. 3. 2 数据源	10
		3. 3. 2. 1 数据源新增	.10
		3.3.2.2 数据源列表操作	.13
	3. 4	数据管理	15
		3. 4. 1 数据同步	15
		3.4.1.1 同步任务新增	.15
		3.4.1.2 数据同步任务操作	21
		3.4.1.3 同步任务列表操作	22
		3.4.2 调度首页	25
		3.4.3 同步任务调度	.26
		3.4.3.1 同步任务调度列表操作	27

1 概述

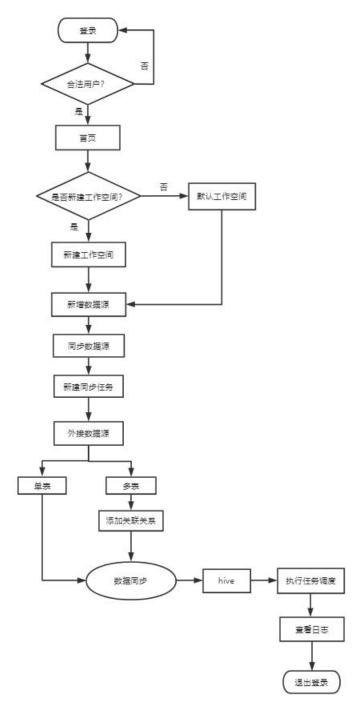
该系统为一站式大数据开发与治理平台(以下简称 iCredit),致力于为数据开发者、数据分析师、数据资产管理者这打造一个具备全域数据开发和治理能力的一站式大数据平台,该版本主要包含工作空间、数据集成两个核心模块,实现数据源的导入和数据同步任务的调度。

2 系统介绍

2.1 产品/项目中的用户角色

角色	人员	职责描述	备注
管理员		创建工作空间;分配权限;系统配置 及管理	
普通用户	数据开发人员	配置数据源以及数据同步任务;管理调度任务	

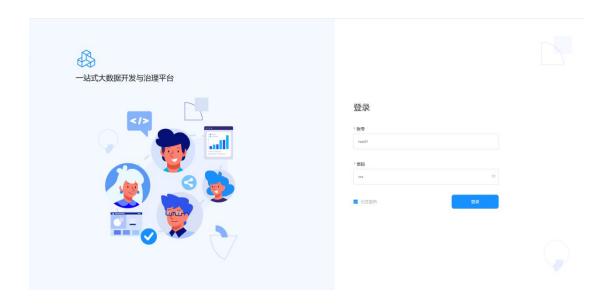
3 基本功能



业务流程图

3.1 登录

如果您是 iCredit 内管理员及其所创建的用户, 在登录页面, 输入如下信息:



用户名和密码填写您的账号密码即可,点击【登录】按钮验证成功即可进入 iCredit 首页。

默认超级管理员的账号密码为:

用户名: root 密码: root

3.2 首页

前提条件

- 用户合法登录 iCredit 平台。
- 用户拥有操作首页的操作权限。

操作流程

在首页页面用户可在工作台看到任务各种状态数量的概览,页面右侧为该用户所创建的任务中的执行情况,工作台包含未运行任务数量,正在执行任务数量、运行失败和运行成功的任务数量。菜单包含个人中心、个性化设置、修改密码和退出登录,提供快捷入口给用户进行个性化设置操作。实践教程包含用户手册,即iCredit平台的使用方法,单击后打开新窗口页面展示PDF文档,支持下载文档。左侧为iCredit平台的开发和治理的流向图,单击模块标题可跳转到相应的模块首页页面,底部的功能演示简单直观的展示了该平台所包含的7个功能模块,支

持用户单击每个方块跳转到相应功能模块的首页。



3.3 工作空间

工作空间是 iCredit 的物理分区,每个工作空间的数据源、服务器等资源相互独立,数据源和数据开发都是基于某个工作空间下的操作,在相应的模块页面,都可以通过手动切换工作空间来操作不同的数据。

工作空间下要完成空间设置和数据源的接入。

3.3.1 空间设置

在工作空间下可配置对应的负责人、成员和角色。下列操作主要以超级管理员用户对工作空间的添加为例。

3.3.1.1 工作空间新增

前提条件

● iCredit 的合法用户登录进入到首页。

● 用户拥有新增工作空间的操作权限。

操作流程

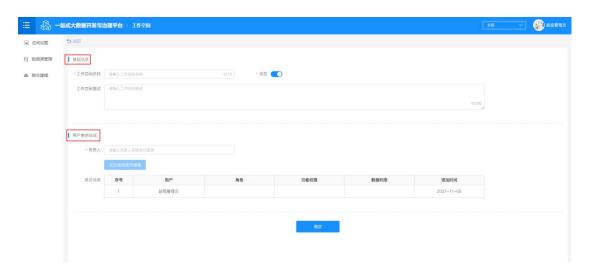
- 1. 登录 iCredit 系统, 进入到 iCredit 首页。
- 2. 单击左上角导航菜单显示图标: ,显示左侧导航栏。



3. 在左侧导航栏中,单击工作空间,进入空间设置列表页面,默认列表只有一条默认工作空间的数据。



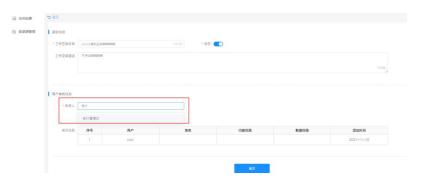
- 4. 在空间设置页面,默认工作空间下,单击【新增工作空间】按钮,进入到新增工作空间页面。
- 5. 在新增工作空间页面中,配置参数。



新增工作空间需要配置的参数如下表所示:

参数	描述
工作空间名称	必填,填写工作空间的名称,名称不允许重复,50字以内不包含特殊字符和空格。
状态	单项选择,默认启用,必选。
工作空间描述	非必填,填写对工作空间的简单描述。
负责人	必选,该工作空间的负责人。
点击添加成员信息	非必选,穿梭框选择成员信息,赋予成员该空间的查看权限。
	约束限制:添加的成员信息只拥有对工作空间的查看权限,其他新增、编辑和删除等权限均是不支持的。
成员信息列表	显示当前用户、负责人和选择的成员信息。

用户需要填写的是空间名称、描述等信息,负责人需要输入框输入才能显示负责 人信息可用户供选择。



如果用户还想将其他的用户都拉到该工作空间中并赋予他们操作该工作空间的权限,则单击【点击添加成员信息】按钮,则显示穿梭框选择成员信息。



选择你想要的成员,单击【保存】按钮,则在成员列表中显示能够操作该空间的 所有成员信息。具体情况如下所示。



- 6. 填写完工作空间的配置参数后,单击【确定】,完成新增工作空间。
- 7. 新增成功的工作空间在右上角下拉框中显示,若要切换新增的工作空间,则在右上角下拉框单击即可切换。由此完成了工作空间的新增。



8. 新增工作空间成功后即可在工作空间下添加 3.3.2 数据源。

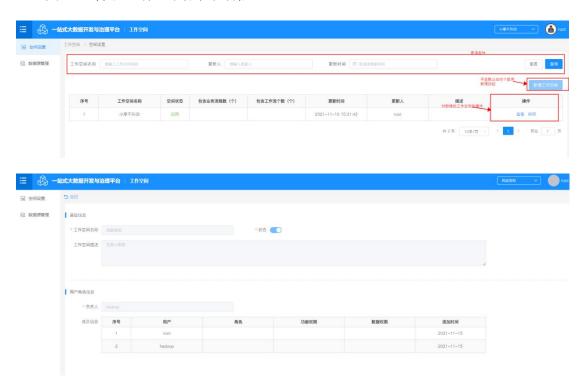
3.3.1.2 工作空间列表操作

前提条件

- 新增工作空间成功。
- 用户拥有操作空间设置的操作权限。

操作流程

1) 在空间设置页面的查条件中输入你所感兴趣的工作空间的信息,单击【查询】 按钮查询到你所想要的工作空间数据,单击【查看】进入到工作空间详情页 面,查看该工作空间的详细信息。



详情页面包含工作空间的新增页面的信息回显。

2) 用户可以对处于"停用"状态的工作空间进行编辑,单击【编辑】按钮后可以对下图信息进行修改。



3) 如果用户想要删除工作空间,则单击【删除】按钮,确认提示内容后单击【确 定】则将该工作空间及其包含的内容都删除。



4) 如果该工作空间内的工作流和调度都无需使用到该数据源,则单击【停用】 按钮将该工作空间内的数据源和任务都停止调度。



5) 如果用户想要将该工作空间重新启用,则单击【启用】按钮,提示:启用成功!则将该工作空间启用,工作空间中的项目可以重新进行调度。

3.3.2 数据源

数据源管理是指对于需要接入平台的数据源的管理,目前支持的类型有关系型数据库(MySQL)。

3.3.2.1 数据源新增

前提条件

- 用户新增工作空间成功。
- 用户拥有新增数据源的操作权限。

操作流程

1. 工作空间页面左侧导航栏中单击数据源管理。

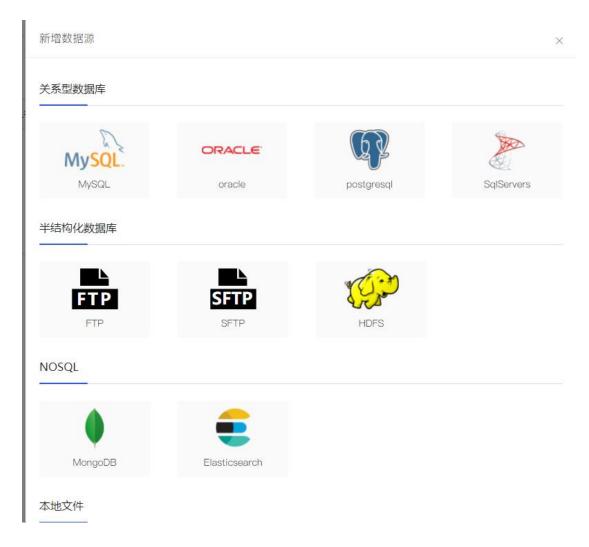


2. 进入数据源管理列表页面。

数据源管理页面显示已添加的数据源信息列表,包含连接信息,启用状态,最近 同步状态和同步时间等信息,同时支持添加数据源、删除数据源、编辑数据源配 置,支持对新增的数据源停用、启用、同步和查看操作。



- 3. 在资源管理页面,单击【新增数据源】按钮,进入新增数据源页面。
- 1) 选择数据源类型:数据源类型目前支持关系型数据库(MySQL),示例图如下所示(这里添加的是 MySQL):



2) 点击 MySQL 类型数据库,进入新增数据源配置参数页面。



用户需要填写的参数如下:

参数	描述
数据源名称	必填,填写数据源的名称,名称不允许重复,50字以内不包含特殊字符和空格。
数据库名	必填,填写数据源所在的数据库名,要与数据库中相匹配。
IP	必填,数据库 IP 地址
端口	必填,数据库端口号。
用户名	必填,数据库用户名。
密码	必填,数据库密码,明文。
启用	必选,单项选择是否启用该数据源。
数据源描述	非必填,对该数据源的简单描述。

- 4. 填写完数据源的配置参数后,单击【测试连接】弹出"测试成功!"即可连接数据源成功,如果测试不成功,请检查一下配置项以及数据库是否正常启动或网络是否正常。
- 5. 单击【确定】则新增数据源成功,返回数据源列表显示,可在列表中查看到新增的数据源及其连接信息。
- 6. 数据源新增成功后,单击操作列表下的【同步】按钮进行数据源同步,将数据源的表结构采集到 iCredit 平台中,平台同时会每隔 10 分钟进行一次表结构

同步。



- 7. 提示同步成功即说明数据源成功同步到 iCredit 平台。
- 8. 数据源同步成功后可到 3.4.1 数据同步中新建数据同步任务。

3.3.2.2 数据源列表操作

前提条件

- 用户新增数据源成功。
- 用户拥有操作数据源的操作权限。

操作流程

1) 通过查询条件输入你想要查询到的数据源信息,单击【查询】在列表中展示你感兴趣的数据源信息,可以对处于"启用"状态的数据源单击【查看】进入数据源详情页面。





详情页面包含数据源的名称、IP、端口等新增时填入的信息回显,还支持对数据源的各个表的表结构信息的查看。

2) 您也可以对该处于"停用"状态的数据源进行编辑,单击【编辑】按钮,可以对下图信息进行修改。



3) 当用户想要删除数据源时,也可以对数据源进行停用后,单击【删除】按钮, 对数据源进行删除操作。



单击确定即可删除成功。

4) 如果用户在工作流中无需再使用到该数据源,则可以对数据源进行停用,单击【停用】按钮即可。



5) 如果您想要将停用状态的数据源重新启用对它进行使用,则单击【启用】按钮,提示: 启用成功!则成功启用该数据源。

3.4 数据管理

3.4.1 数据同步

作为 iCredit 平台的核心枢纽,将不同系统的数据相互打通,实现数据自由 离线或实时流动,并致力于提供复杂网络环境下、丰富的异构数据源之间高速稳 定的数据移动能力,繁杂业务背景下的数据同步解决方案。

3.4.1.1 同步任务新增

前提条件

● 完成 iCredit 的登录进入到首页。

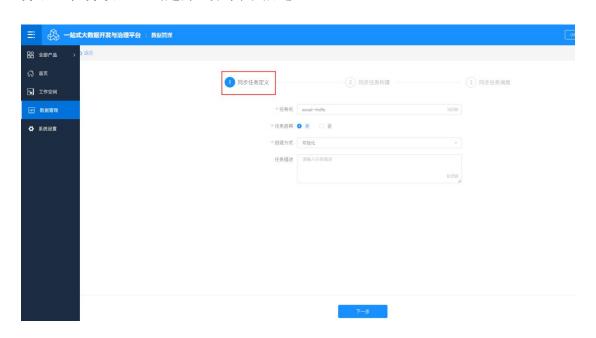
- 完成工作空间的添加和数据源的同步。
- 用户拥有新增数据同步任务的操作权限。

操作流程

- 1. 完成 3.3.2 数据源的同步后,单击左上角图标 □,显示左侧导航栏。
- 2. 单击【数据管理】,进入到数据同步列表首页。



- 3. 单击【新增同步任务】按钮添加一个同步任务,进入到新增同步任务页面,同步任务的创建总共三个步骤,第一步:同步任务定义;第二步:同步任务构建:第三步:同步任务调度。
- 3.4.1.1.1 同步任务定义
- **第一步:同步任务定义。**同步任务定义的内容为该同步任务的基本信息,包含任务名,任务状态,创建方式和简单描述。



该步骤所需要填写的参数如下表所示:

参数	描述
任务名	必填,同步任务的名称,自动生成不具有唯一性。
任务启用	必选,单项选择,默认"是"。
创建方式	必填,下拉框,默认为"可视化"。
任务描述	简单对该同步任务进行描述。

可视化和 SQL 创建方式的区别:

可视化:拖拽方式更快捷,操作更简单,但是最多只能添加四张表进行关联,适用于添加的表较少,建立的表之间关联关系较少或者不熟悉 SQL 语句的情况的使用。

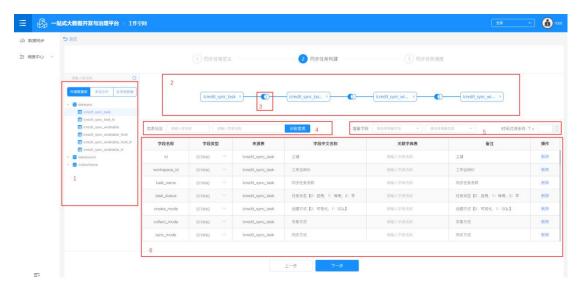
SQL: 需要用户手写 SQL 语句,用户可以根据 SQL 表达式查询表,并且建立表之间的关联关系,没有表数量的限制,可以查询任意的表并且添加他们之间的关联关系,需要对 SQL 表达式的了解程度较高。

单击【下一步】则保存第一步的内容并进入到下一步同步任务的构建中。

3.4.1.1.2 同步任务构建

第二步:同步任务构建。该步骤是添加数据源,对同步到 iCredit 平台的数据源进行识别并简单处理形成宽表同步到 hive 的过程。对于同步任务的创建有两种方式:可视化和 SQL。目前只支持外接数据源的数据同步任务构建。

1) 在左侧树中上部的搜索框可搜索左侧树的库名和表名,支持模糊搜索。根节点为库名,子节点为表名。如果在第一步的创建方式中选择了可视化的创建方式,则在第二步中即可将左侧树中的数据表拖拽到右侧上部的画布区域即可。该种方式最多能够拖拽四张表,拖拽过程表之间自动连接。



点击两张表之间的连接图标,创建两张表之间的关联关系。需要选择关联类型和 关联条件。单击【确定】完成这两张表的关联关系的创建。



当画布区中所有的表的关联关系都被创建完成后,填入宽表信息,即库名和宽表名称(数据源同步到 hive 库,宽表名称自动生成,即数据源同步到 hive 中形成的表的名称,宽表的字段点击识别宽表时显示在列表中),单击【识别宽表】。则在列表中显示识别的数据源宽表字段,可对数据的宽表字段进行删除,修改字段中文名称、关联字典表和字段类型等。

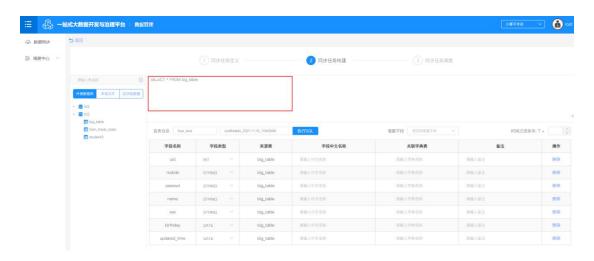


如果用户需要设置增量同步,则选择增量同步字段后,勾选增量存储,则按照增量字段和 cron 表达式进行增量同步,默认在 hive 中每次调度自动新增一个文件夹进行存储数据;之后填写时间过滤条件,T 为每次实际调度的时间,文本框中按照增量类型加 N,如为月增量,N 为 1,每天 0 点调度,则每次 0 点调度时取增量字段的前一个月的数据;如果为日增量,则取得是调度时间前一天的数据。



如果用户没有选择增量字段,则用户新增的该同步任务则为全量同步。

2) 如果在第一步中的创建方式选择了 SQL 方式创建同步任务,那么第二步中的 画布区域为 SQL 表达式的输入框,输入 SQL 语句查询表和创建表之间的关联 关系。



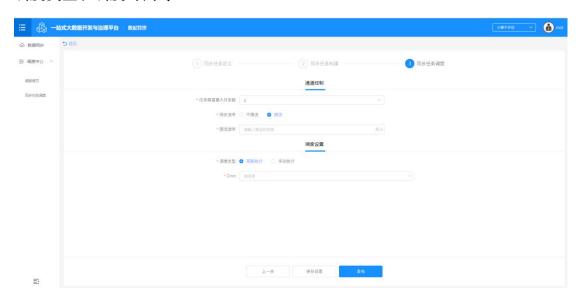
如果执行 SQL 时表达式中所选择的库存在同名情况,则单击【执行 SQL】按钮后提示需要选择正确的库的 IP 地址才能执行。选择正确的库后单击【确定】即可识别数据源生成宽表,在列表显示宽表字段。其他步骤和可视化创建方式一致。



完成第二步的内容构建后单击【下一步】则保存第二步的内容进入到第三步: 同步任务调度。

3.4.1.1.3 同步任务调度

第三步:同步任务调度。该步骤需要配置该同步任务最大的并发数、同步速率、调度类型和调度时间等。



如果用户选择了限流(即设置每秒同步的数据量条数,单位:条/秒),则显示限流速率必填输入框,调度设置时如果选择的是周期执行,则需要选择调度周期,如果是手动执行则不需要选择调度周期,选择周期调度时间后单击【确定】则完成周期调度时间的设置。



该步骤所需要配置的具体参数内容如下表所示:

参数	描述
任务期望最大并发数	必填,下拉框默认为2。
同步速率	必选,默认限流。
	限制:选择限流时需要填写限流速率
调度类型	必选,默认为周期执行。
	限制:选择周期执行须填写调度时间,手动执行则不用
调度时间	单击选择任务调度的时间

填写完第三步内容后单击【发布】,则将该同步任务发布成功,返回同步任务列表页面,任务状态与第一步中选择的状态相同。如果用户单击【保存设置】,则显示的是保存成功,任务状态为草稿状态,停用和草稿状态不能够进行同步任务调度的执行。



在列表的操作列下单击【立即执行】,页面提示操作成功,即可到 <u>3.4.3</u> 同步 任务调度中查看调度任务的执行情况,在 <u>3.4.2</u> 调度首页中能查看到所有的任 务运行结果概览。

3.4.1.2 数据同步任务操作

3.4.1.2.1 下一步

第一二步中的【下一步】按钮只有当页面的配置项都填写并且验证成功后才起作用,否则提示验证结果错误提示,并且配置项不允许为空,当填写配置项并且验证合法后单击【下一步】进入到下一步的内容配置的同时保存该步骤配置内容。

3.4.1.2.2 保存设置

在第三步中点击【保存设置】按钮保存当前数据同步任务的配置数据结果,保存时需要进行配置参数输入的非法检测,当有非法内容存在时不允许保存,不允许有未填的必填项。

3.4.1.2.3 发布

发布时检测:

- 非法:配置项中是否填有非法格式的内容;
- 缺失:用户是否有必填配置项的内容未填写;

当检测到任何一种情况都发布失败,不能发布成功该项任务,显示发布失败提示, 但是允许该项任务被保存。

3.4.1.3 同步任务列表操作

前提条件

- 完成数据同步任务的创建。
- 用户拥有操作数据同步的操作权限。

操作流程

1) 当用户新增有同步任务显示在列表中时,在页面上方查询条件框中输入查询 条件,单击【查询】后可查询到用户所感兴趣的同步任务,在任务列表中单 击【查看】按钮可查看到该同步任务的详细设计内容。

任务详情:



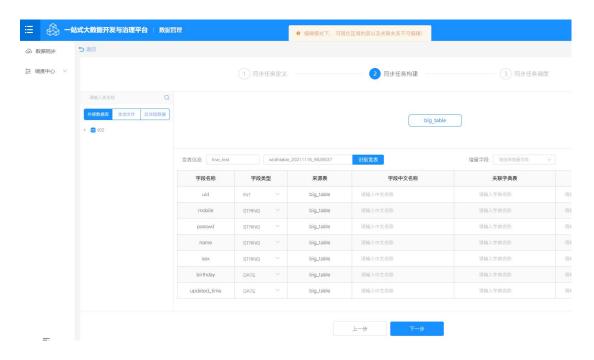
数据源详情:



任务调度详情:



2) 如果用户想要对该同步任务进行修改,则单击列表操作下的【编辑】按钮对 同步任务进行编辑。可编辑的内容如下图所示,页面内容是新增页面填写的 信息的回显,其中第二步同步任务构建中的画布区域部分是不支持编辑的。



3) 如果该同步任务已不再从源表写入数据源进目标表,或不再进行调度任务,用户可以对该数据同步任务进行停用操作,单击【停用】按钮,查看确认页面提示内容后单击【确认】,页面提示:停用成功!,则将该数据同步任务停用成功。



- 4) 如果用户想要重新对停用的数据同步任务启用,则单击【启用】按钮,页面提示: 启用成功!,则将该同步任务重新启用。
- 5) 如果该数据同步任务并没有存在的必要则可单击【删除】按钮,将该同步任 务删除。确认页面提示后单击【确认】,页面提示;删除成功!则该同步任 务删除成功。



3.4.2 调度首页

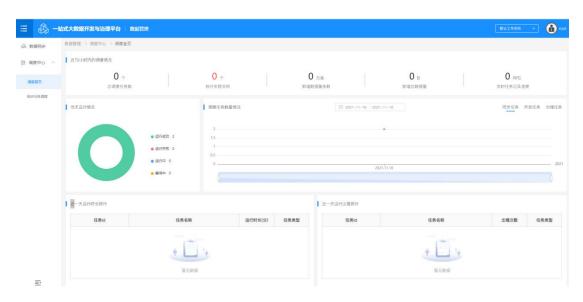
调度首页提供所有调度任务的情况总览。具体的任务调度情况可到 3.4.3 同步任务调度中查看。

前提条件

- 完成 iCredit 的登录进入到首页。
- 用户拥有查看调度首页的操作权限。

操作流程

1. 单击左侧导航栏调度中心,单击【调度首页】,进入调度首页页面如下:



调度首页包含近72小时内的任务调度情况、当天的运行情况、调度任务数量情况、近一天运行时长排行和近一天运行出错排行。

近72小时内的任务调度情况,包含总调度任务数、执行失败实例个数、新增数据量条数、新增总数据量和实时任务记录速度。

当天的运行情况, 饼图的形式展现, 包含运行成功、运行失败、运行中和等待中的任务数量。

调度任务数量情况部分用折线图形式展示,可选择任务调度的时间范围和任务的类型。

近一天的运行时长用列表形式展现可以看到最近一天的任务运行时长的排序,运行时长最长的排在第一位,展示任务的 id,任务名称,运行时长(分)和任务类型。

近一天运行出错排行也是用列表形式展示近一天中执行调度时出错次数最多的任务 id, 任务名称, 出错次数和出错类型。

2. 在该页面可查看到相应时间内的任务执行总体情况。要查看单个任务具体情况可在 3. 4. 3 同步任务调度中查看。

3.4.3 同步任务调度

该模块提供同步调度任务的执行的具体情况,支持历史执行情况和日志的查看。

前提条件

- 用户完成 iCredit 的登录进入到首页。
- 用户创建数据同步任务并执行。
- 用户拥有同步任务调度的操作权限。

操作流程

- 1. 单击左侧导航栏中【调度中心】,单击【同步任务调度】。
- 2. 进入到同步任务调度列表页面,在1中查询条件输入框中输入所要查找的调度任务的信息,单击【查询】按钮则在列表中显示查询到的调度任务。在列表中

能看到任务的 ID、同步任务名称、任务版本、任务状态、调度类型、执行周期、创建和近一次调度时间、近一次的调度状态信息。同步任务执行情况如下。



对于执行的调度任务在调度任务列表中显示,可执行的操作有立即执行、历史日志。

3.4.3.1 同步任务调度列表操作

3.4.3.1.1 立即执行

对于已经执行的调度任务,在列表中支持直接单击【立即执行】按钮重新执行该任务的调度。处于执行中的任务会仍然继续执行不受影响。点击立即执行页面提示"任务立即执行成功",则任务开始同步调度,状态为"执行中",产生一条任务实例在历史日志中能够查看到执行日志。



3.4.3.1.2 历史日志

单击【历史日志】按钮,列表显示该任务的历史执行情况,内容包括任务名称、执行时间、执行状态、执行时长(秒)、同步数据量(条)等内容,能够执行的操作有终止、重跑和查看日志。支持用户根据执行时间和执行状态选择用户所需要的任务实例进行快速查询,在查询条件中输入查询信息后,单击【查询】按钮即可。具体情况如下图所示。



如果历史任务中有执行成功或者失败的任务,即已经产生结果的任务实例,则显示重跑操作,单击【重跑】按钮能够使得该任务重新执行一次。



2) 如果调度任务正在执行当中,用户想要使得该任务立即停止,单击【终止】按钮则直接结束调度任务的执行进程,确认终止后单击【确定】则该次任务调度执行结果为失败。



3) 如果用户想要查看该调度任务实例执行的具体情况,则用户可以单击【查看 日志】按钮查看该任务实例的执行日志。具体情况如下。

[NFO] 2021-11-16 20 23:14.773 - [uskAppid-TASK-1460582069764435970-1460582991173349377-1460584241751224321][116] - create dir success /mp/dolphinscheduler/exec/process/icredit/1460582069764435970-1460582991173349377-1460584241751224321][116] - datax task params ["customConfig-11,"pson-"\"("core"\"("transport)\"("channel\"\"1\"("channel\"\"2\") cored-"\"("cored-\"("transport)\"("channel\"\"2\") cored-"\"("cored-\"("transport)\"("channel\"\"2\") cored-"\"("cored-\"("transport)\"("cored-\"("transport)\"\"("channel\"\"2\") cored-"\"("cored-\"("transport)\"\"("co

cd \$BASEDIR

source (inclingfile)

where (inclingfile)

source (inclingfile)

where (inclined inclined incli

osinfo: Oracle Corporation 1.8 25.241-b07 jymlnfo: Linux amd64 3.10.0-1062.el7.x86_64 cpu num: 4

若想要看所有同步任务调度的总体情况则到 3.4.2 调度首页中查看总体概览。