

ALIBABA CLOUD

阿里云

DataWorks
任务运维

文档版本：20211230

 阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1. 运维中心	06
1.1. 运维中心概述	06
1.2. 查看运维大屏	12
1.3. 实时任务运维	21
1.3.1. 实时计算任务运行与管理	21
1.3.2. 实时同步任务运行与管理	23
1.4. 周期任务运维	28
1.4.1. 查看周期任务	28
1.4.2. 查看周期实例	36
1.4.3. 执行补数据并管理补数据实例	44
1.4.4. 实例DAG图	57
1.4.5. 测试实例	63
1.4.6. 运行诊断	66
1.5. 手动任务运维	68
1.5.1. 手动任务	68
1.5.2. 手动实例	70
1.6. 智能监控	72
1.6.1. 智能监控概述	72
1.6.2. 智能基线	80
1.6.2.1. 基线管理	80
1.6.2.2. 基线实例	83
1.6.2.3. 事件管理	84
1.6.3. 自定义规则	85
1.6.3.1. 自定义规则	85
1.6.3.2. 值班表	92
1.6.4. 自动运维	96

1.6.5. 报警信息	100
1.6.6. 常见问题	101
1.7. 资源运维	102
1.8. 引擎运维	105
1.9. 常见问题	108
1.9.1. 数据未正常产出或任务未运行如何处理	108
1.9.2. 找不到任务实例如何处理	109
2.任务发布	111
2.1. 发布任务	111
2.2. 下线任务	116
3.跨项目克隆	121
3.1. 跨项目克隆说明	121
3.2. 跨项目克隆实践	122

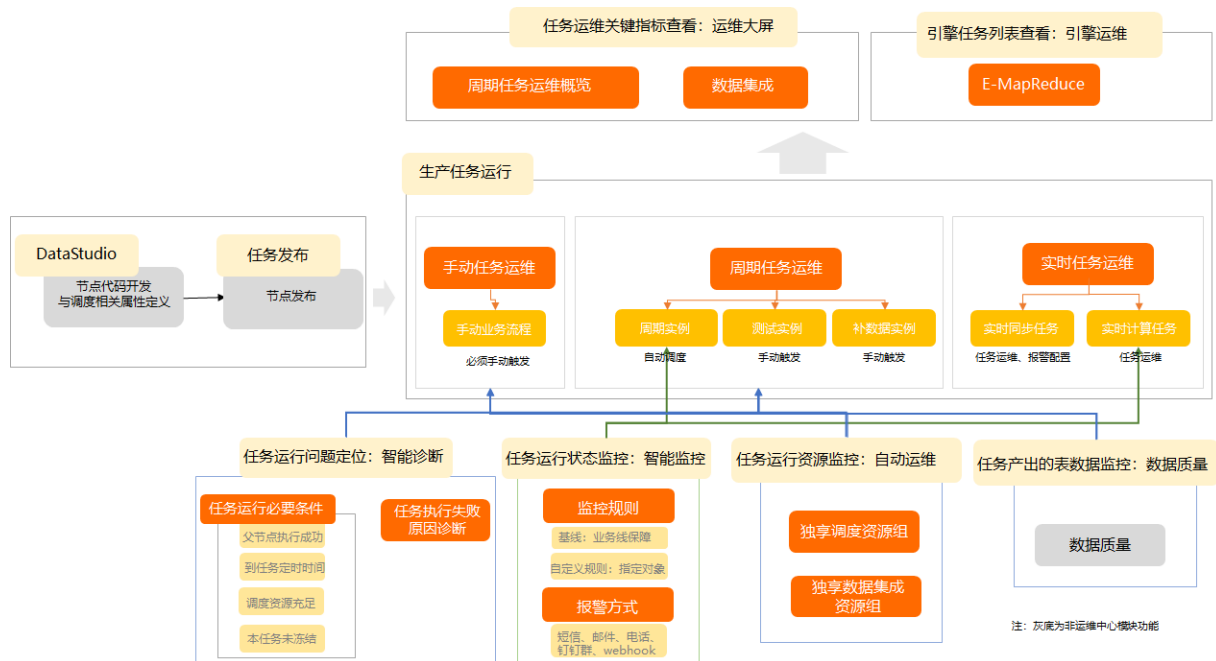
1.运维中心

1.1. 运维中心概述

运维中心包括运维大屏、实时任务运维、周期任务运维、手动任务运维、智能监控、智能诊断、资源运维和引擎运维功能模块，为您提供任务操作与状态、引擎、资源等多方位的运维能力。

运维中心功能模块

当您在DataStudio中完成节点开发，并提交和发布至生产环境后，即可在运维中心对任务进行运维操作，包括周期调度任务的自动调度与手动运行、任务运行详情查看、任务运行状态监控、任务运行使用的资源监控与自动运维；实时任务的执行管控、运行详情查看与监控报警配置；调度任务运维大屏以及数据集成离线同步与实时同步任务运维专区，针对任务运维关键指标的查看。



下表为运维中心各模块功能使用的简单说明：

模块	描述	环境支持情况
查看运维大屏	运维大屏以报表的形式为您展示任务的运行的重要指标。包括调度任务的关键运维指标统计与数据集成离线同步与实时同步任务运维专区。	开发环境运维中心无此模块
实时同步任务运行与管理	在实时任务运维界面，您可以针对实时任务进行启动、停止、下线与监控报警相关操作。	-
手动任务	周期任务运维为您展示任务提交到调度系统后的调度任务列表与调度环境产生的周期实例。周期任务面板展示当前所有的调度节点，您可以在周期任务面板修改周期调度任务使用的资源组、责任人等信息。 DataWorks每晚会根据周期任务自动生成第二天待运行的周期实例，同时支持您手动触发周期任务，分别生成补数据实例和测试实例，您可以通过生成的实例查看任务的运行详情。	开发环境运维中心无法自动调度生成周期实例。

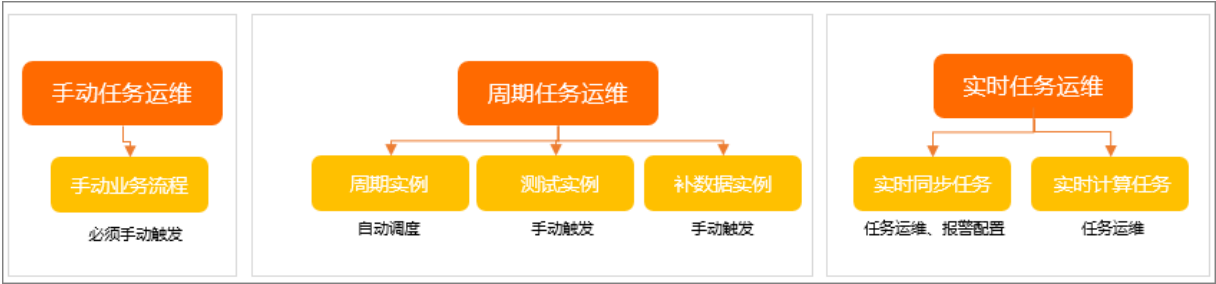
模块	描述	环境支持情况
手动任务运维	手动任务运维为您展示任务提交到调度系统后的手动任务列表与经过手动触发运行后生成的手动实例。手动任务面板展示当前所有的手动触发式任务，您可以在手动任务面板修改任务执行时使用的调度资源组，责任人等信息。 在手动任务面板中，选择类型为手动业务流程，支持您手动触发任务运行并且生成手动实例，您可以查看生成的手动实例的运行详情。	-
智能监控	- 监控对象：周期任务、实时计算任务、资源水位。 监控支持为自动调度的周期任务设置监控报警，监控在自动调度场景下的周期实例运行情况、支持独享资源组水位监控及实时计算任务运行状态监控。 - 监控方式：自定义规则、智能基线 支持针对指定对象（节点、业务流程、工作空间、基线、实时计算任务、独享调度资源组、独享数据集成资源组）的监控报警，同时也支持针对整个“业务线”的基线预警，并通过您配置的报警方式来发送相应的报警信息。 - 报警方式：支持短信、邮件、电话、钉钉群和webhook，已支持设置值班表，将对应时间的报警消息报警给对应的值班人员。	开发环境运维中心 无此模块
资源运维	资源运维主要对 独享调度资源组 及 独享数据集成资源组 的使用情况（使用率、运行中任务列表、排队任务列表）及进行监控和自动运维，支持查看资源组使用趋势。	-
引擎运维	目前支持对EMR引擎的运维。在DataWorks实例任务中，E-MapReduce计算引擎任务会被分为若干作业有序执行，您可以使用DataWorks的引擎运维功能，查看各个E-MapReduce作业的详细信息，及时查找并清理运行有误的作业，避免该类作业阻塞下游任务，影响实例任务的正常运行。	开发环境运维中心 无此模块
运行诊断	DataWorks运维中心为您提供运行诊断功能，帮助您掌握任务运行的全链路信息，以便迅速定位问题。	开发环境运维中心 无此模块

同时，需关注以下注意事项：

- 在标准模式工作空间下，您可以通过界面左上角切换进入到开发运维中心和生产运维中心。
- 仅生产环境运维中心支持自动调度。即只有在生产环境运维中心，您才能在周期实例面板中看到自动调度执行的周期实例。

生产任务运行逻辑

生产任务包括周期任务、手动任务和实时任务。



● 周期任务运维

当节点被提交发布至运维中心后，您可以在运维中心的周期任务中看到该节点（周期任务列表中的任务为所有周期性调度的任务），DataWorks每晚会自动生成第二天待运行的周期实例，同时支持您手动触发周期任务，分别生成补数据实例和测试实例。

实例类型	应用场景	与周期任务的关系 (对应实例如何生成)	实例触发方式 (对应实例如何触发执行)	任务运行条件
周期实例	周期性自动化进行ETL操作。	平台每晚根据某一时间点周期任务快照信息自动生成第二天待运行的周期实例。 ❓ 说明 开发环境运维中心不会自动调度生成周期实例。	平台自动触发周期实例执行。	以周期实例上的任务为例，任务开始运行前需满足以下条件： - 依赖的所有父节点实例状态都为成功状态。 - 已到达任务节点设置的任务定时运行时间。 - 调度资源充足。 - 本任务不是冻结状态。
补数据实例	需要对当前节点及下游节点执行补过去某一段时间或未来某段时间的数据的操作，即需要对历史或未来时间段的数据进行ETL操作。	人为针对当前周期任务手动触发补数据操作，并生成补数据实例。	人为针对对周期任务进行补数据操作后，会生成补数据实例，并同时触发补数据实例执行。	❓ 说明 运维中心中，不同的实例颜色代表实例处于不同的状态中，实例运行状态示意详情可参考：附录：实例状态。
测试实例	需要对当前周期任务进行测试，观察执行结果是否符合预期。 ❓ 说明 测试实例的运行实际上也在真实执行代码逻辑的。	人为针对当前周期任务手动触发测试操作，并生成测试实例。	人为针对周期任务的进行测试操作后，会生成测试实例，并且同时触发测试实例执行。	

您可以在周期任务运维下的周期实例面板、补数据实例面板、测试实例面板中查看各类实例的运行情况，根据实例的状态了解数据产出是否异常，不同实例状态的详情请参见附录：实例状态。

② 说明

- 周期实例根据周期任务每晚（23:30~24:00）定时自动生成，测试实例、补数据实例根据补数据和测试操作时间点的周期任务手动触发，因此补数据和测试会以最新的任务配置情况生成对应的实例。
- 新建节点发布生产环境后，当晚会全量生成第二天待运行的周期实例，但新建节点发布时，您可以通过节点调度配置，选择实例生成方式来控制是否立即生成当天的周期实例，详情请参考[时间属性配置说明](#)。

● 手动任务运维

当手动业务流程被提交发布至**运维中心**后，您可以在**运维中心 > 手动任务运维 > 手动任务**页面中看到该节点，当您需要运行该手动业务流程时，可以在手动任务面板对该任务执行运行操作，手动触发任务运行后，会产生手动实例，您可以在**手动实例**页面查看该实例执行详情。

● 实时任务运维

当实时任务被提交发布至**运维中心**后，您可以在**运维中心 > 实时任务运维**页面对实时任务进行启动、下线及报警配置等相关操作，实时任务运维分为实时计算任务和实时同步任务，您可以分别在实时任务运维下的实时计算任务、实时同步任务页面对其进行管理。详情请参见：[实时计算任务运行与管理](#)、[实时同步任务运行与管理](#)。

任务运行状态监控：智能监控

您可以使用**智能监控**功能监控周期调度任务状态、周期任务自动调度产生的**周期实例**运行状态、实时计算任务运行状态，配置相应监控规则（自定义规则、智能基线）后，平台一旦检测到规则被触发，将会根据您设置的自定义规则及智能基线的报警方式，通过邮件、短信、电话或钉钉群消息等形式通知您，方便您及时发现并处理异常问题。

● 周期任务状态监控

DataWorks每晚根据**周期任务**生成第二天待自动调度运行的**周期实例**，所以为保障周期任务可以正常产生周期实例并且自动调度运行，DataWorks内置了报警规则来对**周期任务**进行定期监控扫描，如有异常便会自动报警。

系统内置全局报警规则（非空间级别报警规则）对周期任务的状态进行监控。报警包括**孤立节点与节点成环**。

② 说明

- DataWorks每天定时9点、12点、16点对周期任务状态进行扫描，如有异常将会自动发送报警。但扫描时间点前10分钟内产生异常不会纳入本次扫描，该异常将会被纳入下一个周期的任务状态扫描中进行扫描。
- 全局规则为系统内置规则，您无须手动新建，默认以短信、邮件的方式报警给节点责任人。但您可以在**规则管理**页面中针对全局规则修改报警接收人。

○ 节点孤立报警

孤立节点指节点在周期任务或周期实例中展开父节点时，没有依赖任何父节点，这类型节点在自动调度时不会被调度调起，所以孤立节点不会自动调度运行，如果该孤立节点下游依赖较多，则会造成严重的后果。孤立节点产生后会自动报警，如果收到孤立节点报警请及时处理。

② 说明 在DataWorks上，除了工作空间根节点外，自定义的每个周期调度的节点都需要有父节点依赖才可以被正常调度运行。

- 节点成环报警

节点成环指的是某节点为上游节点但又同时依赖了自己的下游节点。导致依赖关系成环，这类型节点在自动调度时不会被调度调起，节点成环后会自动报警，如果收到节点成环报警请及时处理。

- 周期实例运行状态监控

在DataWorks上，周期任务周期性调度时将产生周期实例，DataWorks支持通过周期任务自定义规则监控配置，实现对周期实例运行状态的监控，包括指定对象的自定义规则监控报警和重要任务提前预警的基线预警功能。

- 自定义规则报警

通过新建自定义规则实现对指定对象的报警。

您可以对指定对象（任务节点、基线、工作空间、业务流程）上节点的状态（完成、未完成、出错、周期未完成、超时、自动重跑后仍出错等及实时计算任务的出错）设置监控报警。支持短信、邮件、电话、钉钉机器人和webhook的方式发送报警给责任人（节点责任人、指定责任人或值班表中的排班等），详情请参考[自定义规则](#)。

- 基线预警

通过设置基线优先级来保障基线上的任务按时执行，数据顺利产出。

当有任务需要重保并且上游任务依赖较为复杂时，可以使用基线功能将重要任务移到该基线上。当影响基线在预期时间内数据产出的情况发生时，基线将为您快速定位影响当前节点数据产出的关键路径，找到阻塞任务（即影响保障数据产出的关键实例）并第一时间发出报警，保障重要数据在预期时间内顺利完成。当影响基线的任务（基线上的任务及其上游任务）出错或者变慢，您将会收到相应的报警。其报警分为全局基线预警与全局事件报警（非空间级别报警规则）。支持短信、邮件、电话、钉钉机器人和webhook的方式发送报警给责任人（节点责任人、指定责任人或值班表中的排班等），关于基线使用详情请参见：[基线管理](#)。

- 全局基线预警

当基线根据基线的承诺时间、预警余量等预测无法保障基线上的任务在预警时间内完成时，便会触发基线提前预警功能。

- 全局事件报警

当影响基线上数据产出的任务（基线上的任务及其上游任务）出错或者变慢时，将会收到相应的事件报警。

 **说明** 全局规则为系统内置规则，您无须手动配置，但您可以在规则管理页面中针对全局规则修改报警接收人。

- 实时计算任务运行状态监控

您可以通过自定义规则，对实时计算任务的运行状态进行监控，当任务节点出错时通过配置的短信、邮件、电话、钉钉机器人和webhook等方式发送报警给责任人（节点责任人、指定责任人和值班表中的排班等）。详情请参考[自定义规则](#)

任务运行资源监控与运维：自动运维

您可以使用[智能监控 > 规则管理](#)中的自定义规则，来对独享资源组使用率及独享资源组上等待的实例数设置监控报警，并且支持结合[自动运维](#)功能通过短信、邮件、电话、钉钉机器人等方式发送报警给责任人（节点责任人、指定责任人）。关于自定义规则使用详情请参考[自定义规则](#)。

[自动运维](#)功能支持通过自定义规则中的独享资源组监控规则，并在限定条件下对独享资源组进行运维操作来保障重要任务的顺利运行，比如终止运行中的实例。详情请参考[自动运维](#)。

任务运行问题定位：运行诊断

● 调度任务运行的必要条件

节点发布至运维中心后，任务开始运行前需满足以下条件。

 **说明** 当任务未运行时您可以使用智能诊断功能快速进行对问题进行定位与排查，详情可参考文档：[运行诊断](#)。

i. 依赖的所有父节点实例状态都为成功状态

配置节点依赖关系可以保障节点在运行时能取到正确的数据（读取上游表数据时，上游表的最新数据已产生），避免当前节点读取数据时，上游表数据还未产出，导致当前节点读取数据出现问题，所以，当节点存在上游节点依赖时，当前节点的执行，需要等待所依赖的父节点都执行成功。详情请参见：[同周期调度依赖逻辑说明](#)。

说明

- 父节点非成功状态包括父节点未运行、父节点状态失败、父节点被冻结（暂停）、父节点处于运行中。
- 冻结状态的节点将阻塞下游节点执行。

ii. 已到达任务节点设置的定时运行时间

数据开发中定义节点的调度属性时，需要为节点设置调度环境下节点运行的“理想时间”。

- 当当前节点依赖的所有父节点都在当前节点设置的定时运行时间点前执行成功，那么当父节点全都执行成功后，当前节点将进入等待状态，待到达任务节点的定时时间后启动运行。
- 当当前节点依赖的所有父节点在当前节点设置的定时运行时间点后执行成功，那么当父节点全都执行成功且有充足的调度资源时，当前节点将立即运行起来。

iii. 调度资源充足

数据开发中定义节点的调度属性时，您可以设置调度环境下节点运行所使用的独享调度资源组。同时，您也可以在运维中心的周期任务页面进行修改。

当当前节点所有的父节点运行成功，并且已经到达节点设置的定时运行时间时，如果节点指定的调度资源组无可用资源时，节点进入等待资源的状态，直到占用该调度资源的任务执行成功，将资源释放后启动运行。

iv. 本任务不是冻结状态

如果当前节点处于冻结状态，那么节点满足上述三个条件后仍不会运行，并导致依赖该周期实例的下游实例无法运行。如果您需要解冻该节点，详情请参见：[任务冻结与解冻](#)。

● 任务运行失败原因诊断

您可以使用智能诊断功能和DAG的上游分析功能对未运行任务进行快速诊断。

其中，智能诊断功能支持对未开始运行的任务进行任务运行的必要条件检查，支持对运行失败的任务进行原因分析并给出诊断建议等。

智能诊断能力概览：

- 任务运行的全链路诊断。包括父节点运行情况检查、当前节点设置的定时时间检查、调度资源详情（任务使用的调度资源组使用趋势与任务排队数，当前任务等待资源时哪些任务占用资源）、任务执行详情（任务出错时提供原因分析与诊断建议）。

- 当前实例的基本信息与实例运行的关键时间点（上游节点运行完成时间、当前节点设置的定时时间、开始等待调度资源时间、开始运行时间、运行成功时间）。
- 当前节点影响的基线及其当天基线实例运行情况。
- 当前节点平均运行时间、开始执行时间、等待资源消耗时间趋势图及当前节点运行历史的实例列表。

关于智能诊断功能使用详情请参见：[运行诊断](#)。

任务产出的表数据监控：数据质量

您可以使用数据质量功能对周期实例、补数据实例、测试实例产出的表数据进行监控。

周期任务在生产环境通过自动调度或手动触发来运行，分别产生周期实例，补数据实例和测试实例，您可以对周期任务设置数据质量规则来监控周期任务生成的实例（周期实例、补数据实例、测试实例）运行后所产出的表数据是否符合预期。

数据质量通过表配置的分区表达式来匹配节点产出的表分区，数据质量规则关联产出该表数据的调度节点，当节点运行后（运维中心通过测试实例，补数据实例，周期实例的运行及重跑操作产出表数据），会触发该数据质量规则校验，您可以设置规则的强弱来控制节点是否失败退出，从而避免脏数据影响扩大，同时，您可以通过订阅规则来第一时间接收节点产出表数据异常的报警。详情请参考文档[数据质量](#)。

附录：实例状态

运维中心按照不同的颜色与标识来辨识任务正处于运行流程中的哪个阶段，不同的实例颜色与标识代表实例处于不同的状态下。不同的实例颜色与标识所对应的任务状态如下所示。关于任务运行必要条件详情请参见：[任务运行问题定位：运行诊断](#)。

序号	状态类型	状态标识	运行流程图
1	运行成功状态		
2	未运行状态		
3	运行失败状态		
4	正在运行状态		
5	等待状态		
6	暂停/冻结状态		

1.2. 查看运维大屏

运维大屏展示当前您需要重点关注的运维指标、任务整体运行情况、调度资源变化趋势等运维概览，以及离线同步、实时同步任务的运行状态分布、数据同步进度等数据集成信息，帮助您提升任务运维效率。

进入运维大屏

1. 登录[DataWorks控制台](#)。

2. 在左侧导航栏，单击工作空间列表。
3. 选择工作空间所在地域后，单击相应工作空间后的进入数据开发。
4. 单击左上方的图标，选择全部产品 > 任务运维 > 运维中心（工作流），默认进入运维大屏 > 运维概览页面。

查看运维概览

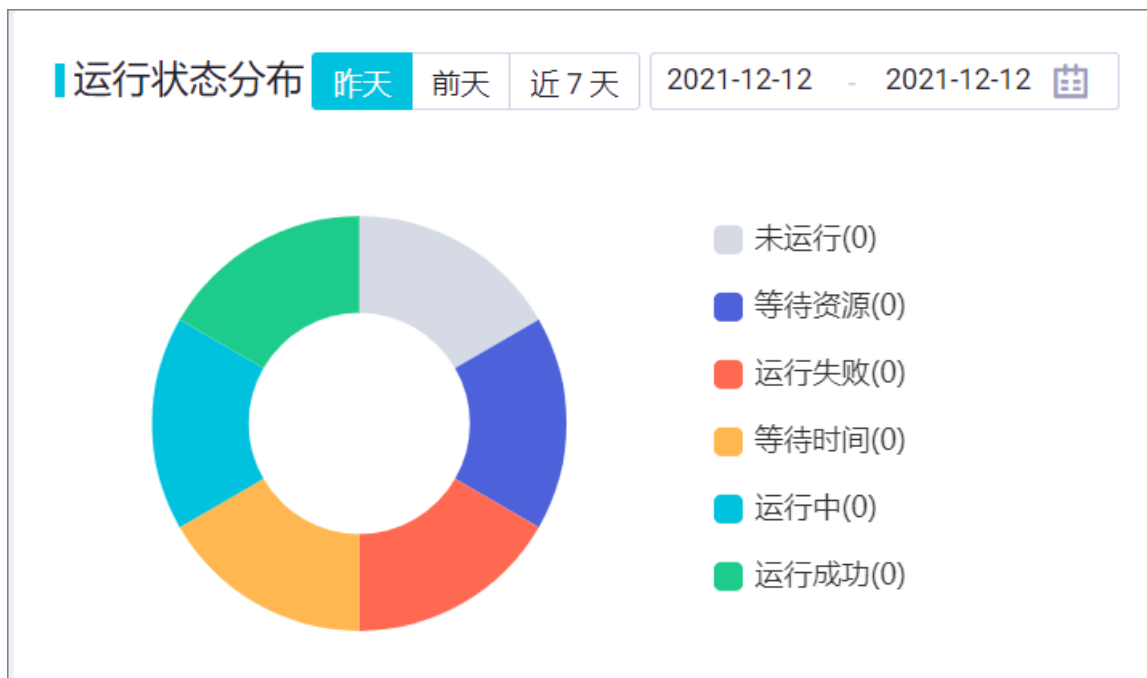
目前运维概览页面的统计范围是周期任务或周期实例，不包括其它类型的任务和实例。您可以在运维大屏页面查看以下内容：

- 重点关注区域展示以下内容：
 - 您当前需要重点关注的周期实例，包括失败实例、运行慢实例、等资源实例。同时满足以下条件的实例即为运行慢实例：
 - 实例处于运行中的状态。
 - 实例运行超过30分钟。
 - 实例的运行时长与过去10天的平均运行时长相比，已经超过15分钟及以上。
 - 您当前需要重点关注的孤立节点、暂停节点和过期节点：
 - 孤立节点此处统计的为周期任务中没有父节点的节点个数。当节点没有上游依赖时，该节点会变为孤立节点，将无法再自动调度运行。例如，修改父节点输出名称后，造成依赖丢失。关于孤立节点产生原因、影响及解决方案详情，您可以参考文档：[什么是孤立节点，如何处理孤立节点？](#)。
 - 暂停节点此处统计的为暂停（冻结）状态的周期任务。周期任务暂停后，该周期任务产生的实例也将是冻结状态，冻结状态的实例将不会运行，并且会阻塞其下游节点执行。关于冻结操作详情介绍与影响说明可参考文档：[任务冻结与解冻](#)。
 - 过期节点此处统计的为节点调度生效日期已过的节点。节点在调度生效日期内将自动产生周期实例并运行，调度生效日期范围外将无法产生周期实例并且自动调度。

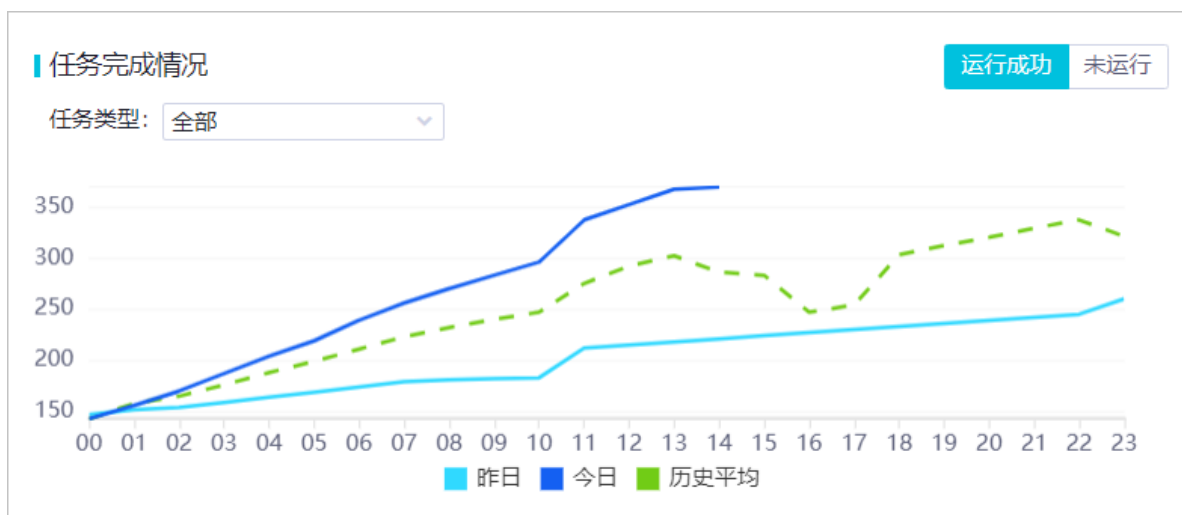


重点关注根据您的需求可以选择统计昨天、前天、近7天或指定时间段内（最多支持7天）的数据。您可以单击某个指标进入相应的节点或实例列表，及时关注相关指标并进行治理，以避免影响业务。

- 运行状态分布区域展示当前所有业务日期为昨天、前天、近7天或指定时间段内（最多支持7天）正常调度实例的运行状态分布，统计的是页面请求发生时的数据。您可以单击饼图中的某个色块，进入相应状态的实例列表。



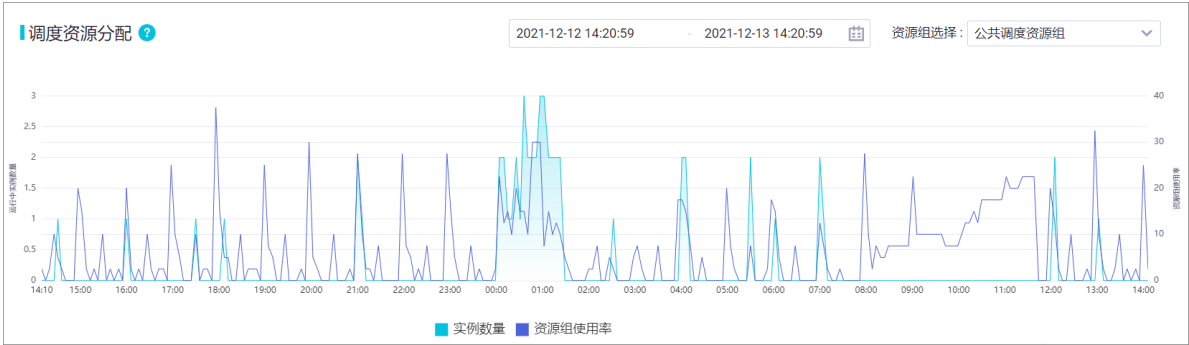
- **任务完成情况** 区域展示页面请求当天00:00~23:00内，当前工作空间中周期实例的完成情况。您可以查看昨日、今日和历史平均的运行成功或未运行的任务，并可以指定需要查看的任务类型。



对比正常周期性调度昨日、今日和历史平均的任务完成情况，统计出折线图。如果三条折线偏移过多，表示在某个时间段内有异常情况，您需要进一步检查与分析。

🔍 **说明** 由于运维中心统计方法的变更（仅统计生产环境的实例），今日任务完成情况折线会产生明显波动。

- **调度资源分配** 区域默认展示指定时间段内（最多支持7天）调度资源组使用率与各时间点运行在该资源组上的实例数量的变化趋势。您可以单击右上方的资源组选择下拉列表切换资源组分别查看当前工作空间下的各调度资源组，其资源组使用率与运行的实例数变化趋势。



实例数量为对应时间段运行在该资源组上实例数量，资源组使用率为对应时间段运行在该资源组上的实例所占用的资源的百分比。

说明

- 资源组使用率不记录数据集成任务占用的资源。
- 资源组使用率与资源组上运行的实例数量为资源组级别，例如：如果您使用的独享调度资源组为多个工作空间共享，那么此处统计的为多个工作空间下，该资源组总的资源使用率和实例数量变化趋势。

- 运行时长排行区域展示运行时长、等待资源时长和运行变慢时长的排行。该区域每日更新数据，对运行在前一天，且运行成功的实例的运行时长进行排行。

运行时长排行

运行时长排行

等待资源时长排行

变慢排行

上次更新: 2021-10-08

	节点名称	负责人	执行时长
1	7000048746...	测试小时任务是否并行	2分
2	7000048746...	测试小时任务是否并行	1分25秒
3	7000048746...	测试小时任务是否并行	1分8秒
4	7000048834...	测试告警	1分4秒
5	7000048377...	oss_odps_di	59秒

- 近一个月出错排行区域每日更新数据，统计最近一个月的任务出错排行，为您展示任务出错次数最多的前10个节点。您可以查看节点ID、节点名称和出错次数。

近一个月出错排行

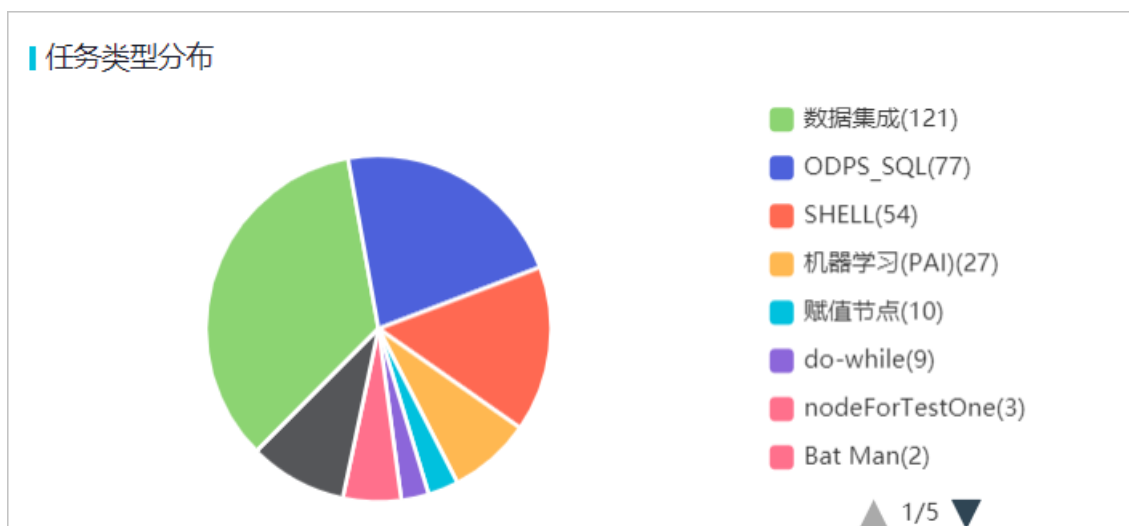
上次更新: 2020-06-18

	节点ID	节点名称	负责人	出错次数
1		mi		256
2				87
3				46
4				43
5				18
6				17
7				16
8				16
9				16
10				16

- 调度任务数量趋势区域每日更新数据，展示在一段时间内，生产环境的周期任务、周期实例的数量变化趋势。您最多可以查看近一年的调度任务数量变化。



- 任务类型分布区域展示页面请求发生时，各种类型的任务数量。饼图最多支持展示8种任务类型，如果统计的任务超过8种类型，会进行合并展示。



查看离线同步运维

在运维大屏页面，单击数据集成，默认展示某个时间范围内离线同步的信息。您可以在右上方设置统计数据的时间范围。

启动

是否重置位点：☐ 重置位点

启动时间点位：

时区：

Failover: 分钟内, Failover次数超过 任务自动结束

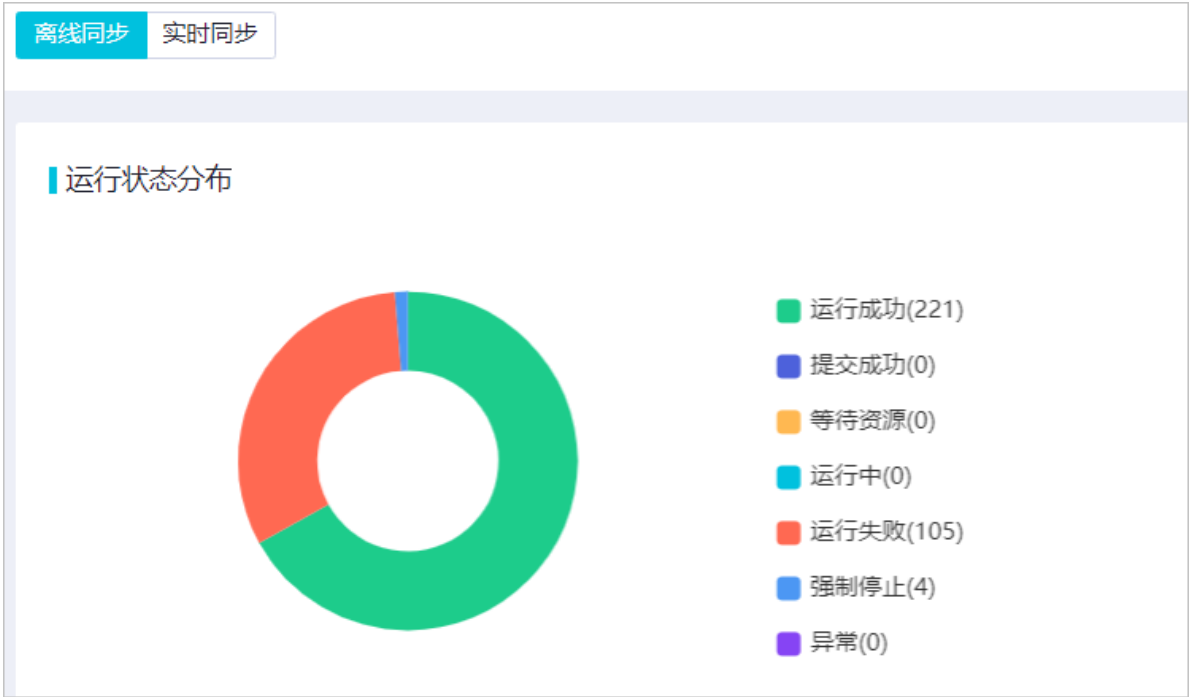
* 脏数据策略: ☒ 零容忍, 不允许 ☐ 不限制 ☐ 有限控制

确定

取消

您可以查看离线同步任务的以下内容：

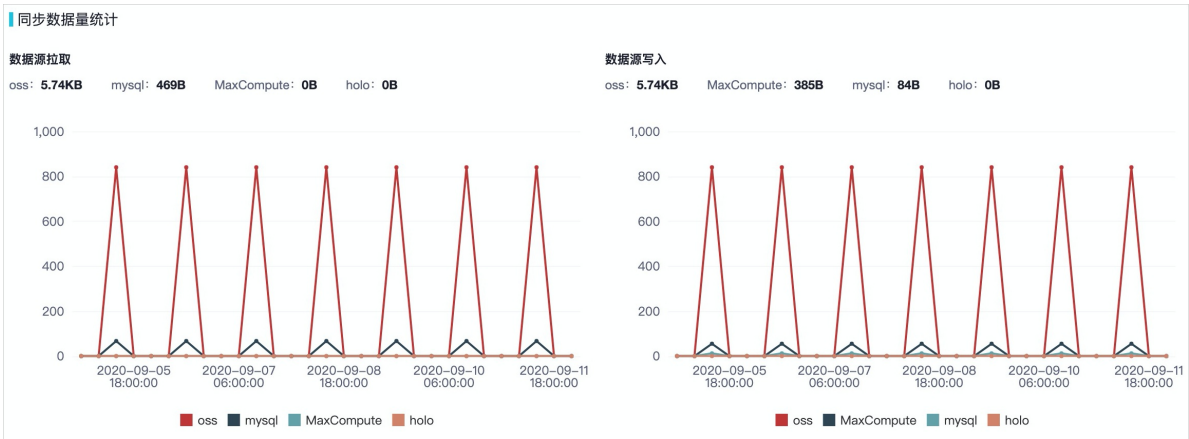
- 运行状态分布区域展示所选业务日期内调度实例的运行状态分布，统计的是页面请求发生时的数据。您可以单击饼图中的某个色块，进入相应状态的实例列表。



- 数据同步进度区域展示所选业务日期内离线同步的数据量，包括总数据量、公网总流量和记录总量。



- 同步数据量统计区域根据数据源类型展示所选业务日期内，同步数据量的拉取和写入曲线。



- 最新榜单Top10区域分别展示10个最新失败实例和最新成功实例，以便您全局查看最新的同步任务状态。
- 同步任务执行详情区域支持通过筛选提交时间、任务状态和任务名称等条件，快速搜索任务实例，并可以单击实例ID，查看该实例的运行详情。

同步任务执行详情

提交时间: 起始日期 - 结束日期 任务状态: 全部 任务名称: 来源数据源类型: 全部 来源数据源名称: 数据集成资源组: 全部 去向数据源类型: 全部 去向数据源名称: 是否有公网流量: 全部

实例ID	任务名称	任务状态	提交时间	启动时间	结束时间	来源数据源	去向数据源	数据集成资源组	运行并发数	是否有公网流量	已同步数据量	已同步记录数	脏数
331888674		运行成功	2020-09-21 13:10:13	2020-09-21 13:10:21	2020-09-21 13:10:51	xc_mongo_rds	odps_first		2	是	223B	4	0
331868045		运行成功	2020-09-21 12:10:08	2020-09-21 12:10:12	2020-09-21 12:10:48	xc_mongo_rds	odps_first		2	是	223B	4	0
331860823		运行失败	2020-09-21 12:00:13	2020-09-21 12:00:24	2020-09-21 12:00:32	odps_first	odps_first		2	否	0B	0	0

查看实时同步运维

在运维大屏 > 数据集成页面，单击实时同步，查看实时同步任务的以下内容：

- 运行状态分布区域展示当前实时同步实例的运行状态分布，统计的是页面请求发生时的数据。您可以单击饼图中的某个色块，进入相应状态的实例列表。



- 概览区域展示当前工作空间下，所有实时任务的数据速度和记录速度的总和。



- 任务延迟Top10区域展示延迟最高的10个任务，以便您迅速定位高延迟任务。

任务延迟 Top10			
	任务名称	业务延迟	数据集或共享资源组
1	OB_2_DATAHUB	23 毫秒	
2	mysql_to_datahub	24 毫秒	

- 任务报警信息区域展示最近的报警信息，以便您快速查看异常情况。

任务报警信息							
发生时间	任务ID	任务名称	接收人	报警级别	指标	设定阈值	报警时间值
2020-09-15 14:57:47	287			WARNING	心跳告警		
2020-09-15 14:57:16	408			WARNING	lxxx	0	
2020-09-15 14:52:47	287			CRITICAL	心跳告警		
2020-09-15 14:52:47	287			CRITICAL	心跳告警		
2020-09-15 14:52:47	287			WARNING	心跳告警		
2020-09-15 14:52:47	287			WARNING	心跳告警		

- Failover信息区域展示指定时间内的Failover消息，以便总览任务Failover的状态。

Failover信息

在最近 分钟内共Failover 0 次

时间	实例ID	任务名称	Failover事件
----	------	------	------------

1.3. 实时任务运维

1.3.1. 实时计算任务运行与管理

DataWorks的实时计算任务页面为您展示了任务实例的基本信息及运行详情，方便您查看目标任务的运行情况。同时，您还可以为目标任务设置监控报警，在任务运行出错时及时发现并处理异常。

使用限制

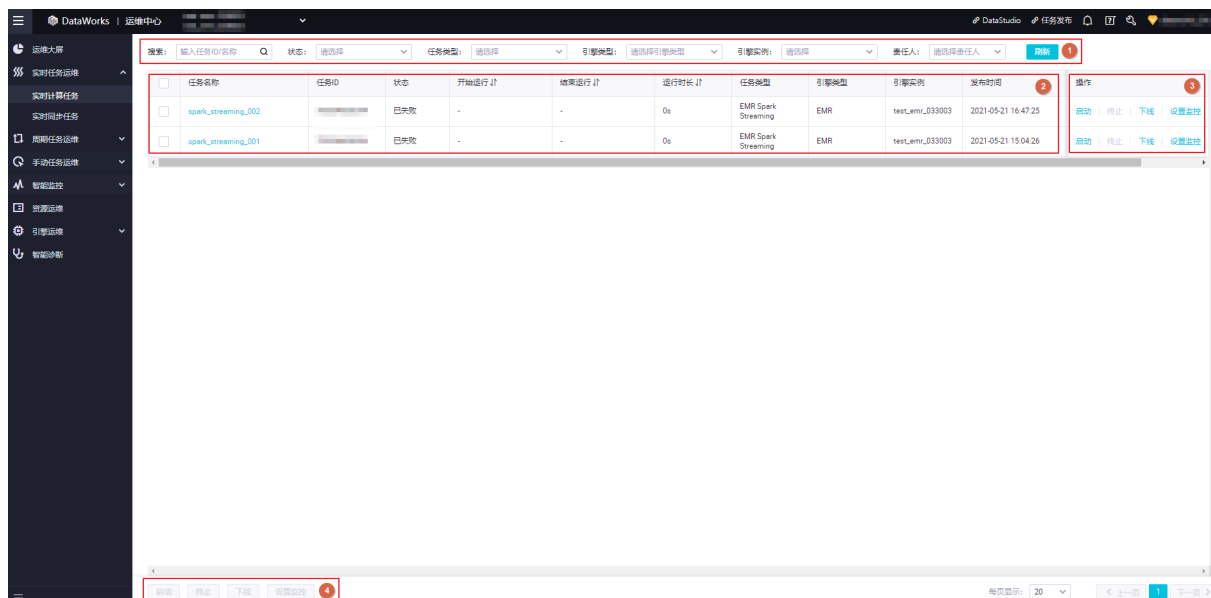
DataWorks当前仅支持对EMR Spark Streaming及EMR Streaming SQL类型的节点提交的任务进行运维。

进入实时计算任务页面

1. 登录DataWorks控制台。
2. 在左侧导航栏，单击工作空间列表。
3. 选择工作空间所在地域后，单击相应工作空间后的进入数据开发。
4. 单击左上方的图标，选择全部产品 > 任务运维 > 运维中心（工作流）。
5. 在左侧导航栏，单击实时任务运维 > 实时计算任务，进入实时计算任务页面。

查看实时计算任务列表

运维中心的实时计算任务页面为您展示了生产环境的实时计算任务列表，您可以在该页面查看任务的基本信息，执行启动、终止、下线等运维操作，并设置任务出错监控。



截图显示了DataWorks的实时计算任务列表界面。左侧是导航栏，包含“运维中心”、“实时任务运维”、“实时计算任务”、“实时同步任务”、“离线任务运维”、“手动任务运维”、“智能监控”、“资源运维”、“引擎运维”和“智能诊断”等选项。右侧是任务列表，包含任务名称、任务ID、状态、开始运行时间、结束运行时间、运行时长、任务类型、引擎实例、发布时间和操作按钮。当前列表中有两个任务：spark_streaming_002 和 spark_streaming_001，状态均为“已失败”。操作按钮包括“启动”、“终止”、“下线”和“设置监控”。

任务名称	任务ID	状态	开始运行时间	结束运行时间	运行时长	任务类型	引擎实例	发布时间	操作
spark_streaming_002		已失败	-	-	0s	EMR Spark Streaming	test_emr_033003	2021-05-21 16:47:25	启动 终止 下线 设置监控
spark_streaming_001		已失败	-	-	0s	EMR Spark Streaming	test_emr_033003	2021-05-21 15:04:26	启动 终止 下线 设置监控

区域	描述
1	在该区域，您可以通过任务ID、任务名称、状态、任务类型、引擎类型、引擎实例及责任人等筛选条件过滤出需要查询的实时计算任务。  说明 - 任务ID或任务名称搜索的结果会受到其它筛选条件的影响。只有同时满足所有筛选条件时，才会显示相应的查询结果。 - 任务名称搜索支持模糊匹配，输入关键词后，即可显示包含关键词的所有实时计算任务。
2	在该区域，您可以查看目标任务的基本信息及运行情况。 任务名称：为您展示任务的名称。单击任务名称，即可进入目标任务的详情页面。 任务ID：为您展示任务的ID。 运行状态：任务的运行状态，包括未运行、生成中、等待资源、启动中、运行中、已停止、已失败、重启中以及下线中等状态。 开始运行：任务的开始运行时间。 结束运行：任务的结束运行时间。 运行时长：任务的运行时长，单位为秒。 任务类型：任务的类型。  说明 DataWorks当前仅支持对EMR Spark Streaming及EMR Streaming SQL类型的节点提交的任务进行运维。 引擎类型：任务所使用的引擎类型。DataWorks当前仅支持对E-MapReduce引擎的实时计算任务进行运维。 引擎实例：任务所绑定的E-MapReduce引擎名称。 发布时间：任务的发布时间，格式为yyyy-MM-dd HH:mm:ss。 责任人：任务所在工作空间的责任人。  说明 您可以根据任务的开始运行、结束运行以及运行时长，按照时间的升序或降序对目标任务进行排序。
3	在该区域，您可以对目标任务进行如下操作。 启动：启动未运行的实时计算任务。 终止：终止运行中的实时计算任务。 下线：下线已提交发布的实时计算任务。 设置监控：用于监控目标实时计算任务的运行情况。当任务运行出错时，会根据您设置的短信、邮件、电话、钉钉群机器人或WebHook等方式发送报警至相应的报警接收人。 配置报警接收人所使用的邮件、电话信息，详情请参见设置报警接收人的手机或邮箱信息，配置钉钉群机器人及获取Webhook地址，详情请参见发送报警消息至钉钉群。
4	您可以勾选多个目标任务后，在该区域批量启动、终止、下线目标任务，或为目标任务设置监控报警。

查看实时计算任务详情

在任务列表单击目标任务实例名称，即可进入任务的详情页面。您可以在任务的详情页面查看任务的详细信息、运行日志、操作日志以及任务的运行代码，方便您快速定位对目标任务执行的相关操作记录及操作内容，及时获取任务运行出错的报错详情并跟进处理。

详细信息	运行日志	操作日志	代码																				
基本信息 <table><tbody><tr><td>名称: [模糊]</td><td>任务ID: [模糊]</td></tr><tr><td>任务类型: EMR_SPARK_STREAMING</td><td>责任人: [模糊] </td></tr><tr><td>所属工作空间: [模糊]</td><td>最新版本: 1</td></tr><tr><td>发布时间: 2021-05-21 16:47:25</td><td>调度资源组:</td></tr><tr><td>Region: 上海</td><td>引擎类型: EMR</td></tr><tr><td>引擎实例: [模糊]</td><td></td></tr></tbody></table> 运行信息 <table><tbody><tr><td>实例ID: [模糊]</td><td>状态: 已失败</td></tr><tr><td>执行版本: 1</td><td>开始运行: -</td></tr><tr><td>开始等待资源: -</td><td>等待资源耗时: -</td></tr><tr><td>结束时间: -</td><td>运行时长: 0s</td></tr></tbody></table>				名称: [模糊]	任务ID: [模糊]	任务类型: EMR_SPARK_STREAMING	责任人: [模糊] 	所属工作空间: [模糊]	最新版本: 1	发布时间: 2021-05-21 16:47:25	调度资源组:	Region: 上海	引擎类型: EMR	引擎实例: [模糊]		实例ID: [模糊]	状态: 已失败	执行版本: 1	开始运行: -	开始等待资源: -	等待资源耗时: -	结束时间: -	运行时长: 0s
名称: [模糊]	任务ID: [模糊]																						
任务类型: EMR_SPARK_STREAMING	责任人: [模糊] 																						
所属工作空间: [模糊]	最新版本: 1																						
发布时间: 2021-05-21 16:47:25	调度资源组:																						
Region: 上海	引擎类型: EMR																						
引擎实例: [模糊]																							
实例ID: [模糊]	状态: 已失败																						
执行版本: 1	开始运行: -																						
开始等待资源: -	等待资源耗时: -																						
结束时间: -	运行时长: 0s																						

1.3.2. 实时同步任务运行与管理

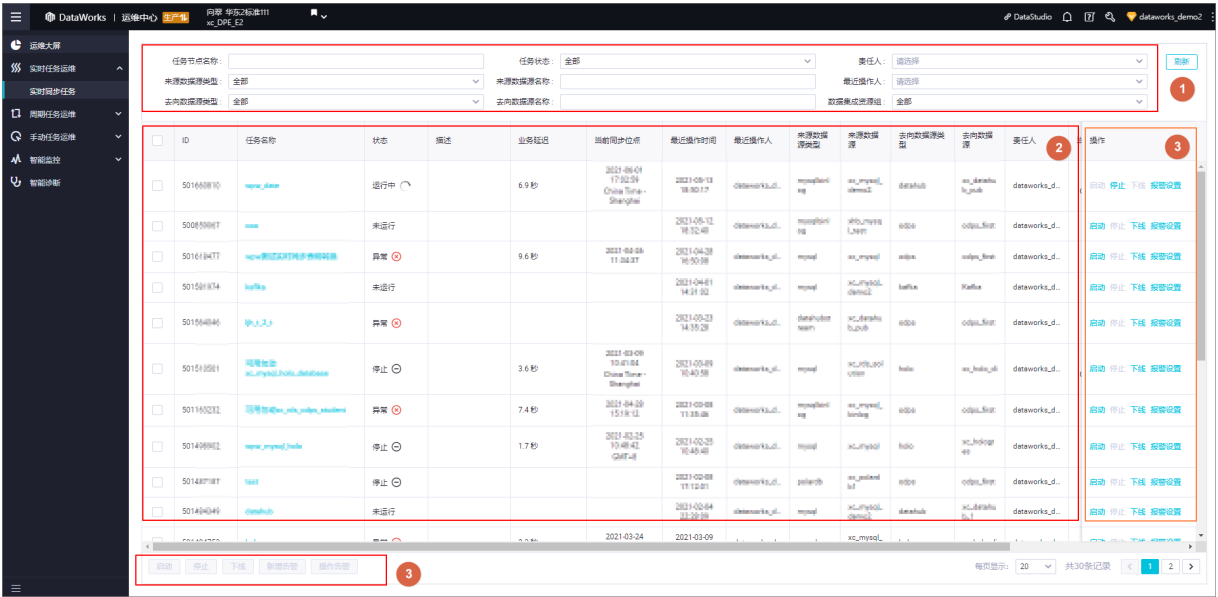
实时同步任务页面为您展示实时同步任务提交至调度系统后，经过调度系统运行后产生的实时同步任务实例。

进入实时任务运维页面

1. 登录DataWorks控制台。
2. 在左侧导航栏，单击工作空间列表。
3. 选择工作空间所在地域后，单击相应工作空间后的进入数据开发。
4. 单击左上方的图标，选择全部产品 > 任务运维 > 运维中心（工作流）。
5. 在左侧列表，单击实时任务运维后的图标。
6. 在下拉列表下单击实时同步任务，即可查看所有实时同步任务实例信息。

查看实时任务列表

实时任务实例页面以列表形式对被调度的实时同步任务进行运维和管理，包括查看实时同步任务基本信息、检查运行信息和基本属性配置以及启动和下线实例等功能。



区域	描述
1	在该区域，您可以通过筛选条件过滤出需要查询的实例。 您可以根据任务节点名称、任务状态、责任人、来源数据源类型、来源数据源名称、最近操作人、去向数据源类型、去向数据源名称、数据集资源组等条件进行精确筛选。 说明 任务名搜索的结果会受到其它筛选条件的影响，只有同时满足所有筛选条件的结果才可以展示。

区域	描述
2	在该区域，您可以查看任务实例信息，包括： ● ID：为您展示任务实例的ID。 ● 任务名称：为您展示任务实例名称。单击任务名称，即可进入任务实例详情页面。 ● 运行状态：包括运行中、未运行、等待资源、异常、停止五种状态。 ● 描述：对实例的简单描述。 ● 业务延迟：当前时间距离同步位点的时间间隔。 ● 当前同步位点：当前消费到增量数据的时间。 ● 最近操作时间：最后操作实例的时间点。 ● 最近操作人：最后对实例操作的 ● 来源数据源类型：来源数据源的类型。 ● 来源数据源：数据同步的来源数据源。 ● 去向数据源类型：去向数据源的类型。 ● 去向数据源：数据同步的去向数据源。 ● 责任人：实时任务所在的工作空间责任人。 ● 当前起始位点：目前开始消费增量数据的时间点。 ● 数据读取速度（BPS）：数据同步读取数据的速度。 ● 记录速度：写入日志的数据同步读取数据速度。 ● 数据集成资源组：任务实例运行使用的数据集成资源组。
3	在该区域，您可以对所选择的任务实例进行如下操作。 ● 启动：启动未运行的同步任务。 ● 停止：停止运行中的同步任务。 ● 下线：下线未运行、异常及停止的同步任务。 ● 新增告警：勾选对应同步任务后，单击新增告警，即可进行同步任务设置告警规则。 ● 操作告警：勾选对应同步任务后，单击操作告警，即可对同步任务告警进行删除、启用和停止操作，同时支持通过选择告警指标类型来批量修改所有对应的告警规则。

实时任务详情页面

在任务列表单击实例名称，即可进入任务详情页面。在页面上方可以切换至不同子页面，任务详情页面可以查看任务的运行信息、日志、基本属性、任务配置、Failover记录和DDL记录。

运行信息 日志 基本属性 任务配置 Failover记录 DDL记录

任务状态: 运行中

数据读取

步骤类型	线程号	总行数	总字节数	总等待时间	速度 (条记录/秒)	速度 (MB/秒)	业务延时	脏数据条数
MySQL1	0	1	8B	17.24 day	0	0B	0.26 s	0

数据写入

步骤类型	线程号	总行数	总字节数	总等待时间	速度 (条记录/秒)	速度 (MB/秒)	业务延时	脏数据条数
DataHub1	0	1	8B	3 ms	0	0B	0.26 s	0

数据库事件状态

总览: 插入: 1 次, 更新: 0 次, 删除: 0 次, DDL: 0 次

事件关联表

来源表名称:

来源表名称	插入	更新	删除
xc_demo.wipw_date	1	0	0

同步任务相关操作

启动同步任务

在操作区域单击启动，即可进入任务的启动页面。设置好相关参数，单击确定，即可启动运行该任务。

启动

是否重置位点: ☐ 重置位点

启动时间点位:

时区:

Failover: 分钟内, Failover次数超过 任务自动结束

* 脏数据策略: ☒ 零容忍, 不允许 ☐ 不限制 ☐ 有限控制

确定

取消

当勾选重置位点选项时，启动时间点位和时区为必填参数。

停止同步任务

单击停止，在弹出的对话框中选择单击停止，即可停止运行中的任务。

下线同步任务

单击下线，在弹出的对话框中选择单击下线，即可下线当前任务。

报警设置

- 在报警事件页面，设置发生时间范围、报警级别和规则查看报警事件。
- 在报警规则页面，显示当前同步任务实例已创建的报警规则。您可以在报警规则列表查看已创建规则的指标和状态等信息，也可以在操作区域对规则进行编辑、暂停、删除和模拟测试等操作。如果需要新

创建规则，请单击右侧的新建规则。

新建规则

* 名称:

描述:

* 指标:

任务状态

* 阈值:

WARNING 在 5 分钟内, 无心跳

CRITICAL 在 5 分钟内, 无心跳

* 报警间隔:

5 分钟内只发一次报警

WARNING:

☐ 邮件 ☐ 短信 ☐ 电话 ☐ 钉钉

CRITICAL:

☐ 邮件 ☐ 短信 ☐ 电话 ☐ 钉钉

* 接收人 (非钉钉):

请选择

确定

取消

参数	描述	是否必选
名称	填写新建的规则名称。	是
描述	对新建的规则进行简单描述，便于快速排查定位。	否
指标	设置描述规则的维度指标类型，指标类型和阈值需配合设置。指标类型包括：任务状态、业务延迟、Failover、脏数据和DDL不支持。	是
阈值	触发报警的阈值需要和指标类型配合设置。 - 当指标类型为任务状态时，设置触发报警规则的最短时间间隔。 - 当指标类型为业务延迟时，设置业务延迟大于等于特定时间后的报警持续时间。 - 当指标类型为Failover时，设置特定时间段内Failover超过多少次才能触发报警。 - 当指标类型为脏数据时，设置特定时间段内脏数据超过多少条才能触发报警。 - 当指标类型为DDL不支持时，仅需选择报警级别。	是
报警间隔	设置两次报警之间最小时间间隔，默认为5分钟，最短时间间隔可设置1分钟。	是

参数	描述	是否必选
WARNING	设置WARNING级别报警通知方式，包括邮件、短信、电话、钉钉。  说明 - 以邮件方式接收告警：若子账号接收告警需要由主账号在用户信息中为子账号完善个人邮箱信息 - 以短信方式接收告警：若子账号接收告警需要由主账号在用户信息中为子账号完善个人手机号信息。 - 以电话方式接收告警：若子账号接收告警需要由主账号在用户信息中为子账号完善个人手机号信息。 - 以自定义钉钉机器人接收告警：要求是普通群，且自定义机器人需要添加关键字 "DataWorks"。	否
CRITICAL	设置CRITICAL级别报警通知方式，包括邮件、短信、电话、钉钉。	否
接收人（非钉钉）	设置发送报警的联系人。	是

1.4. 周期任务运维

1.4.1. 查看周期任务

周期任务是指调度系统按照调度配置自动定时运行的任务。您可以通过列表及DAG图的方式，快速查看任务的详细信息并执行相关操作。

背景信息

周期任务页面默认展示当前登录账号下的业务流程。

 **注意** 请勿操作project name_root节点，该节点为工作空间根节点。周期任务的实例均依赖于该节点，如果冻结该节点，周期任务实例将无法运行。

实例生成方式包括T+1次日生成和发布后即时生成：

- T+1次日生成：

- 23:30之前提交发布的任务，第二天实例生效。
- 23:30之后提交发布的任务，第三天实例生效。
- **发布后即时生成**：23:30之后发布的实时转实例功能不生效。详情请参见[实例生成方式：发布后即时生成实例](#)。

如果您无法找到任务实例，则请参见[找不到任务实例如何处理](#)处理。

周期任务页面通过列表和DAG图的方式，为您展示已提交的周期任务的详细内容。

- 周期任务列表

您可以通过条件筛选，展示该条件下目标任务的基本信息，并对该任务进行测试、补数据等相关操作，详情请参见[查看周期任务列表](#)。

- DAG图

DAG图为您直观的呈现了目标周期任务的上下游节点任务概况，支持查看节点详情、血缘关系，测试节点任务等功能。同时，为您提供了节点聚合、上游分析、下游分析功能，方便您按照工作空间、责任人、优先级等维度，对图中的节点进行聚合，查看目标维度节点的总数量。您还可以对指定节点进行上下游分析，快速定位阻塞当前节点运行的上游节点，并查看当前节点的下游节点数量，掌握任务的运行情况。详情请参见[查看周期任务DAG图](#)。

使用限制

DAG图的功能限制如下：

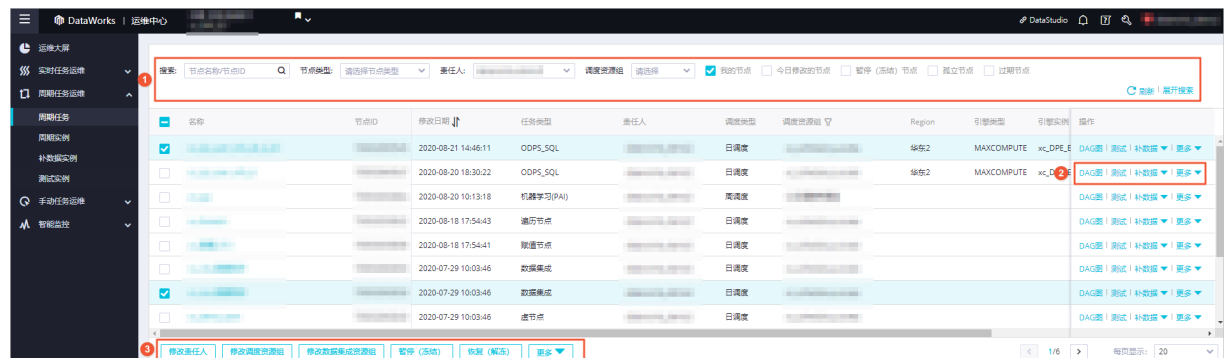
- 仅DataWorks专业版及以上版本用户才可以使用DAG图的节点聚合、上游分析、下游分析功能。
- DataWorks基础版和标准版用户可以免费试用节点聚合、上游分析、下游分析功能至2021年5月31日，到期后，必须升级DataWorks至专业版才可以继续使用该功能。升级DataWorks版本，详情请参见[DataWorks增值版本](#)。
- 当前仅华南1（深圳）地域支持使用DAG图的节点聚合、上游分析、下游分析功能。

进入周期任务页面

1. 进入[数据开发](#)页面。
 - i. 登录[DataWorks控制台](#)。
 - ii. 在左侧导航栏，单击[工作空间列表](#)。
 - iii. 选择工作空间所在地域后，单击相应工作空间后的[进入数据开发](#)。
2. 单击左上方的图标，选择全部产品 > 任务运维 > 运维中心（工作流）。
3. 在左侧导航栏，单击[周期任务运维](#) > [周期任务](#)。进入周期任务页面，查看周期任务列表或DAG图。

查看周期任务列表

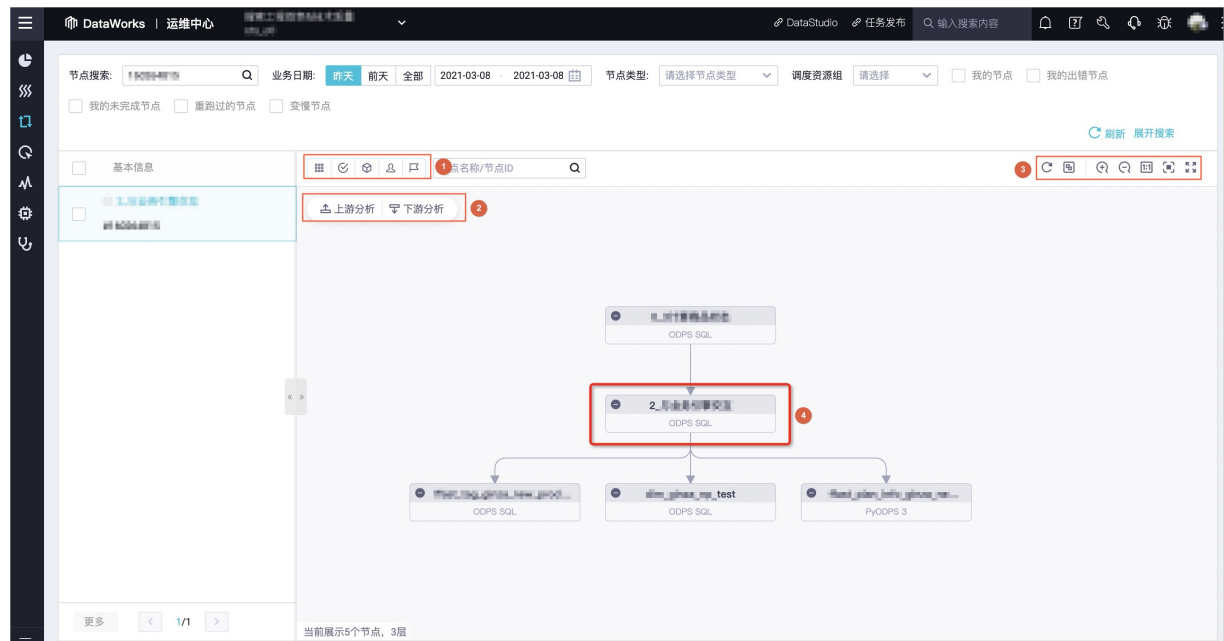
单击页面中间的图标，展开周期任务列表。



操作	描述
筛选	在上图的模块①区域，您可以通过筛选条件过滤需要查询的任务。 您可以根据节点名称、节点ID、解决方案、业务流程、节点类型、责任人、Region、调度资源组、引擎类型、引擎实例、基线、我的节点、今日修改的节点、暂停（冻结）节点、孤立节点和过期节点等条件进行精确筛选。  说明 任务名搜索的结果会受到其它筛选条件的影响，只有同时满足所有筛选条件的结果才可以展示。
DAG图	单击操作栏中的DAG图，即可打开该节点的DAG图。您可以在DAG图中查看节点的属性、操作日志、代码等信息。
测试	单击操作栏中的测试，即可对当前节点进行测试，详情请参见 测试实例 。
补数据	单击操作栏中的补数据，即可对当前节点进行补数据，详情请参见 执行补数据并管理补数据实例 。
更多	单击操作栏中的更多，可以进行暂停、恢复、查看实例等更多操作： - 单击暂停（冻结），即可将当前节点置为暂停（冻结）状态。暂停后，正常生成实例，但当前任务生成的实例及下游依赖实例均不调度。 - 单击恢复（解冻），即可将当前节点置为恢复（解冻）状态。恢复后，当前任务生成的实例及下游依赖实例正常调度。 - 单击查看实例，即可查看此节点的周期实例。 - 单击添加报警，即可为节点配置报警。 - 单击修改责任人，即可修改节点责任人。 - 单击添加到基线，即可将当前节点添加到基线。 - 如果工作空间存在多个调度资源组，请单击修改调度资源组。在修改调度资源组对话框中，选择资源组，单击确定。 - 单击配置质量监控，即可配置数据质量，对数据进行校验。 - 单击查看血缘，即可查看节点的血缘关系图。 - 单击上下游，即可跳转至节点基本信息页面，查看节点的上游列表和下游列表。
批量操作	在上图的模块③区域，您可以批量选择任务，进行修改责任人、修改调度资源组、修改数据集资源组、暂停（冻结）、恢复（解冻）、添加报警、添加到基线和下线节点等操作。

查看周期任务DAG图

单击相应任务后的DAG图，打开该任务的DAG图。您可以在DAG图进行如下操作：



- 节点聚合

当周期任务的节点数量或层级过多时，您可以按照工作空间、责任人、优先级等维度进行节点任务聚合，查看指定目标维度的节点数量。方便您了解各个维度节点任务的数量，合理分配并执行节点任务。以下示例为您展示，周期任务节点不聚合及按优先级聚合的显示情况。

- 周期任务节点不聚合时，显示如下图所示。



- 周期任务节点按照优先级聚合时，显示如下图所示。您可以快速了解到当前周期任务的子节点共有6个优先级为1的节点任务。



- 上游节点分析

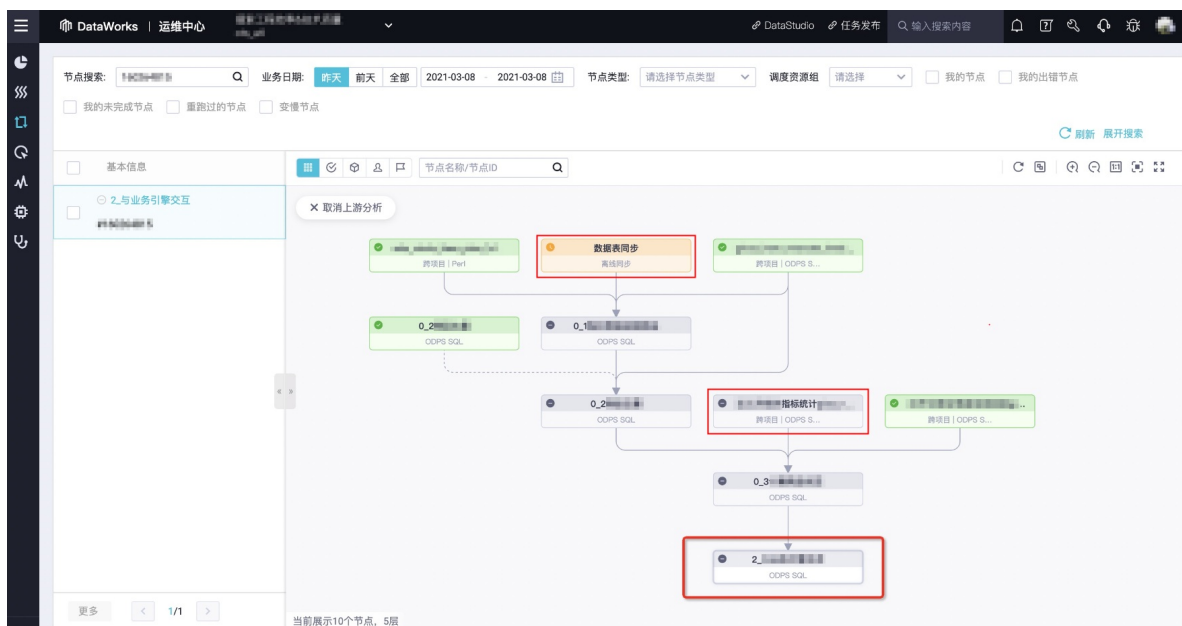
周期任务通常存在上下游依赖关系，如果目标节点长时间未运行，则您可以对其进行上游节点分析，在DAG图中查看阻塞当前节点运行的上游节点，快速定位并及时处理相关运行报错，提升任务的执行效率。

说明 仅支持对未运行的节点执行上游分析操作。

示例2_节点任务长时间未运行，则您可以选中该节点，单击左上角的上游分析



分析结果得出，导致阻塞当前节点运行的上游节点任务为数据表同步及指标统计，如下图所示。通过该分析结果，您可以快速定位阻塞节点，并及时处理相关报错。



下游节点分析

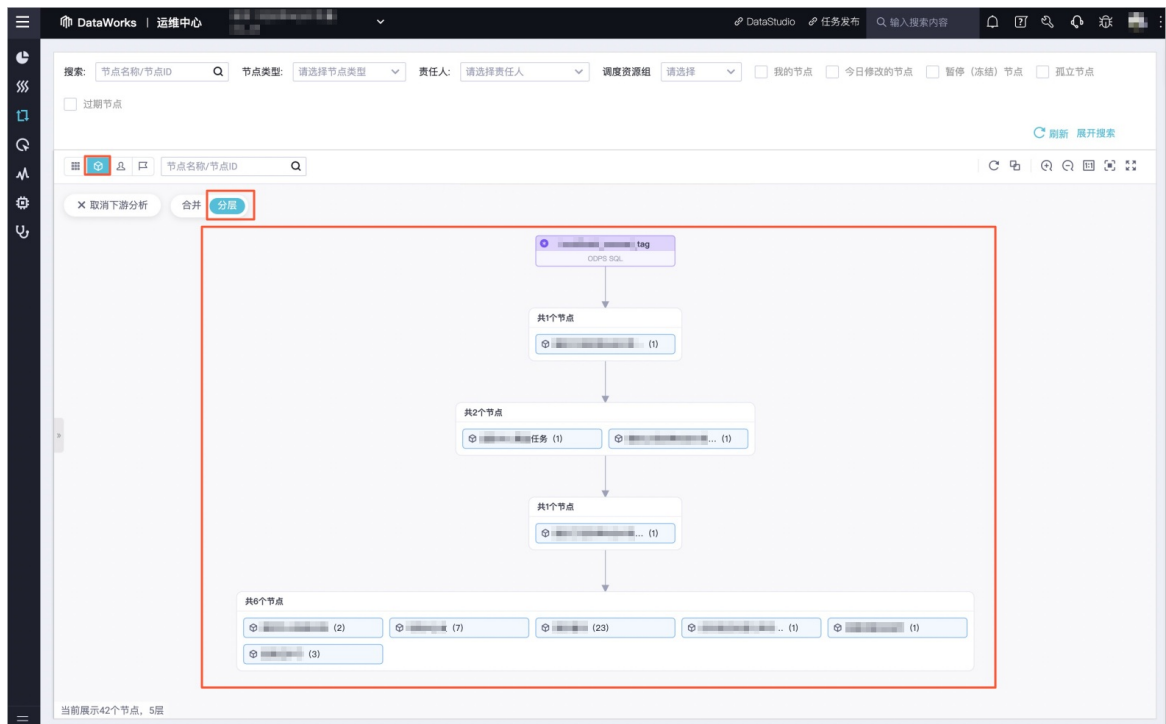
当周期任务的节点数量或层级过多时，您可以进行下游节点分析，通过工作空间、责任人、优先级等维度进行筛选，统计不同层级目标维度的节点数量，或统计所有层级目标维度的节点总数量。

说明

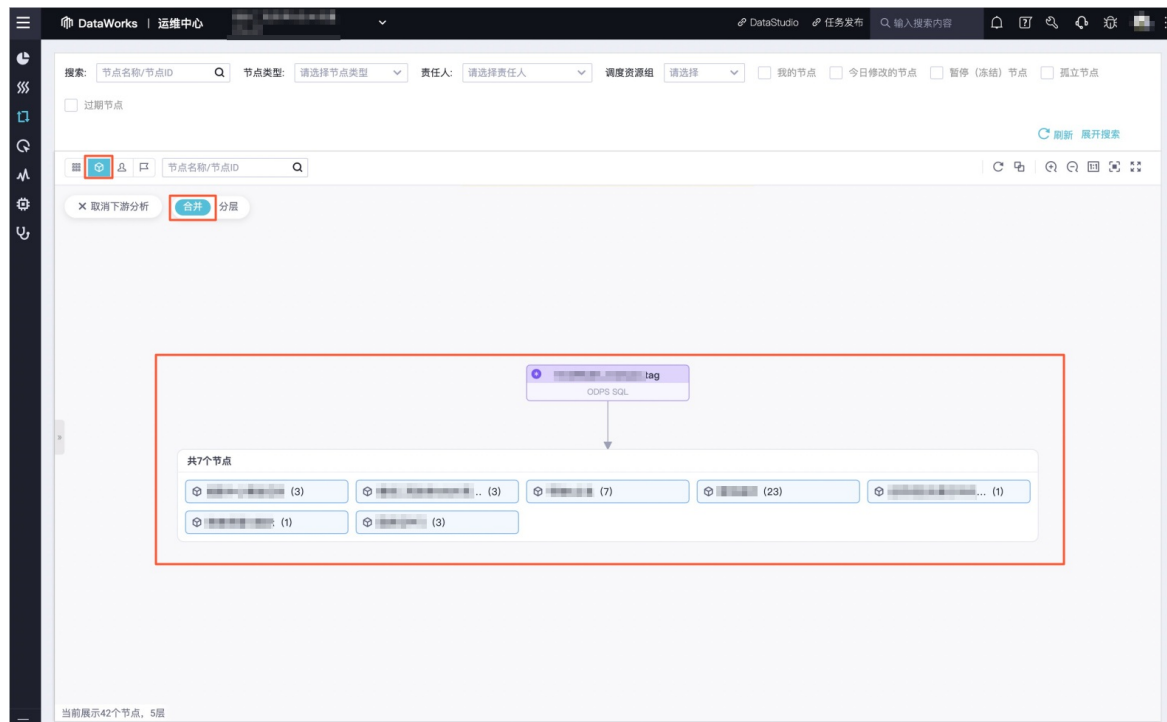
- 下游节点默认按照责任人聚合，统计所有层级该维度的节点总数量。
- 下游分析使用分层展示下游节点分析结果时，最多展示6层。如果您希望展示更多层级，请单击左上角的继续分析。

示例对tag节点进行下游分析，结果如下：

- 使用分层方式，并按照所属工作空间聚合，在不同层级展示不同工作空间的节点数量。



- 使用合并方式，并按照所属工作空间聚合，将所有下游节点合并至同一层，展示所属不同工作空间的节点数量。

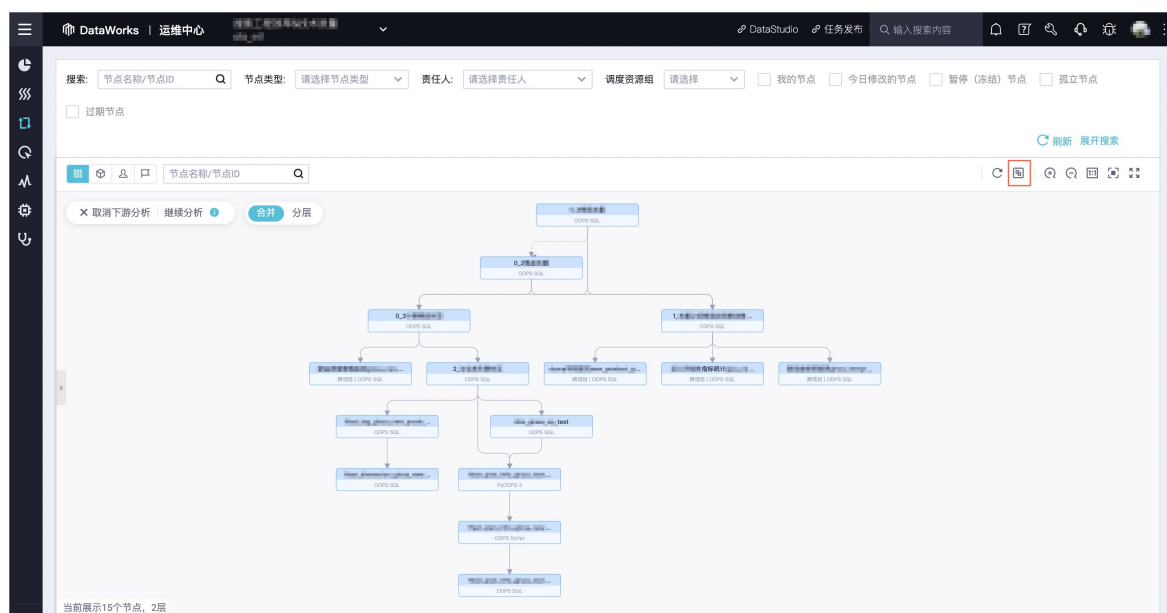


- 选择DAG图的显示样式。

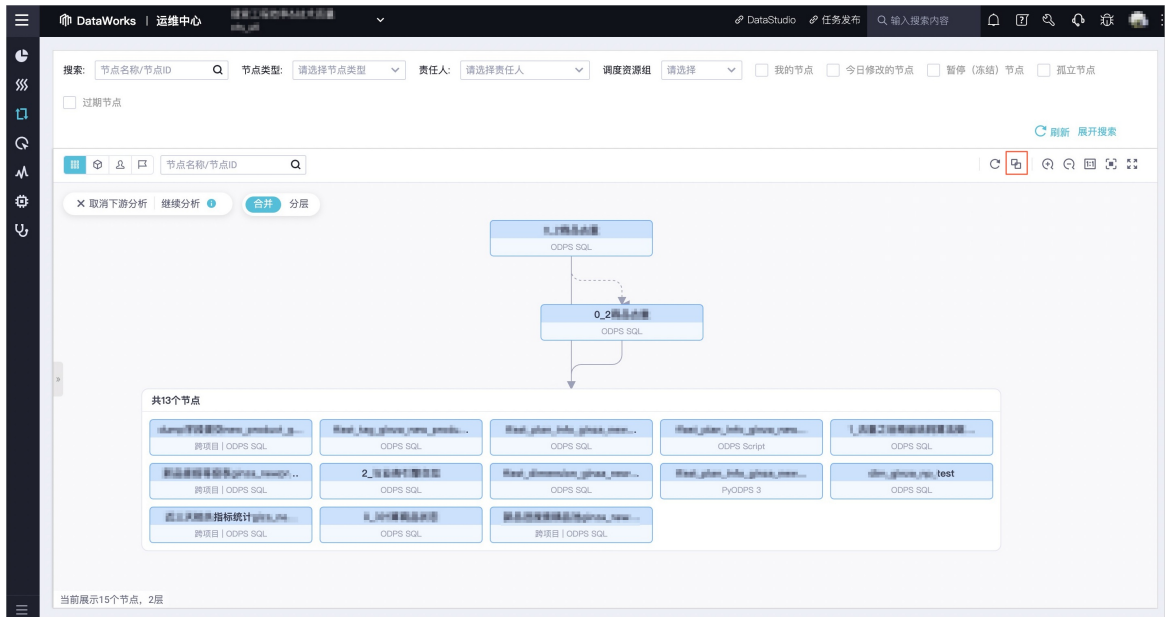
您可以根据实际需要，单击DAG图所在界面右上角的功能图标，调整DAG图的显示样式。例如，全屏展示、适配屏幕展示等。

示例对0_2节点的下游节点，展示拆分组及成组的DAG图。

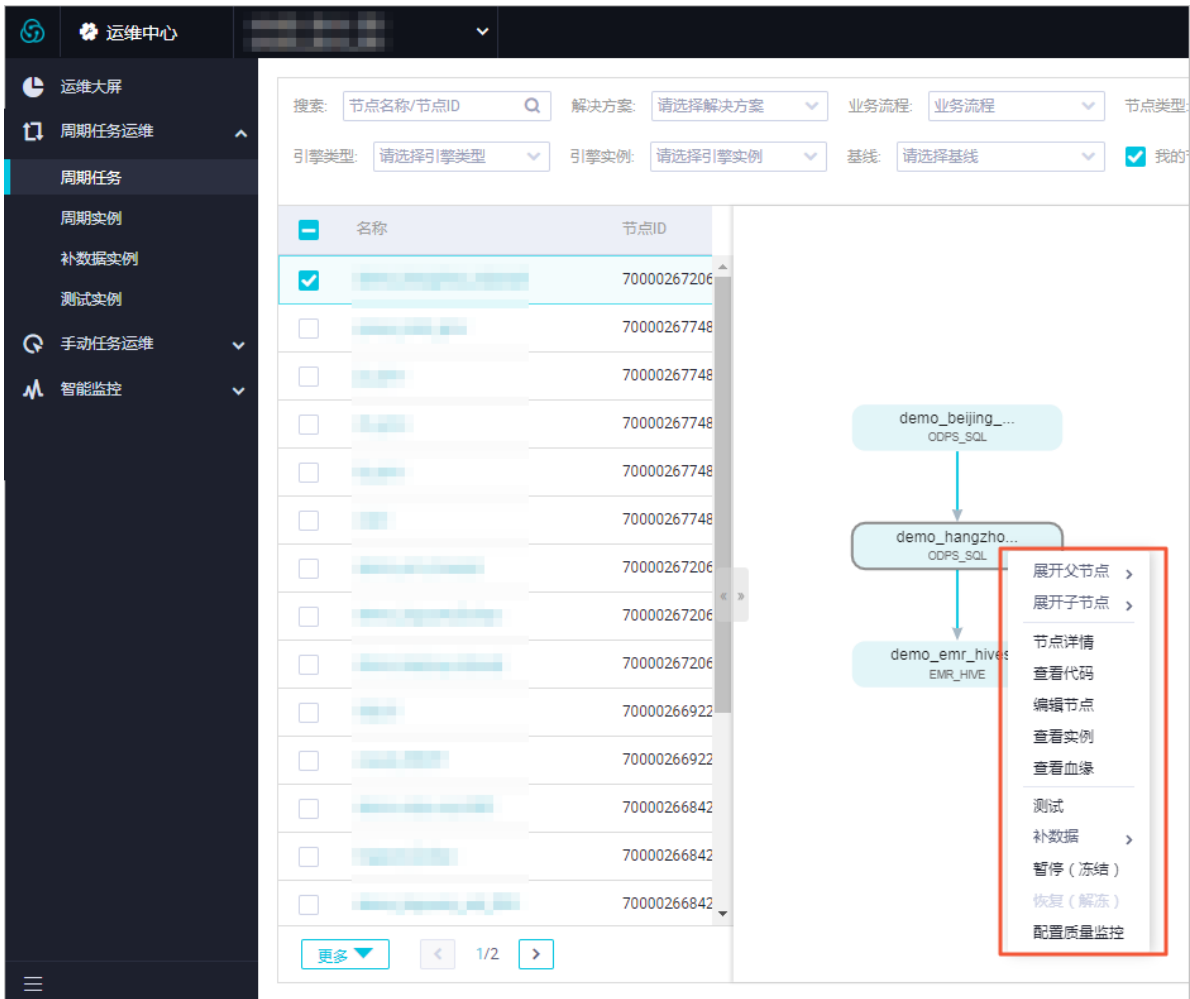
- 拆分组的DAG图如下所示。通过该方式，您可以直观的了解所有节点的上下游关系。



- 成组的DAG图如下所示。通过该方式，以每行呈现5个节点任务，有序的为您展示下游节点，方便您快速获取下游节点的总数量。



- 右键单击目标节点，执行相关操作。



操作	描述
展开父节点	当一个业务流程有3个及以上节点时，运维中心展示任务时会自动隐藏节点。您可以通过展开父子层级查看更多的节点依赖关系。层级越大，展示越全面。
展开子节点	
节点详情	单击后，即可跳转至节点基本信息页面，查看当前节点的输入表、输出表、上游列表和下游列表等信息。
查看代码	查看当前节点的代码。
编辑节点	单击后，即可跳转至数据开发页面，对当前节点的内容进行修改。
查看实例	查看当前节点的周期实例。
查看血缘	查看当前节点的血缘关系图。
测试	单击后，您需要在冒烟测试对话框中，输入冒烟测试名称并选择业务日期，单击确定，进入测试实例页面。
补数据	包括当前节点、当前节点及下游节点和海量节点模式。
暂停（冻结）	将当前节点置为暂停（冻结）状态，并停止调度。
恢复（解冻）	恢复暂停（冻结）的节点的调度。
配置质量监控	配置当前节点的数据质量，对数据进行校验。

1.4.2. 查看周期实例

周期实例是周期任务达到启用调度所配置的周期性运行时间时，被自动调度的实例快照。您可以通过列表及DAG图的方式，快速查看实例的详细信息并执行相关操作。

周期任务每调度一次，便生成一个实例工作流。您可以对已调度的实例任务进行日常的运维管理，例如查看运行状态，对任务进行终止、重跑和解冻等操作。

② 说明

- 周期任务定时生成周期实例，实例会按最新的代码运行任务。如果您的任务在实例生成后修改了代码并重新提交发布，则未运行的实例会拉取最新的代码运行任务。
- 如果任务失败未报警，请首先检查是否已在[个人信息](#)页面配置了您的手机号码与邮箱地址。
- 您需要购买DataWorks专业版及以上版本，周期实例即可支持运行诊断功能，详情请参见[运行诊断](#)。

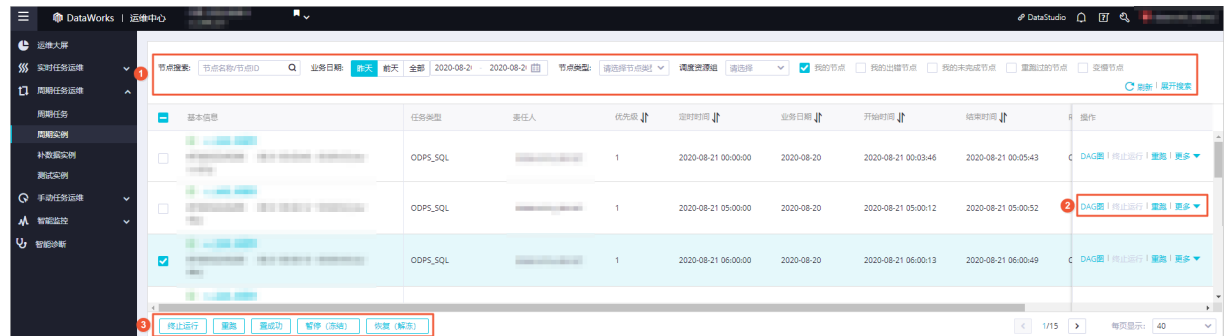
使用限制

- 目前仅周期实例支持在任务失败时产生报警。
- 如果周期实例和补数据实例均启动运行，为了保证周期实例的正常运行，您需要终止补数据实例的运行。
- DAG图的功能限制如下：
 - 仅DataWorks专业版及以上版本用户才可以使用DAG图的节点聚合、上游分析、下游分析功能。

- DataWorks基础版和标准版用户可以免费试用节点聚合、上游分析、下游分析功能至2021年5月31日，到期后，必须升级DataWorks至专业版才可以继续使用该功能。升级DataWorks版本，详情请参见DataWorks增值版本。
- 当前仅华南1（深圳）地域支持使用DAG图的节点聚合、上游分析、下游分析功能。

周期实例列表

周期实例列表以列表形式对被调度的任务进行运维和管理，包括检查运行日志、重跑任务、终止正在运行的任务等。

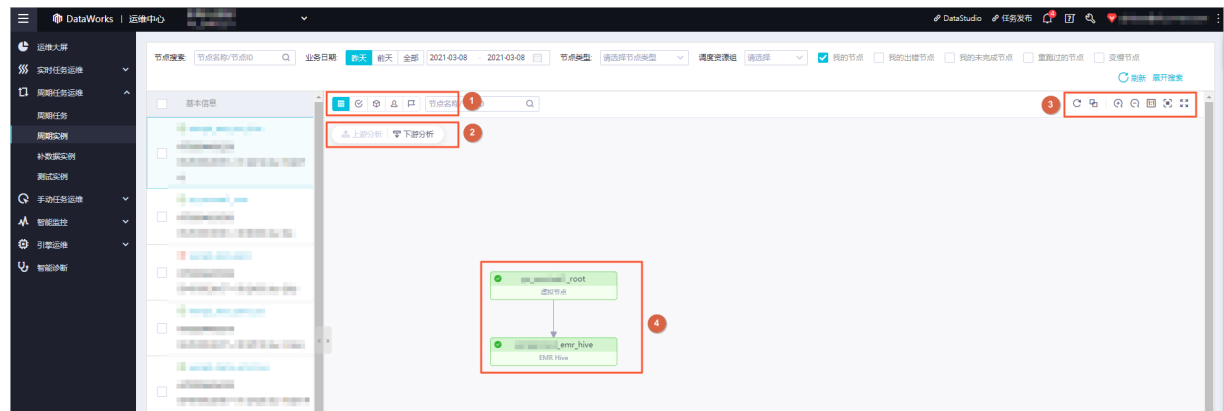


操作	描述
筛选	在上图的模块1区域，您可以通过筛选条件过滤出需要查询的实例。 您可以根据节点名称、节点ID、业务日期、节点类型、运行日期、解决方案、业务流程、Region、引擎类型、引擎实例、基线、责任人、调度类型、运行状态、我的节点、我的出错节点和我的未完成节点等条件进行精确筛选。 说明 默认筛选的业务日期是当前时间的前一天。
终止运行	仅支持对状态为等待时间、等待资源和运行中的实例进行终止运行操作，进行该操作后，该实例的状态被置为失败。
重跑	可以重跑某任务，任务执行成功后可以触发下游未运行状态任务的调度。常用于处理出错节点和漏跑节点。 说明 仅支持重跑成功和失败状态的任务。
重跑下游	可以重跑某任务及其下游任务，需要您自定义勾选，勾选的任务将被重跑，任务执行成功后可以触发下游未运行状态任务的调度。常用于处理数据修复。 说明 仅支持勾选未运行、成功和失败状态的任务，如果勾选了其它状态的任务，页面会提示已选节点中包含不符合运行条件的节点，并禁止提交运行。
置成功	修改当前节点的状态为成功，并运行下游未运行状态的任务。常用于处理出错节点。 说明 仅支持置成功状态为失败的任务，不能置成功业务流程。

操作	描述
暂停（冻结）	周期实例中的冻结只针对当前实例，且正在运行中的实例。
恢复（解冻）	可以将冻结状态的实例解冻。 - 如果该实例还未运行，则上游任务运行完毕后，会自动运行。 - 如果上游任务都运行完毕，则该任务会直接被置为失败，需要手动重跑后，实例才会正常运行。
批量操作	在上图的模块3区域，您可以批量选择任务，进行终止运行、重跑、置成功、冻结和解冻等操作。

实例DAG图

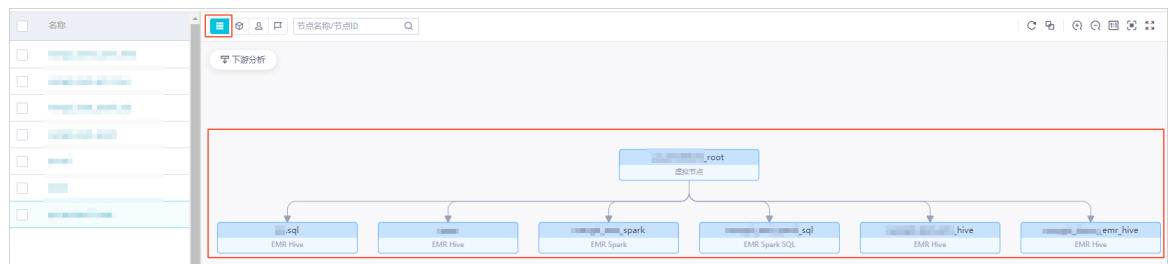
单击相应任务后的DAG图，打开该任务的DAG图。您可以在DAG图进行如下操作：



- 节点聚合

当周期实例任务的节点数量或层级过多时，您可以按照任务状态、工作空间、责任人、优先级等维度进行节点任务聚合，查看指定目标维度的节点数量。方便您了解各个维度节点任务的数量，合理分配并执行节点任务。以下示例为您展示，周期任务节点不聚合及按优先级聚合的显示情况。

- 周期任务节点不聚合时，显示如下图所示。



- 周期任务节点按照优先级聚合时，显示如下图所示。您可以快速了解到当前周期任务的子节点共有6个优先级为7的节点任务。

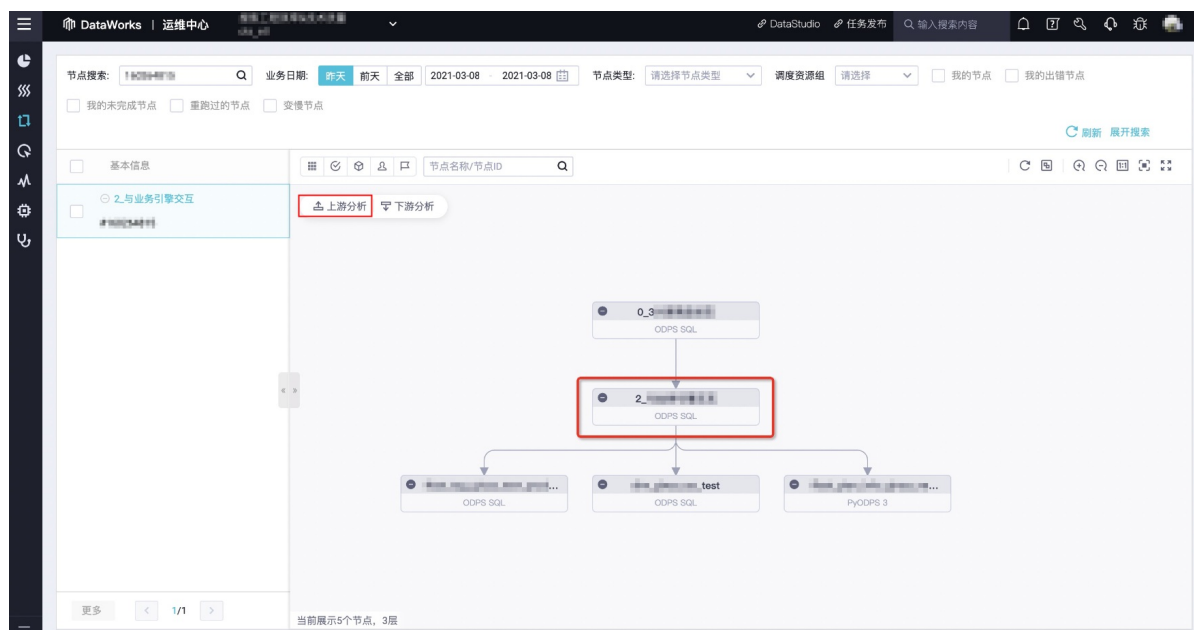


上游节点分析

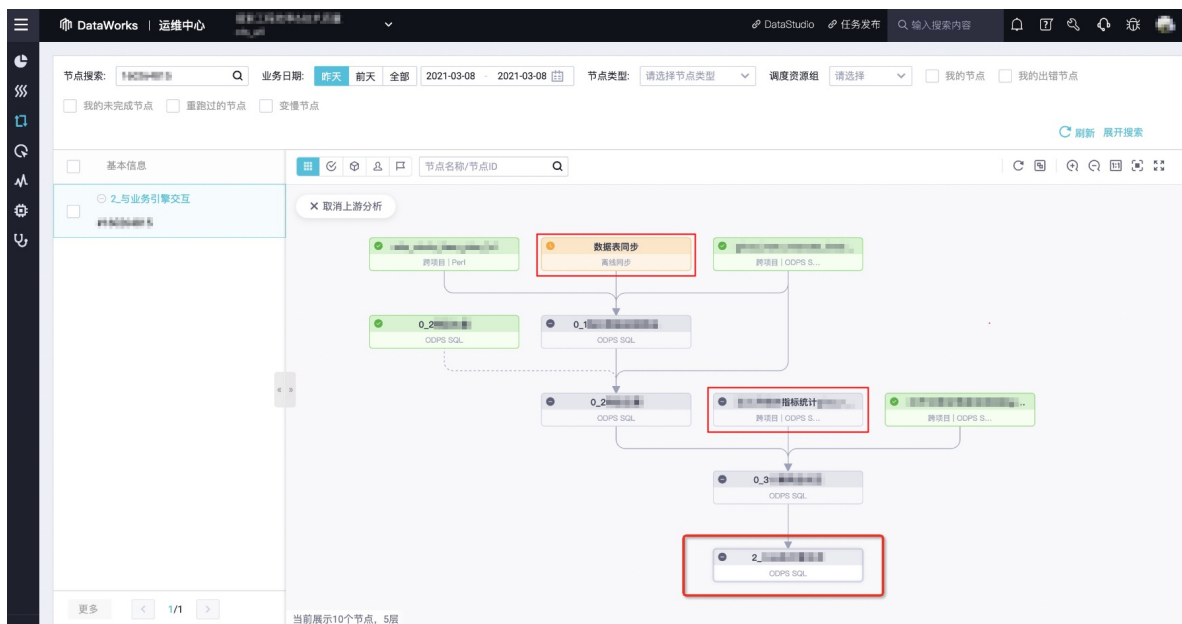
周期任务通常存在上下游依赖关系，如果目标节点长时间未运行，则您可以对其进行上游节点分析，在DAG图中查看阻塞当前节点运行的上游节点，快速定位并及时处理相关运行报错，提升任务的执行效率。

说明 仅支持对未运行的节点执行上游分析操作。

示例2_节点任务长时间未运行，则您可以选中该节点，单击左上角的上游分析



分析结果得出，导致阻塞当前节点运行的上游节点任务为数据表同步及指标统计，如下图所示。通过该分析结果，您可以快速定位阻塞节点，并及时处理相关报错。



- 下游节点分析

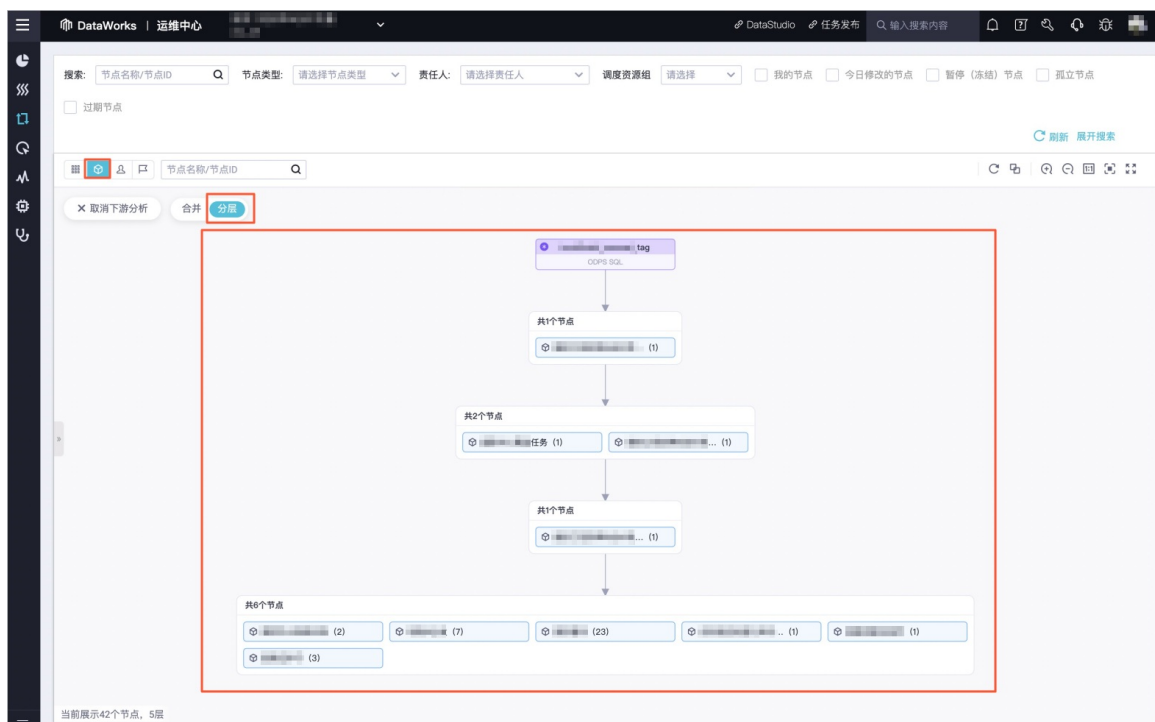
当周期任务的节点数量或层级过多时，您可以进行下游节点分析，通过任务状态、工作空间、责任人、优先级等维度进行筛选，统计不同层级目标维度的节点数量，或统计所有层级目标维度的节点总数量。

② 说明

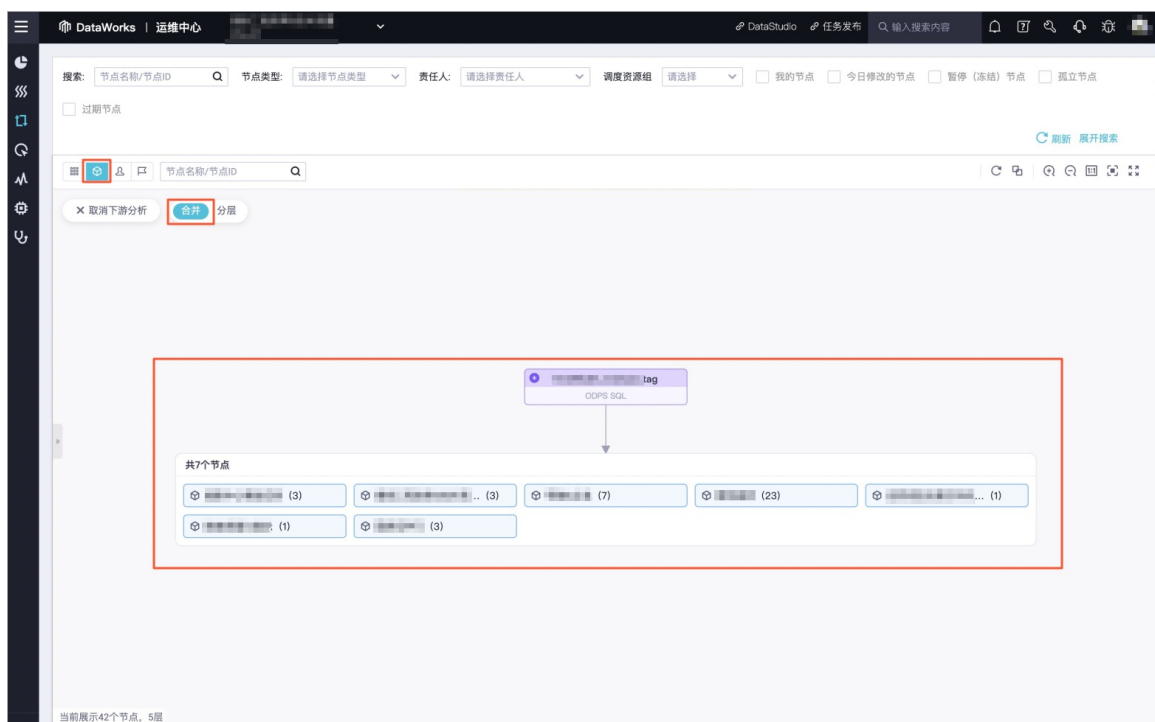
- 下游节点默认按照责任人聚合，统计所有层级该维度的节点总数量。
- 下游分析使用分层展示下游节点分析结果时，最多展示6层。如果您希望展示更多层级，请单击左上角的[继续分析](#)。

示例对tag节点进行下游分析，结果如下：

- 使用分层方式，并按照所属工作空间聚合，在不同层级展示不同工作空间的节点数量。



- 使用合并方式，并按照所属工作空间聚合，将所有下游节点合并至同一层，展示所属不同工作空间的节点数量。

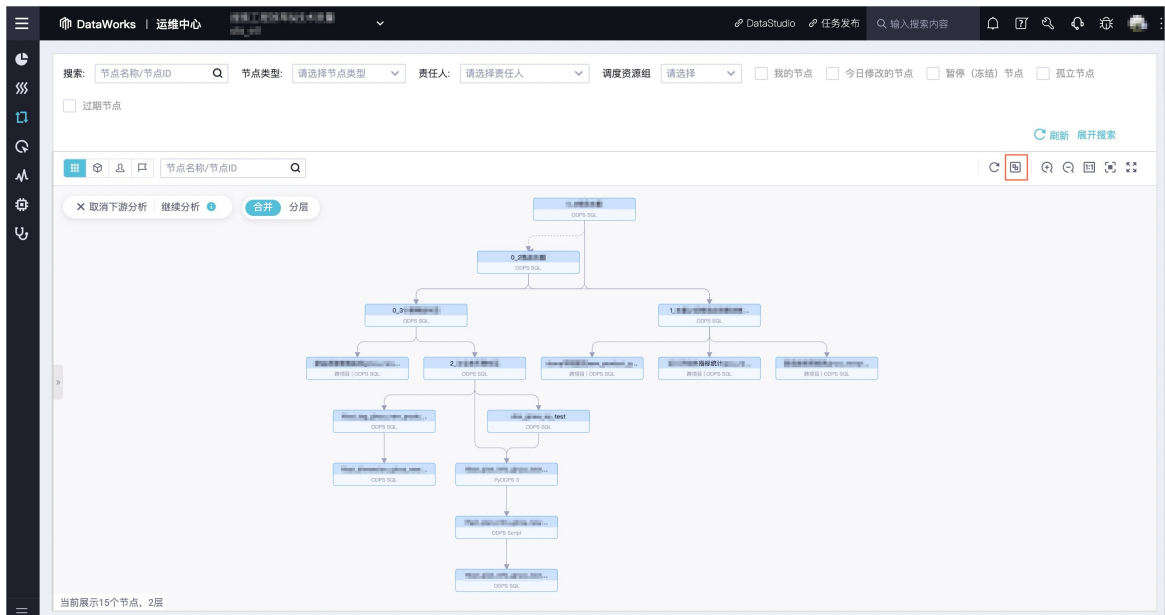


- 选择DAG图的显示样式。

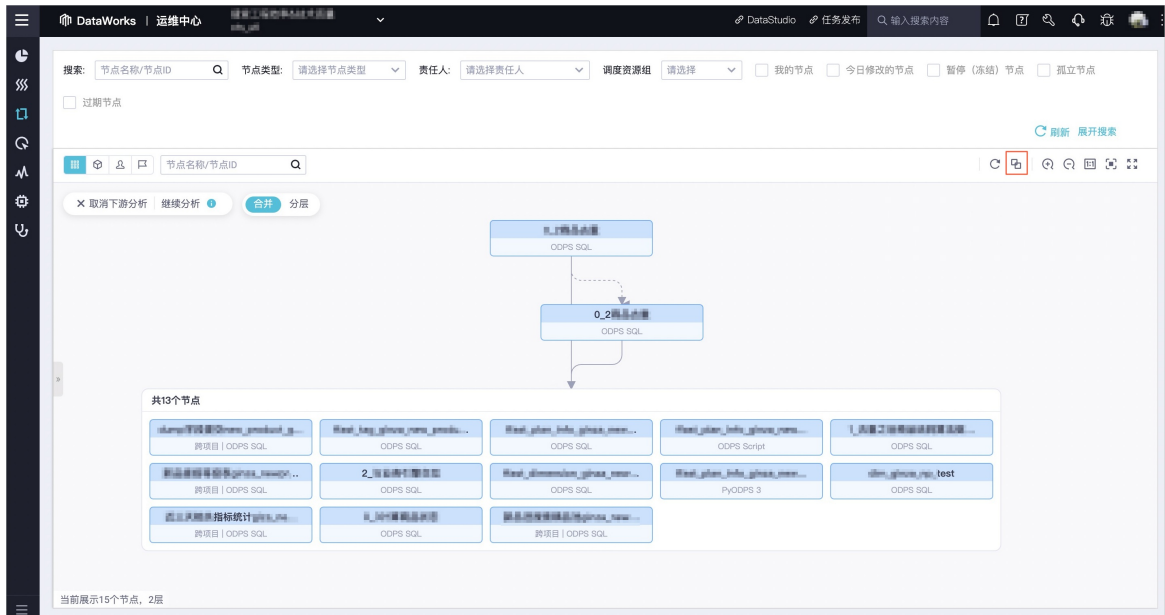
您可以根据实际需要，单击DAG图所在界面右上角的功能图标，调整DAG图的显示样式。例如，全屏展示、适配屏幕展示等。

示例对0_2节点的下游节点，展示拆分组及成组的DAG图。

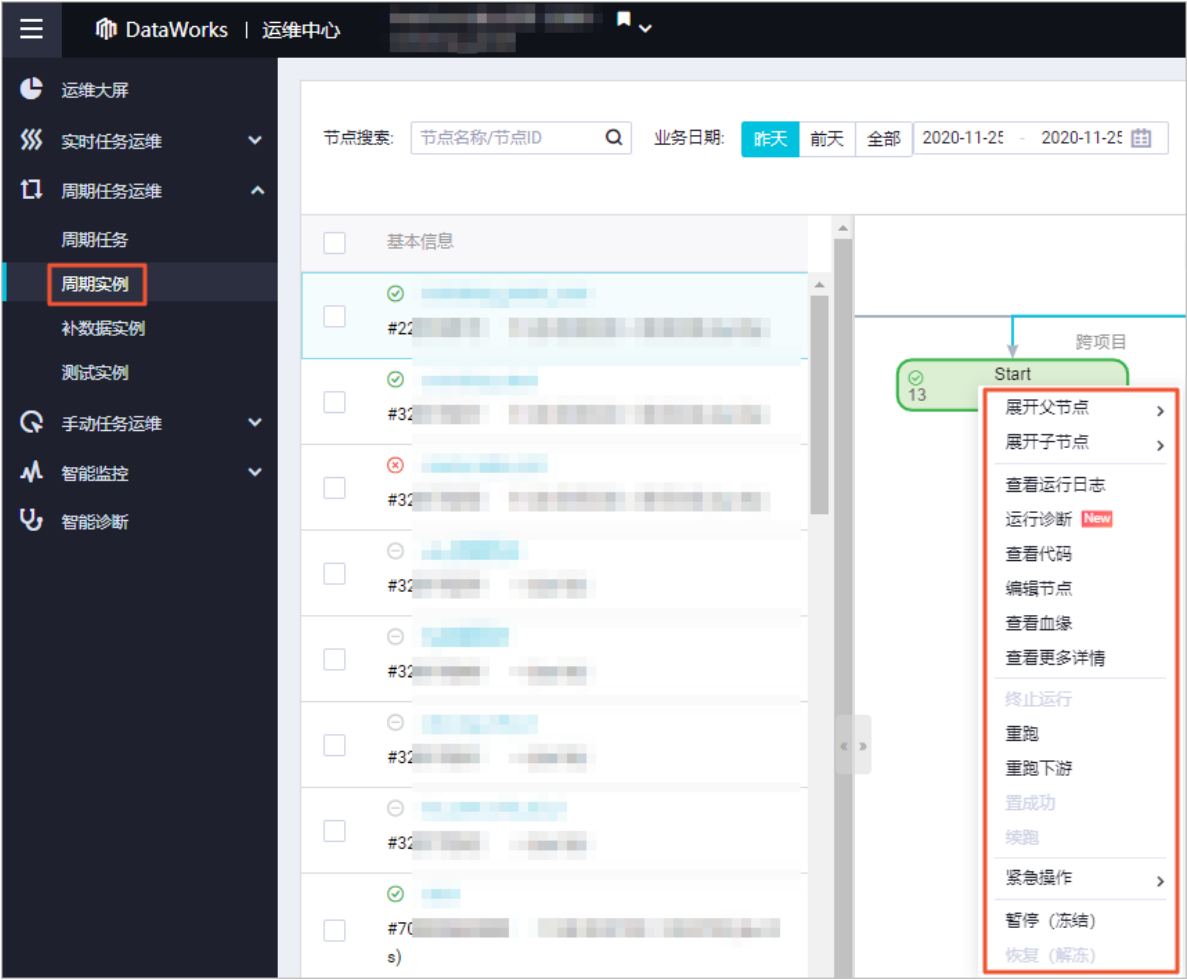
- 拆分组 的DAG图如下所示。通过该方式，您可以直观的了解到所有节点的上下游关系。



- 成组 的DAG图如下所示。通过该方式，以每行呈现5个节点任务，有序的为您展示下游节点，方便您快速获取下游节点的总数量。



- 右键单击目标节点，执行相关操作。



操作	描述
展开父节点/子节点	当一个业务流程有3个及以上节点时，运维中心展示任务时会自动隐藏节点。您可以通过展开父子层级查看全部节点的内容。
查看运行日志	查看当前实例正在运行、成功、失败等状态的运行日志。
运行诊断	DataWorks运维中心为您提供运行诊断功能，帮助您掌握任务运行的全链路信息，以迅速定位问题。详情请参见 运行诊断 。
查看代码	查看当前实例的代码。
编辑节点	单击后，即可跳转至数据开发页面，对当前节点的内容进行修改。
查看血缘	查看当前实例的血缘关系。
查看更多详情	单击后，即可查看当前节点的属性、上下文、运行日志、操作日志和代码等信息。
终止运行	仅等待运行、运行中状态的实例可以进行终止运行的操作。进行此操作后，该实例将为失败状态。

操作	描述
重跑	可以重跑某任务，任务执行成功后可以触发下游未运行状态任务的调度。常用于处理出错节点和漏跑节点。  说明 仅支持重跑未运行、成功和失败状态的任务。
重跑下游	可以重跑某任务及其下游任务，需要您自定义勾选，勾选的任务将被重跑，任务执行成功后可以触发下游未运行状态任务的调度。常用于处理数据修复。  说明 仅支持勾选未运行、完成和失败状态的任务，如果勾选了其它状态的任务，页面会提示已选节点中包含不符合运行条件的节点，并禁止提交运行。
置成功	将当前实例的状态改为成功，并运行下游未运行状态的任务。常用于处理出错节点。  说明 仅支持将状态为失败的任务置成功，不能置成功业务流程。
续跑	任务执行失败后，可以续跑此任务。
紧急操作	当前实例在非常紧急的情况下的操作，紧急操作只对当前节点本次有效。 选择去除依赖，即可解除当前节点的依赖关系。常用于上游失败并与此实例没有数据关系时，启动此节点。
暂停（冻结）	周期实例中的冻结仅针对当前实例，且正在运行中的实例。
恢复（解冻）	可以将冻结状态的实例解冻。 - 如果该实例还未运行，则上游任务运行完毕后，会自动运行。 - 如果上游任务都运行完毕，则该任务会直接被置为失败。需手动重跑后，方会正常运行。

1.4.3. 执行补数据并管理补数据实例

DataWorks的补数据功能用于在指定日期范围内运行周期任务，支持补当前节点及其下游节点，您可以在补数据实例页面查看实例的运行状态，以及对补数据实例进行终止、重跑和解冻等操作。本文为您介绍如何对周期任务进行补数据，并管理补数据实例。

背景信息

周期任务开发完成并提交发布后，任务会按照调度配置定时运行。如果您希望在指定时间段运行周期任务，可以使用补数据功能。DataWorks周期任务的补数据操作，支持补当前节点及其下游节点，详情请参见[补数据](#)，其支持的补数据模式如下：

- **当前节点**：用于对当前节点进行补数据操作。
- **当前节点及下游节点**：用于对当前节点及下游节点进行批量补数据操作，通常在当前节点的下游节点数量较少时使用该功能，支持指定部分下游节点。
- **海量节点模式**：用于对当前节点及下游节点进行批量补数据操作，通常在当前节点的下游节点数量较多时使用该功能，支持按照项目筛选下游节点。同时，支持节点白名单和黑名单，选中或过滤目标节点。

- **高级模式**：用于灵活选择一批节点进行批量补数据操作，节点之间可以不存在依赖关系。您可以在DAG图上使用可视化方式添加节点，也可以在任务列表中添加节点。
 - 您可以在当前周期任务的DAG图中，结合DAG的聚合功能，按照工作空间、所属责任人或优先级等维度将节点划分为组，直接勾选该节点组，快速添加目标类别的节点为补数据节点。实例的DAG图，详情请参见[实例DAG图](#)。
 - 您可以在周期任务界面，通过目标条件进行筛选过滤，在左侧任务列表中，勾选符合条件的节点为需要补数据的节点。

使用限制

- 仅华南1（深圳）、中东部1（迪拜）地域支持周期任务使用高级模式进行补数据。
- DataWorks支持批量终止补数据实例，但暂不支持批量删除补数据实例。补数据实例在过期后（30天左右）将自动删除。

注意事项

- 当补一个区间的数据时，在启动补数据的第一天存在一个运行失败的实例，则当天的补数据实例会被置为失败状态，第二天的实例也不会启动运行（只有当天的全部任务实例都成功，第二天的任务实例才会开始运行）。
- 自依赖的周期任务执行补数据时，如果补数据任务中第一个实例前一天的周期实例没有运行，则该补数据任务也无法触发运行。如果补数据任务的第一个实例前一天没有周期实例，则补数据实例会直接触发运行。
- 如果周期实例和补数据实例均启动运行，为了保证周期实例的正常运行，您需要终止补数据实例的运行。
- 补数据实例过多或并行数过高可能导致周期调度的资源紧张，请您根据需求合理配置。

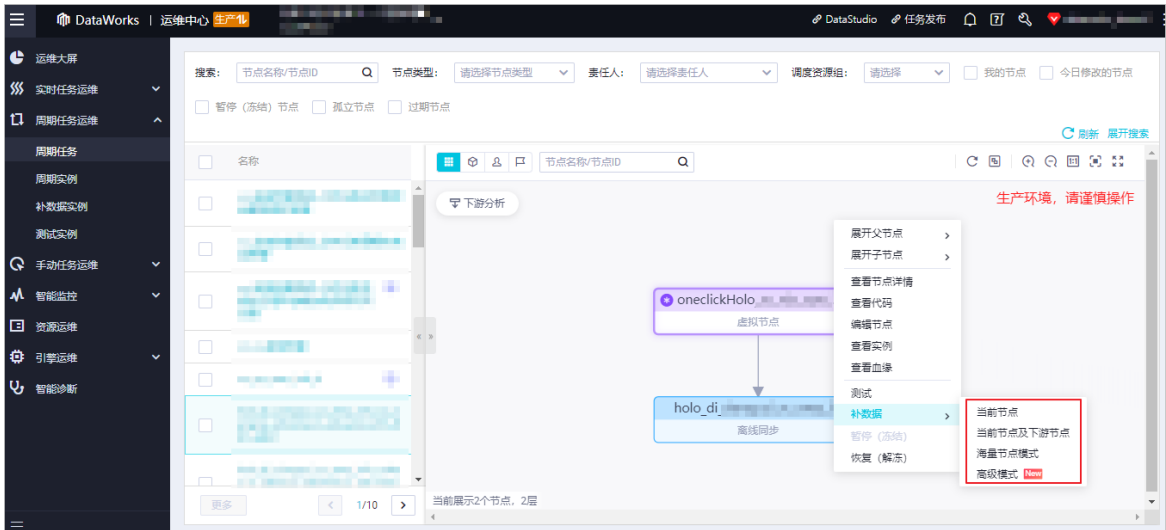
补数据

1. 进入数据开发页面。
 - i. 登录[DataWorks控制台](#)。
 - ii. 在左侧导航栏，单击工作空间列表。
 - iii. 选择工作空间所在地域后，单击相应工作空间后的进入数据开发。
2. 单击左上方的图标，选择全部产品 > 任务运维 > 运维中心（工作流）。
3. 在左侧导航栏，单击周期任务运维 > 周期任务，进入周期任务页面。
4. 目标节点补数据。
 - i. 单击周期任务列表中的目标任务，打开该任务的DAG图。

您也可以单击图标展开周期任务列表，单击目标任务操作列的DAG图，打开该任务的DAG图。

- ii. 右键单击目标节点，鼠标悬停至补数据，选择相应的补数据模式，配置补数据的相关参数。

 **说明** 您也可以在周期任务页面，单击图标展开周期任务列表，单击目标任务操作列的补数据，选择相应的补数据模式，对该周期任务进行补数据。



补数据的模式及参数配置如下：

- 使用当前节点模式补数据。

补数据

⚠ 补数据实例过多或并行数过高可能导致周期调度的资源紧张。

补数据名称：

P_oneclickHolo_...

当前任务：

oneclickHolo_...

业务日期：

2021-08-23 - 2021-08-23

删除

添加

是否并行：

☐

?

顺序：

☒ 业务日期正序

☐ 业务日期倒序

确定

取消

参数配置如下。

参数	描述
补数据名称	系统会自动生成补数据名称，您可以根据需要进行修改。
当前任务	当前选中的任务。

> 文档版本：20211230

46

参数	描述
业务日期	选择执行补数据任务的业务日期，业务日期精确到天。 - 如果您需要对该任务的多个不连续的时间段进行补数据，则可以单击添加，配置多个时间段。 - 当配置的业务日期晚于当前日期，则您可以选择立即运行定时时间大于此刻的补数据实例，后续当实际时间大于配置的业务日期时间，系统会立即运行该补数据实例。 例如，当前日期为 2021-08-24，业务日期为 2021-09-17，则等实际日期为 2021-09-18（立即大于业务日期）时，该补数据实例会立即启动运行。  说明 建议补数据的时间不要太长，以免出现任务需要等待资源的情况。
是否并行	您可以通过选择是否并行，控制同时生成多少个补数据实例来进行补数据，具体如下： - 不并行：按照业务日期顺序串行执行，前一个补数据实例运行结束后下个补数据实例才会运行。 - 并行：根据您配置的并行数，按照业务日期将补数据实例拆分成若干组并行执行。同一时刻有多个不同业务日期的补数据实例运行。
并行数	用于配置当前补数据任务生成的补数据实例个数，当进行补数据时，所有补数据实例并行执行。  说明 当是否并行参数选择并行时，需要配置当前参数。 并行数取值最小为2组，最大为10组。多个实例并行执行的情况如下： - 如果业务日期的跨度时间少于配置的并行组数，则并行执行。例如业务日期是1月11日~1月13日，并行数配置为4组，则只会生成三个补数据实例（每个补数据实例对应一个业务日期），三个实例同时并发执行。 - 如果业务日期的跨度大于选择的并行组数，则系统会结合配置的业务日期顺序兼有串行和并行。例如业务日期是1月11日~1月13日，并行数配置为2组，则会生成两个补数据实例（其中一个补数据实例会有两个业务日期，这两个业务日期对应的任务串行执行），两个补数据实例并行执行。
顺序	您可以选择按照业务日期正序或业务日期倒序进行补数据。

- 使用当前节点及下游节点模式补数据。

补数据

补数据实例过多或并行数过高可能导致周期调度的资源紧张。

补数据名称:

P_oneclickHolo_

业务日期:

2021-08-23 - 2021-08-23

删除

添加

是否并行:

☐

顺序:

☒ 业务日期正序

☐ 业务日期倒序

选择需要补数据的节点

搜索:

按名称进行搜索...

Q

层级:

请选择

☐	任务名称	任务类型	责任人	工作空间
☐		虚拟节点		
☐		离线同步		
☐		离线同步		

确定

取消

参数配置如下。

参数	描述
补数据名称	系统会自动生成补数据名称，您可以根据需要进行修改。
业务日期	选择执行补数据任务的业务日期，业务日期精确到天。 ■ 如果您需要对该任务的多个不连续的时间段进行补数据，则可以单击添加，配置多个时间段。 ■ 当配置的业务日期晚于当前日期，则您可以选择立即运行定时时间大于此刻的补数据实例，后续当实际时间大于配置的业务日期时间，系统会立即运行该补数据实例。 例如，当前日期为 2021-08-24，业务日期为 2021-09-17，则等实际日期为 2021-09-18（立即大于业务日期）时，该补数据实例会立即启动运行。 ? 说明 建议补数据的时间不要太长，以免出现任务需要等待资源的情况。

参数	描述
是否并行	您可以通过选择是否并行，控制同时生成多少个补数据实例来进行补数据，具体如下： ■ 不并行：按照业务日期顺序串行执行，前一个补数据实例运行结束后下个补数据实例才会运行。 ■ 并行：根据您配置的并行数，按照业务日期将补数据实例拆分成若干组并行执行。同一时刻有多个不同业务日期的补数据实例运行。
并行数	并行数取值最小为2组，最大为10组。多个实例并行执行的情况如下： ■ 如果业务日期的跨度时间少于配置的并行组数，则并行执行。例如业务日期是1月11日~1月13日，并行数配置为4组，则只会生成三个补数据实例（每个补数据实例对应一个业务日期），三个实例同时并发执行。 ■ 如果业务日期的跨度大于选择的并行组数，则系统会结合配置的业务日期顺序兼有串行和并行。例如业务日期是1月11日~1月13日，并行数配置为2组，则会生成两个补数据实例（其中一个补数据实例会有两个业务日期，这两个业务日期对应的任务串行执行），两个补数据实例并行执行。
顺序	您可以选择按照业务日期正序或业务日期倒序进行补数据。
选择需要补数据的节点	您可以根据任务名称、层级等条件进行筛选过滤，选择需要补数据的目标节点。  说明 ■ 通过任务名称进行搜索时支持模糊匹配，即输入关键字，即可显示所有名称中包含关键字的节点。 ■ 选择需要补数据的节点区域为您展示的是当前节点及其包含的所有下游节点（包含下游节点的子节点），您可以选择当前节点及其下游节点的部分或全部节点任务（包含当前节点）进行补数据。

- 使用海量节点模式补数据。

补数据

补数据实例过多或并行数过高可能导致周期调度的资源紧张。

补数据名称:

P_oneclickHolo

业务日期:

2021-08-23 - 2021-08-23

删除

添加

顺序:

业务日期正序

业务日期倒序

按工作空间选择需要补数据的节点

☒ 包含当前节点 oneclickHolo

您的所有项目

项目筛选

Date

Date

dem

flink

Flink

flink

shar

195 项

>

<

补数据的项目

项目筛选

暂无数据

0 项

节点白名单

节点搜索

节点黑名单

节点搜索

确定

取消

参数配置如下。

参数	描述
补数据名称	系统会自动生成补数据名称，您可以根据需要进行修改。
业务日期	选择执行补数据任务的业务日期，业务日期精确到天。 如果您需要对该任务的多个不连续的时间段进行补数据，则可以单击添加，配置多个时间段。 当配置的业务日期晚于当前日期，则您可以选择立即运行定时时间大于此刻的补数据实例，后续当实际时间大于配置的业务日期时间，系统会立即运行该补数据实例。 例如，当前日期为 2021-08-24，业务日期为 2021-09-17，则等实际日期为 2021-09-18（立即大于业务日期）时，该补数据实例会立即启动运行。 说明 建议补数据的时间不要太长，以免出现任务需要等待资源的情况。

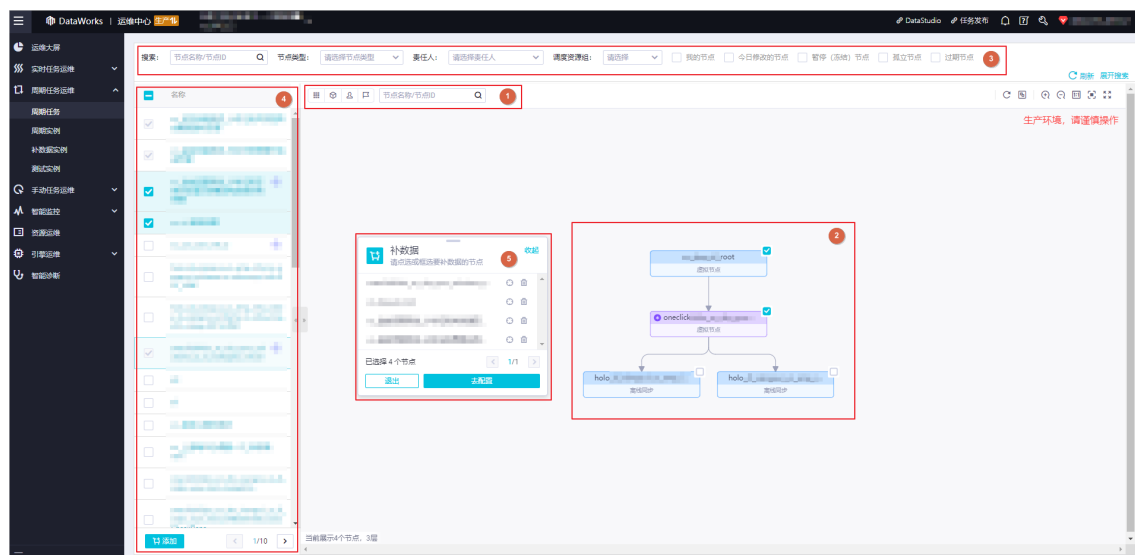
> 文档版本：20211230

50

参数	描述
顺序	您可以选择按照业务日期正序或业务日期倒序进行补数据。
按工作空间选择需要补数据的节点	您可以根据业务需求，在您的所有项目区域，选择目标工作空间，将其添加至补数据的项目区域，为目标工作空间中所选节点进行补数据。  说明 ■ 筛选DataWorks工作空间支持模糊匹配，即输入关键字，即可显示所有名称中包含关键字的工作空间。 ■ 您只能选择当前所在地域中的DataWorks工作空间进行补数据操作。 ■ 您可以为该补数据任务配置节点白名单或黑名单，选中或过滤无需执行补数据的节点。 ■ 您可以选择该补数据任务是否包含当前节点，取值如下： ■ 如果勾选包含当前节点，则补的是当前节点及下游节点的数据。 ■ 如果未勾选包含当前节点，则当前节点会空跑，补的是当前节点下游节点的数据。
节点白名单	除选中的DataWorks工作空间中包含的节点外，仍需要进行补数据的节点。  说明 目前仅支持通过节点ID搜索目标节点。
节点黑名单	选中的DataWorks工作空间中不需要进行补数据的节点。  说明 目前仅支持通过节点ID搜索目标节点。

- 使用高级模式补数据。

在高级模式中，您可以使用DAG图的聚合功能及节点的类型、责任人等过滤条件，批量为节点之间不存在依赖关系的多个节点进行补数据。



高级模式补数据步骤如下：

a. 选择需要补数据的节点。

- 在当前周期任务的DAG图中，您可以结合DAG图的聚合功能（区域1），按照不聚合、按所属工作空间聚合、按责任人聚合或按优先级聚合等维度将节点划分为组，直接勾选该节点组（区域2），快速添加目标类别的节点为补数据节点。DAG图的聚合功能，详情请参见[实例DAG图](#)。
- 您也可以在周期任务界面，通过节点名称、节点类型、责任人、调度资源组等筛选条件进行过滤（区域3），在左侧任务列表中，勾选符合条件的周期任务为需要补数据的节点（区域4），单击下方的添加，将目标周期任务添加至补数据任务中。

说明 该方式是将整个目标周期任务添加至补数据任务中，因此系统会对该周期任务中的所有节点进行补数据。如果您希望对周期任务中的部分节点进行补数据，则可以单击目标周期任务的名称，进入DAG图界面，选择需要进行补数据的节点。

b. 查看补数据节点。

补数据节点添加完成后，您可以在区域5的补数据面板，查看已选择的节点。同时，您还可以执行如下管理操作：

- 单击目标节点后的图标，跳转至该节点的DAG图界面，您可以重新勾选该DAG图中需要补数据的节点。
- 单击目标节点后的图标，在当前补数据任务中删除该节点。

c. 在区域5的补数据面板，单击去配置，配置补数据任务的相关参数。

补数据

补数据实例过多或并行数过高可能导致周期调度的资源紧张。

补数据名称:

P_oneclickHolo_

已选任务数:

4个

修改

业务日期:

2021-08-23 - 2021-08-23

删除

添加

是否并行:

☐

?

顺序:

☒ 业务日期正序 ☐ 业务日期倒序

确定

取消

参数配置如下。

参数	描述
补数据名称	系统会自动生成补数据名称，您可以根据需要进行修改。
已选任务数	当前补数据任务中包含的节点个数。您可以单击 修改 ，修改当前补数据任务包含的节点。
业务日期	选择执行补数据任务的业务日期，业务日期精确到天。 - 如果您需要对该任务的多个不连续的时间段进行补数据，则可以单击添加，配置多个时间段。 - 当配置的业务日期晚于当前日期，则您可以选择立即运行定时时间大于此刻的补数据实例，后续当实际时间大于配置的业务日期时间，系统会立即运行该补数据实例。 例如，当前日期为 2021-08-24，业务日期为 2021-09-17，则等实际日期为 2021-09-18（立即大于业务日期）时，该补数据实例会立即启动运行。 ? 说明 建议补数据的时间不要太长，以免出现任务需要等待资源的情况。
是否并行	您可以通过选择是否并行，控制同时生成多少个补数据实例来进行补数据，具体如下： - 不并行：按照业务日期顺序串行执行，前一个补数据实例运行结束后下个补数据实例才会运行。 - 并行：根据您配置的并行数，按照业务日期将补数据实例拆分成若干组并行执行。同一时刻有多个不同业务日期的补数据实例运行。

参数	描述
并行数	并行数取值最小为2组，最大为10组。多个实例并行执行的情况如下： - 如果业务日期的跨度时间少于配置的并行组数，则并行执行。例如业务日期是1月11日~1月13日，并行数配置为4组，则只会生成三个补数据实例（每个补数据实例对应一个业务日期），三个实例同时并发执行。 - 如果业务日期的跨度大于选择的并行组数，则系统会结合配置的业务日期顺序兼有串行和并行。例如业务日期是1月11日~1月13日，并行数配置为2组，则会生成两个补数据实例（其中一个补数据实例会有两个业务日期，这两个业务日期对应的任务串行执行），两个补数据实例并行执行。
顺序	您可以选择按照业务日期正序或业务日期倒序进行补数据。

5. 单击确定，启动补数据。

管理补数据实例

补数据任务启动后，会生成相应的补数据实例，您可以在运维中心页面，单击左侧导航栏的周期任务运维 > 补数据实例，查看补数据实例的基本信息及运行详情，并进行终止运行、重跑等管理操作。进入运维中心，详情请参见[进入运维中心](#)。

周期任务运维

补数据实例

搜索: 1

节点名称/节点ID

补数据名称: 选择补数据名称

创建人: 选择创建人

创建日期: 2021-08-24

运行状态: 选择运行状态

业务日期: 选择业务日期

节点类型: 选择节点类型

责任人: 请选择责任人

地域: 请选择region

调度资源组: 请选择

引擎实例: 请选择

基础: 请选择基础

我的节点

我发起的

重置

刷新

收起搜索

任务名称: 2

位置状态

运行状态

创建人

创建日期

节点数

业务日期

操作

+ P_...

- P_...

2021-01-01 00:00:00

名称

责任人

定时时间

开始运行时间: 2021-08-24 10:43:32

结束时间: 2021-08-24 10:44:37

运行时长: 1m4s

操作: 终止运行 | 重跑 | 更多

名称

责任人

定时时间

开始运行时间: 2021-08-24 10:44:44

结束时间: 2021-08-24 10:45:53

运行时长: 1m8s

操作: 终止运行 | 重跑 | 更多

名称

责任人

定时时间

开始运行时间: 2021-08-24 10:46:19

结束时间: 2021-08-24 10:46:36

运行时长: 16s

操作: 终止运行 | 重跑 | 更多

名称

责任人

定时时间

开始运行时间: 2021-08-24 10:45:56

结束时间: 2021-08-24 10:46:15

运行时长: 19s

操作: 终止运行 | 重跑 | 更多

名称

责任人

定时时间

开始运行时间: 2021-08-24 10:43:35

结束时间: 2021-08-24 10:44:27

运行时长: 52s

操作: 终止运行 | 重跑 | 更多

名称

责任人

定时时间

开始运行时间: 2021-08-24 10:43:29

结束时间: 2021-08-24 10:43:29

运行时长: 0s

操作: 终止运行 | 重跑 | 更多

终止运行

重置

4

每页显示: 20

1

下一页

区域	描述
----	----

区域	描述
1	在该区域，您可以通过筛选条件过滤出需要查询的实例。 您可以根据节点名称、节点ID、补数据名称、创建人、创建日期、运行状态、业务日期、我的节点、我发起的等条件进行精确筛选。  说明 - 在该区域，您需要单击展开搜索，才会显示节点类型调度资源组引擎实例等更多的筛选条件。 - 节点名称支持模糊匹配，即输入关键字，即可显示所有名称中包含关键字的节点。
2	在该区域，您可以查看补数据实例的相关信息，包括： 任务名称：为您展示补数据实例的名称。单击实例名称前的图标，即可显示该实例的运行日期，实例包含的节点及节点的运行信息（即区域3）。 检查状态：当前补数据实例的检查状态。 运行状态：包括运行中、未运行、等待资源、异常、停止五种状态。 创建人：创建补数据实例的阿里云账号。 创建日期：创建补数据实例的日期。 节点个数：补数据实例包含的节点数量。 业务日期：运行补数据实例的日期。 在该区域，您还可以对补数据实例执行如下操作： 批量终止：您可以选择批量终止补数据实例中运行状态为等待运行和运行中的实例，执行此操作后，对应实例将被置为失败状态。  说明 - DataWorks支持批量终止补数据实例，但暂不支持批量删除补数据实例。补数据实例在过期后（30天左右）将自动删除。 - DataWorks当前不支持对状态为未运行、运行成功或运行失败的实例执行终止运行操作。 批量重跑：批量重新运行补数据实例。  说明 DataWorks当前仅支持批量重跑状态为运行失败的补数据实例。 复用：您可以选择复用某次补数据操作的节点集，方便您快速选定需要进行补数据的节点。

区域	描述
3	在该区域，您可以查看补数据实例所包含节点的相关信息，包括： ● 名称：为您展示补数据实例所包含节点的名称。单击节点名称，即可进入节点的详情页面。 ● 责任人：节点任务所在的工作空间责任人。 ● 定时时间：节点任务设置的定时运行时间。 ● 开始运行时间：节点任务运行的起始时间。 ● 结束时间：节点任务运行的结束时间。 ● 运行时长：节点任务运行的时长。 在该区域，您还可以对节点任务执行如下操作： ● 终止运行：终止运行状态为等待运行和运行中的节点，执行此操作后，该节点将被置为失败状态。 ? 说明 DataWorks当前不支持对状态为未运行、运行成功或运行失败的节点执行终止运行操作。 ● 重跑：重新运行目标节点任务。 ? 说明 仅支持重跑状态为运行成功和运行失败的节点。 ● 更多 > 重跑下游：重新运行目标节点任务的下游节点。 ● 更多 > 置成功：将该节点任务的状态置为成功。 ● 更多 > 暂停（冻结）：将当前节点置为暂停（冻结）状态，并停止调度。 ● 更多 > 恢复（解冻）：将暂停（冻结）的节点恢复调度。 ● 更多 > 查看血缘：查看节点的血缘关系图。
4	您可以在区域3勾选多个目标节点任务，在本区域（区域4）单击终止运行或重跑，即可批量终止运行或重新运行所选的节点任务。

实例状态说明

序号	状态类型	状态标识
1	运行成功状态	
2	未运行状态	
3	运行失败状态	
4	正在运行状态	
5	等待状态	

序号	状态类型	状态标识
6	冻结状态	

1.4.4. 实例DAG图

本文为您介绍实例（包括自动调度的周期实例、手动触发生成的补数据实例和测试实例）DAG图查看与DAG图的相关功能。

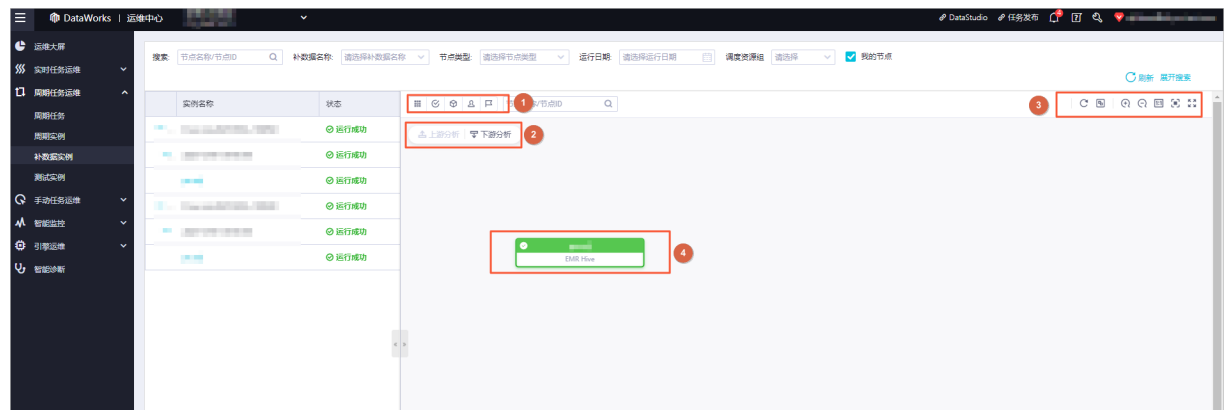
使用限制

DAG图的功能限制如下：

- 仅DataWorks专业版及以上版本用户才可以使用DAG图的节点聚合、上游分析、下游分析功能。
- DataWorks基础版和标准版用户可以免费试用节点聚合、上游分析、下游分析功能至2021年5月31日，到期后，必须升级DataWorks至专业版才可以继续使用该功能。升级DataWorks版本，详情请参见[DataWorks增值版本](#)。
- 当前仅华南1（深圳）地域支持使用DAG图的节点聚合、上游分析、下游分析功能。

实例DAG图

单击相应任务后的DAG图，打开该任务的DAG图。您可以在DAG图进行如下操作：



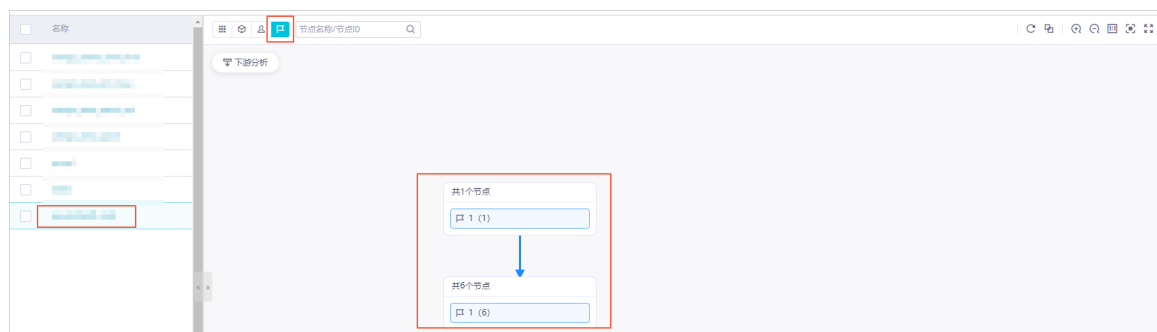
- 节点聚合

当周期实例任务的节点数量或层级过多时，您可以按照任务状态、工作空间、责任人、优先级等维度进行节点任务聚合，查看指定目标维度的节点数量。方便您了解各个维度节点任务的数量，合理分配并执行节点任务。以下示例为您展示，周期任务节点不聚合及按优先级聚合的显示情况。

- 周期任务节点不聚合时，显示如下图所示。



- 周期任务节点按照优先级聚合时，显示如下图所示。您可以快速了解到当前周期任务的子节点共有6个优先级为7的节点任务。



上游节点分析

周期任务通常存在上下游依赖关系，如果目标节点长时间未运行，则您可以对其进行上游节点分析，在DAG图中查看阻塞当前节点运行的上游节点，快速定位并及时处理相关运行报错，提升任务的执行效率。

说明 仅支持对未运行的节点执行上游分析操作。

示例2_节点任务长时间未运行，则您可以选中该节点，单击左上角的上游分析。



分析结果得出，导致阻塞当前节点运行的上游节点任务为数据表同步及指标统计，如下图所示。通过该分析结果，您可以快速定位阻塞节点，并及时处理相关报错。



下游节点分析

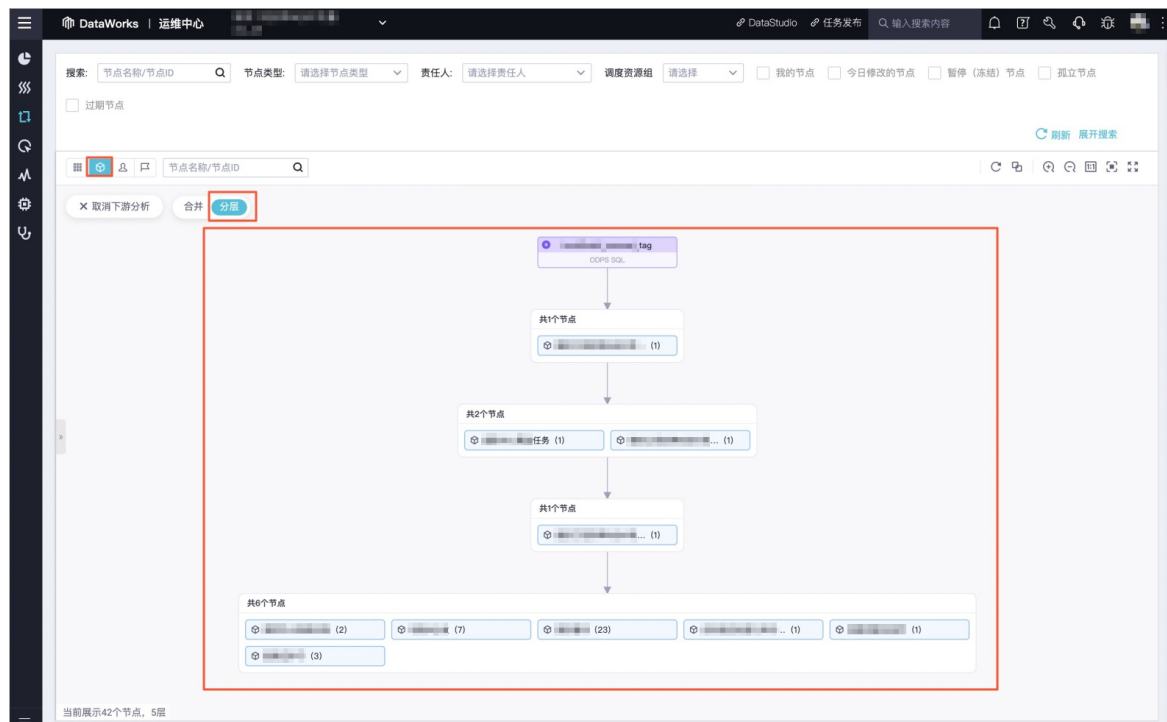
当周期任务的节点数量或层级过多时，您可以进行下游节点分析，通过任务状态、工作空间、责任人、优先级等维度进行筛选，统计不同层级目标维度的节点数量，或统计所有层级目标维度的节点总数量。

说明

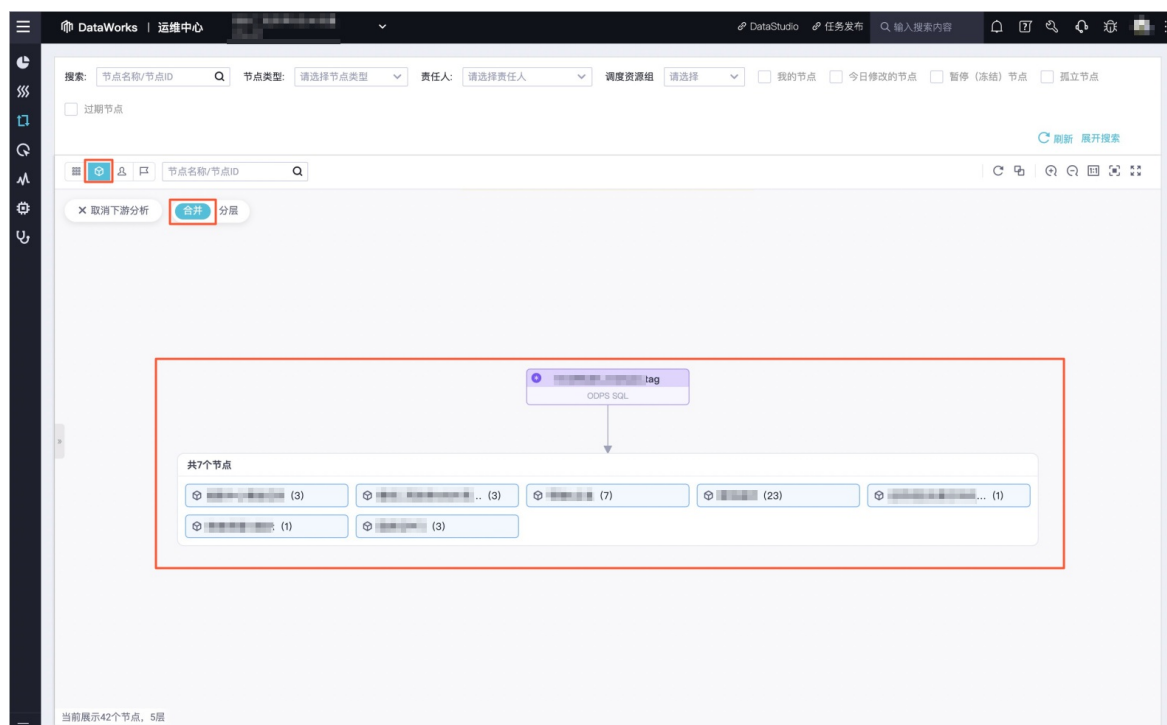
- 下游节点默认按照责任人聚合，统计所有层级该维度的节点总数量。
- 下游分析使用分层展示下游节点分析结果时，最多展示6层。如果您希望展示更多层级，请单击左上角的继续分析。

示例对tag节点进行下游分析，结果如下：

- 使用分层方式，并按照所属工作空间聚合，在不同层级展示不同工作空间的节点数量。



- 使用合并方式，并按照所属工作空间聚合，将所有下游节点合并至同一层，展示所属不同工作空间的节点数量。

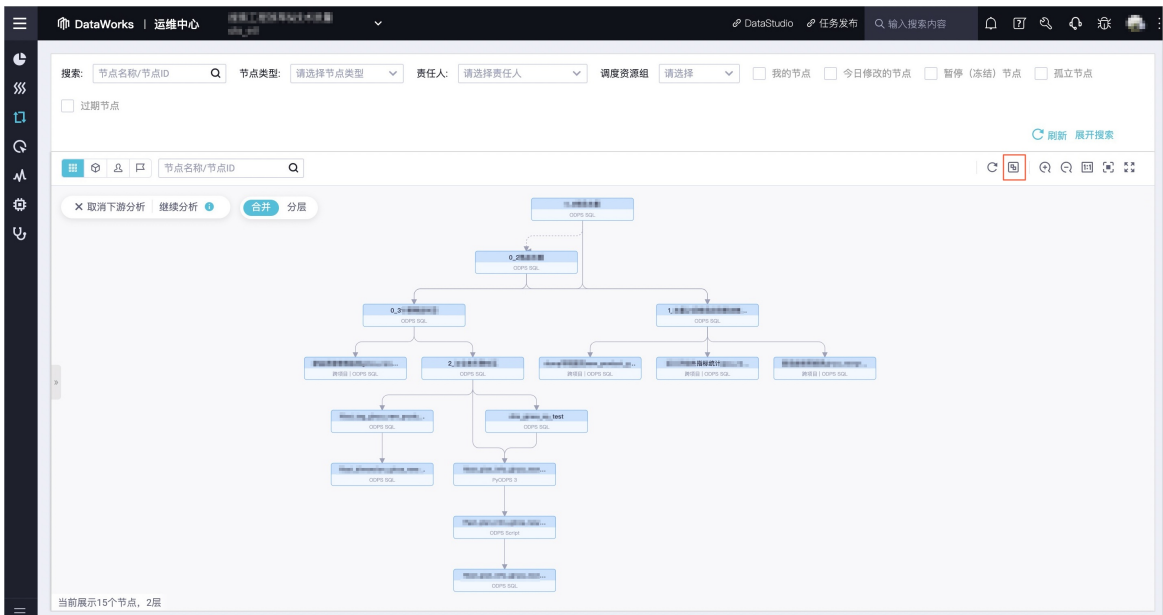


- 选择DAG图的显示样式。

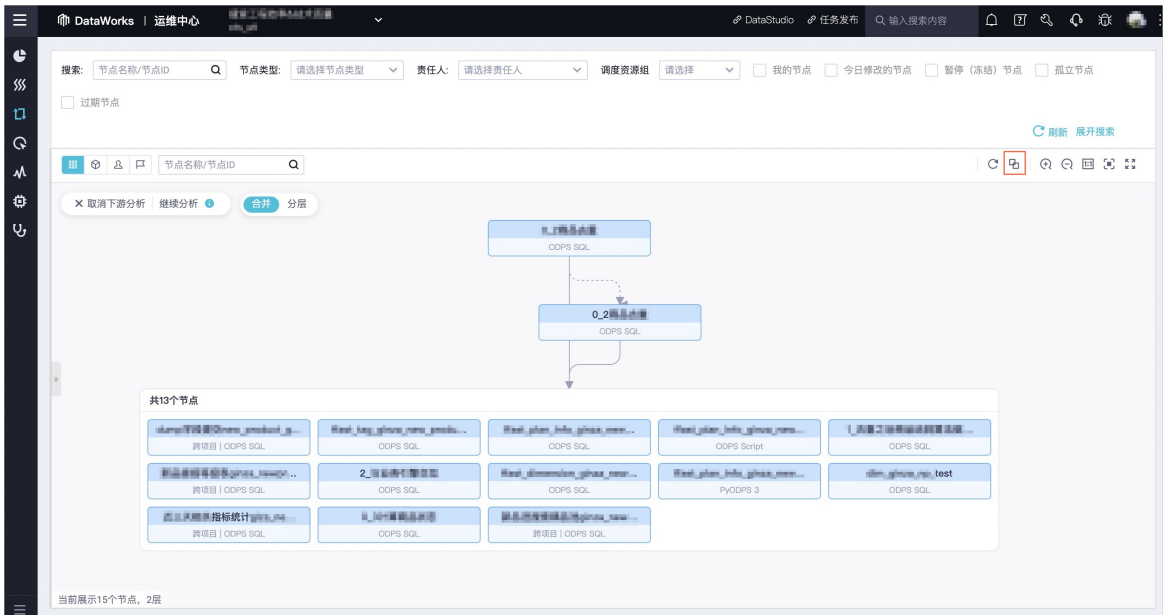
您可以根据实际需要，单击DAG图所在界面右上角的功能图标，调整DAG图的显示样式。例如，全屏展示、适配屏幕展示等。

示例对0_2节点的下游节点，展示拆分组及成组的DAG图。

- 拆分组 的DAG图如下所示。通过该方式，您可以直观的了解到所有节点的上下游关系。



- 成组 的DAG图如下所示。通过该方式，以每行呈现5个节点任务，有序的为您展示下游节点，方便您快速获取下游节点的总数量。



- 右键单击目标节点，执行相关操作。



说明 单击右上角的刷新图标，只能刷新实例的DAG状态，不能刷新实例的运行日志。

操作	描述
展开父节点/子节点	当一个业务流程有3个及以上节点时，运维中心展示任务时会自动隐藏节点。您可以通过展开父子层级查看全部节点的内容。
查看运行日志	查看当前实例正在运行、成功、失败等状态的运行日志。
查看代码	查看当前实例的代码。
编辑节点	单击后，即可跳转至数据开发页面，对当前节点的内容进行修改。
查看血缘	查看当前实例的血缘关系。
终止运行	仅支持对状态为等待运行和运行中的实例进行终止运行操作，进行此操作后，该实例的状态被置为失败。
重跑	失败的任务或状态异常的任务重跑实例。
重跑下游	当前节点的下游重跑实例，如果存在多个下游实例，会将这些实例全部重跑。
置成功	修改当前实例的状态为成功，并运行下游未运行状态的任务。常用于处理出错节点。 说明 仅支持置成功状态为失败的任务，不能置成功业务流程。

操作	描述
紧急操作	当前实例在非常紧急的情况下的操作，紧急操作只对当前节点本次有效。 选择 去除依赖 ，即可解除当前节点的依赖关系。常用于上游失败并与此实例没有数据关系时，启动此节点。
暂停（冻结）	将当前实例置为暂停（冻结）状态，并停止调度。
恢复（解冻）	恢复暂停（冻结）的节点的调度。

1.4.5. 测试实例

测试实例是对周期任务进行测试时产生的实例，可以对测试实例进行运维管理。

 **注意** 您需要购买DataWorks专业版及以上版本，测试实例即可支持运行诊断功能，详情请参见[运行诊断](#)。

使用限制

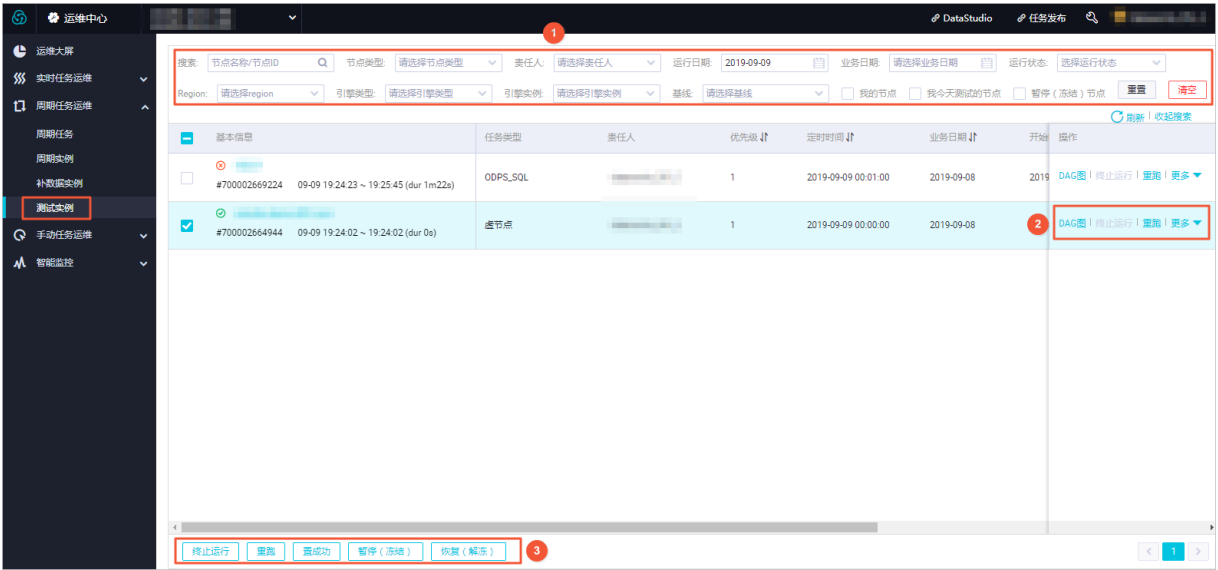
测试实例必须等周期实例完成后才能启动运行。

进入测试实例页面

1. 登录[DataWorks控制台](#)。
2. 在左侧导航栏，单击**工作空间列表**。
3. 单击相应工作空间后的**进入数据开发**。
4. 单击左上角的图标，选择**全部产品 > 运维中心**。
5. 在左侧导航栏，单击**周期任务运维 > 测试实例**，查看测试实例列表或DAG图。

测试实例列表

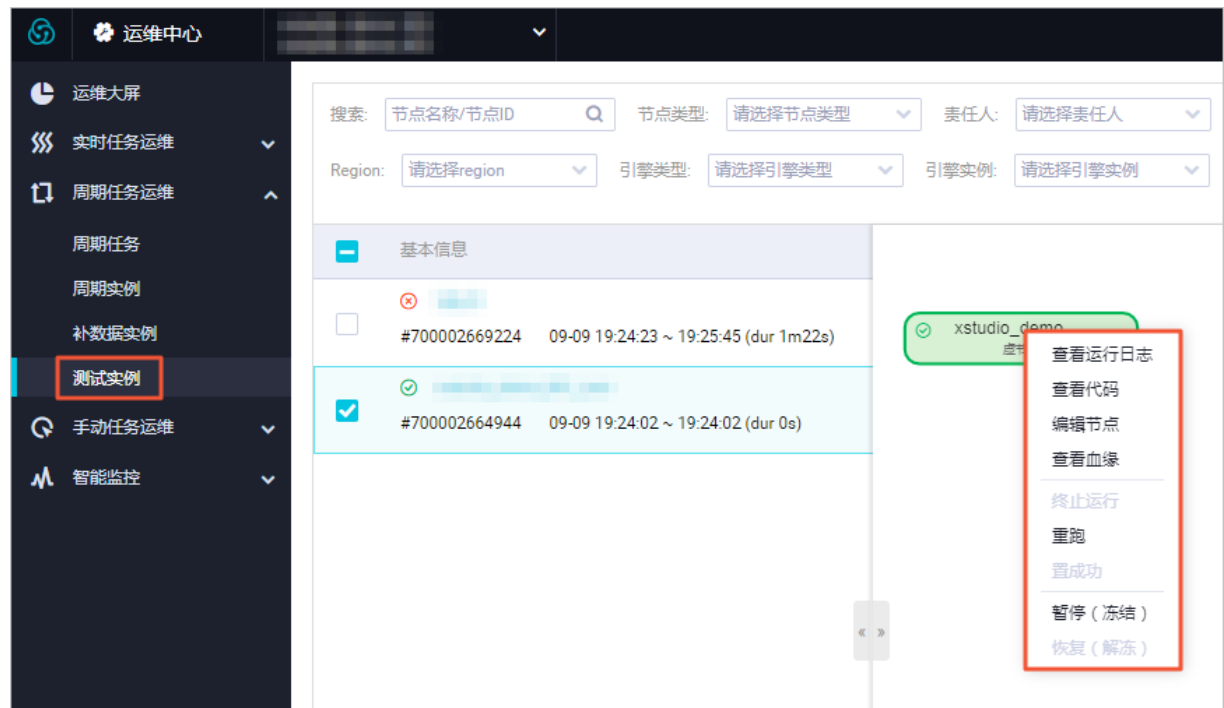
测试实例列表以列表形式对被调度的任务进行运维和管理，包括重跑、置成功、暂停（解冻）、恢复（解冻）、查看血缘和查看运行日志等。



操作	说明
筛选	在上图的模块1区域，您可以通过筛选条件过滤出需要查询的实例。 您可以根据节点名称、节点ID、责任人、运行日期、业务日期、运行状态、Region、引擎类型、引擎实例、基线、我的节点、我今天测试的节点和暂停（冻结）等条件进行精确筛选。
终止运行	仅支持对状态为等待运行和运行中的实例进行终止运行操作，进行此操作后，该实例的状态被置为失败。
重跑	可以重跑某任务，任务执行成功后可以触发下游未运行状态任务的调度。常用于处理出错节点和漏跑节点。 说明 仅支持重跑成功和失败状态的任务。
更多	单击操作栏中的更多，可以进行置成功、暂停（冻结）、恢复（解冻）、查看血缘和查看运行日志等更多操作。
批量操作	在上图的模块3区域，您可以批量选择任务，进行终止运行、重跑、置成功、冻结和解冻等操作。

测试实例DAG图

单击实例名或操作栏中的DAG图，即可打开该实例的DAG图。您可以在DAG图中，右键单击实例进行相关操作。



操作	说明
查看运行日志	查看当前实例正在运行、成功、失败等状态的运行日志。
查看代码	查看当前实例的代码。
编辑节点	单击后，即可跳转至数据开发页面，对当前节点的内容进行修改。
查看血缘	查看当前实例的血缘关系。
终止运行	仅仅支持对状态为等待运行和运行中的实例进行终止运行操作，进行此操作后，该实例的状态被置为失败。
重跑	可以重跑某任务，任务执行成功后可以触发下游未运行状态任务的调度。常用于处理出错节点和漏跑节点。 说明 仅支持重跑成功和失败状态的任务。
置成功	修改当前节点的状态为成功，并运行下游未运行状态的任务。常用于处理出错节点。 说明 仅支持置成功状态为失败的任务，不能置成功业务流程。
暂停（冻结）	将当前节点置为暂停（冻结）状态，并停止调度。

操作	说明
恢复（解冻）	可以恢复（解冻）状态为暂停（冻结）状态的实例。 - 如果该实例还未运行，则上游任务运行完毕后，会自动运行。 - 如果上游任务都运行完毕，则该任务会直接被置为失败。需手动重跑后，才会正常运行。

1.4.6. 运行诊断

DataWorks运维中心为您提供运行诊断功能，帮助您掌握任务运行的全链路信息，以迅速定位问题。

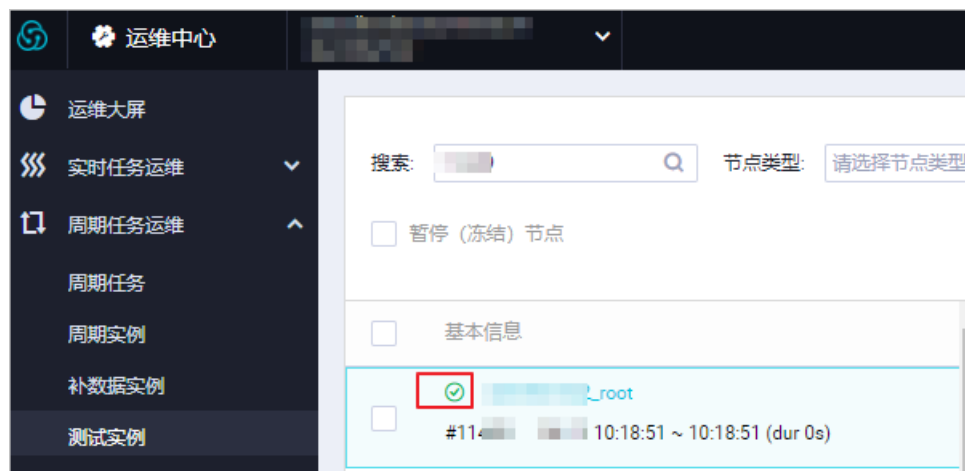
进入运行诊断

 **注意** DataWorks专业版及以上版本，支持使用运行诊断功能。您目前可以免费体验，但建议您升级到专业版以获取更多产品能力。

1. 登录[DataWorks控制台](#)。
2. 在左侧导航栏，单击工作空间列表。
3. 选择工作空间所在地域后，单击相应工作空间后的[进入数据开发](#)。
4. 单击左上角的图标，选择全部产品 > 任务运维 > 运维中心。
5. 在左侧导航栏，单击周期任务运维 > 周期实例。

您也可以进入补数据实例、测试实例和手动任务运维 > 手动实例页面，以同样的方式进入相应实例的运行诊断页面。

6. 单击相应实例前的状态图标，进入相应实例的运行诊断页面。



您还可以通过以下方式进入相应实例的运行诊断页面：

- 右键单击相应实例的DAG图，单击查看运行日志。在日志详情页面，单击立即前往任务运行诊断。
- 右键单击相应实例的DAG图，单击运行诊断。
- 展开实例列表，单击相应实例后的更多 > 运行诊断。

运行诊断

实例的运行诊断包括上游依赖、调度资源和任务执行：

● 上游依赖

运行诊断的上游依赖页面，为您展示该实例上游节点的运行情况，能够迅速定位阻塞本节点运行的上游节点。

运行诊断影响基线历史实例

基础信息

名称：di

节点ID：320170369

任务状态：运行成功

定时时间：2020-03-03 00:25:00

开始运行：2020-03-03 00:25:01

实例状态：实例运行失败

引擎实例：1

实例ID：510382724513

任务类型：数据集成

开始等待资源：2020-03-03 00:25:00

结束运行：2020-03-03 00:25:04

所属工作空间：

调度资源组：默认资源组

负责人：

调度类型：日调度

等待资源时间：1秒钟

执行参数：bizdate=20200302

引擎类型：

是否可重跑：否

实例运行诊断

上游依赖

定时检查

调度资源

任务执行

直接上游依赖 ①

名称	运行状态	调度类型	实例类型	工作空间	业务日期	定时时间	周期序号	运行次数	操作
vi 510382728686 2020-03-03 00:00:13 ~ 2020-03-03 00:24:00	运行成功	日调度	周期实例		2020-03-02	2020-03-03 00:24:00	1	1	运行诊断

● 调度资源

运行诊断的调度资源页面，为您展示资源使用率。您可以根据该页面展示的情况，合理安排任务的定时时间，以错峰运行。

in2

运行诊断影响基线历史实例

实例运行诊断

上游依赖

定时检查

调度资源

任务执行

2020-03-16 12:00:03~2020-03-16 12:01:09
(dur: 1m6s)
Gateway:11.194.111.220
运行失败

调度资源信息

资源组：default_resource_group正在运行的任务/正在等待的任务：0 / 0

每天0点~9点是DataWorks所有租户的任务高峰期，公共资源会被租户共享，可能出现任务等待资源的情况，您可以前往控制台添加专享资源组或自定义资源组来缓解资源紧张的情况。

资源使用趋势

任务等待资源：1秒钟

任务运行中：1分钟6秒钟



占用资源的任务

时间范围：2020-03-16 11:48:00 ~ 2020-03-16 12:12:00

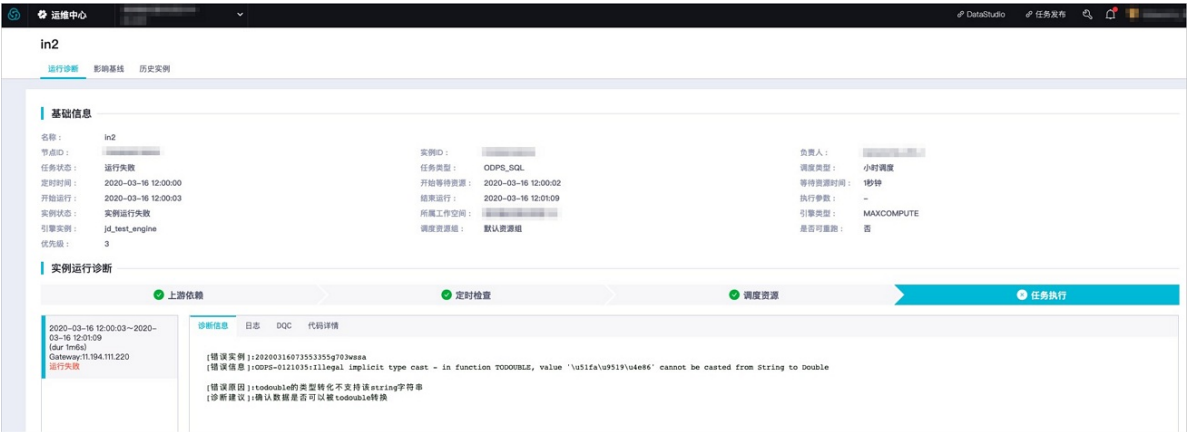
名称	节点ID	实例ID	任务状态	调度类型	实例类型	定时时间	开始运行	结束运行	运行环境	所属工作空间	负责人	运行诊断
in3			运行成功	小时调度	周期实例	2020-03-10 16:00:00	2020-03-10 16:00:10	2020-03-10 16:00:10	生产			运行诊断
in1			运行成功	小时调度	周期实例	2020-03-16 12:00:00	2020-03-16 12:00:02	2020-03-16 12:00:09	生产			运行诊断

● 任务执行

67

> 文档版本：20211230

对于运行失败的任务，任务执行页面可以根据日志信息，为您提供智能诊断建议。您还可以查看Instance列表进行进一步的分析。



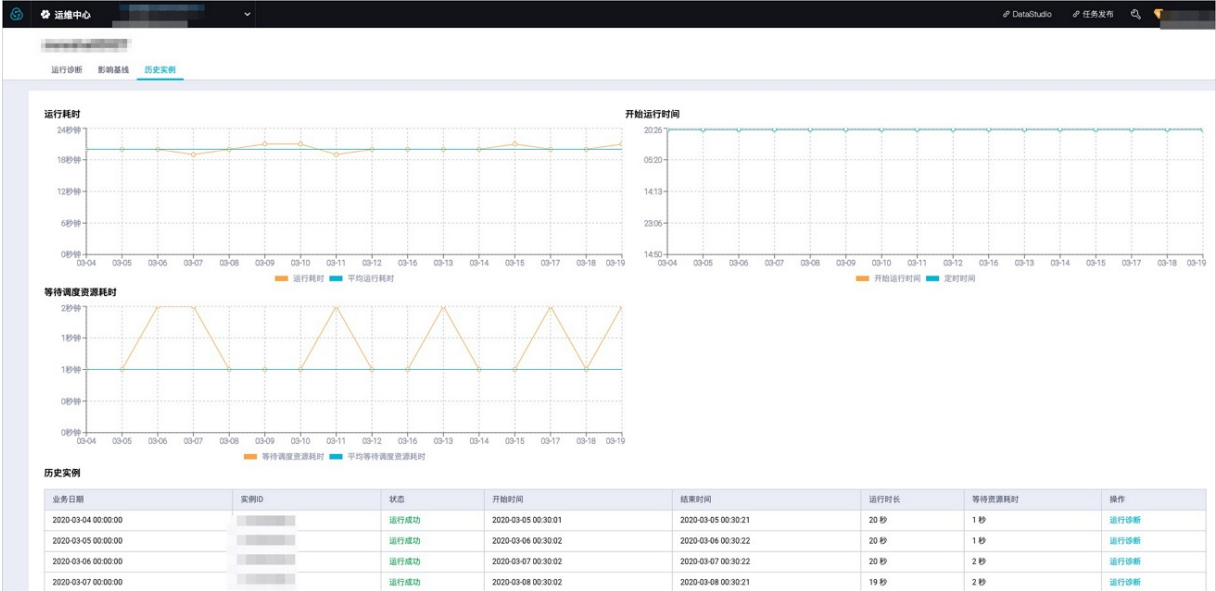
注意 仅MaxCompute类型的任务支持查看Instance列表。

影响基线

在顶部菜单栏，单击影响基线，即可查看该节点所在的基线。

历史实例

在顶部菜单栏，单击历史实例，即可查看最近15次实例的运行耗时、开始运行时间、等待调度资源耗时和历史实例。



1.5. 手动任务运维

1.5.1. 手动任务

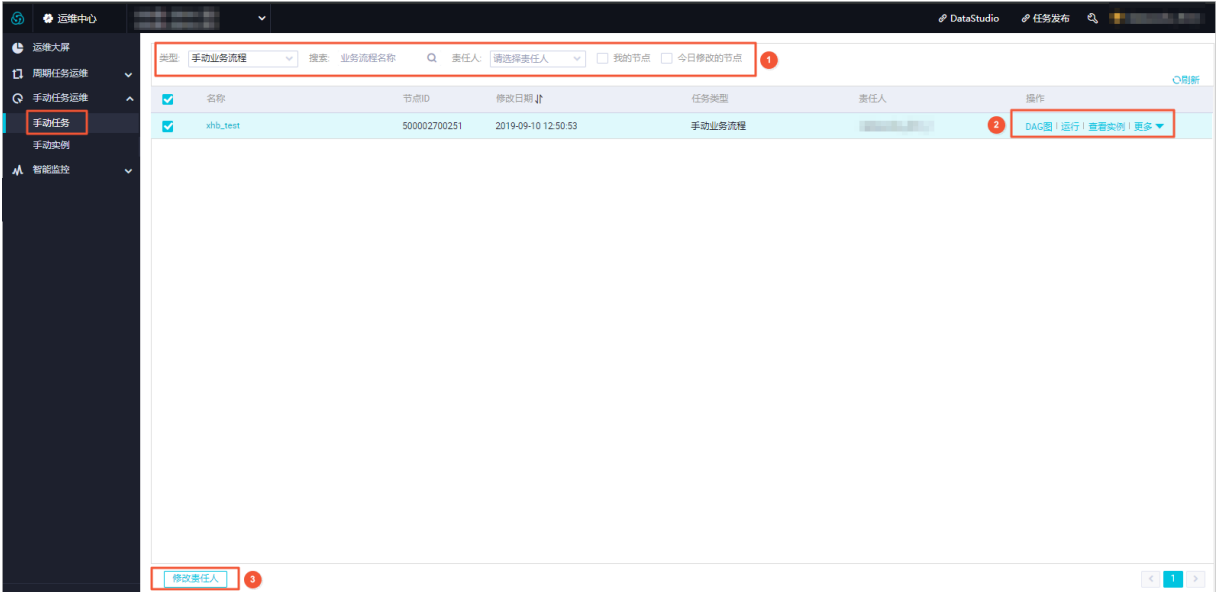
手动任务是指新建任务时，调度类型选择手动任务后，提交至调度系统的任务。

说明

- 手动任务提交至调度系统后，不会自动运行，只有手动触发才会运行。
- 目前DataWorks V1.0创建的手动任务显示在手动任务选项下，DataWorks V2.0及以上版本创建的手动任务显示在手动业务流程选项下。

手动业务流程列表

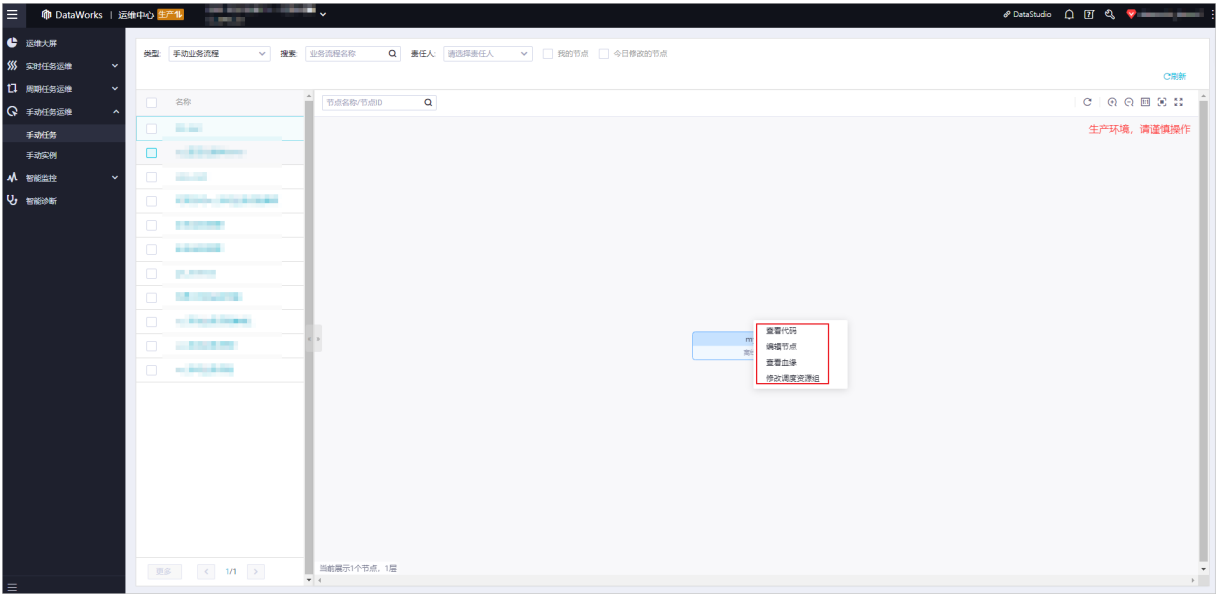
手动业务流程列表以列表的形式展示已提交的手动任务。



操作	描述
筛选	在上图的模块1区域，您可以通过筛选条件过滤出需要查询的任务。 您可以根据类型、业务流程名称、责任人、我的节点和今日修改的节点等条件进行精确筛选。 说明 任务名搜索的结果会受到其它筛选条件的影响，只有同时满足所有筛选条件的结果方可展示。
DAG图	单击操作栏中的DAG图，即可打开当前业务流程的DAG图。您可以在DAG图中查看业务流程的代码、血缘等信息。
运行	单击操作栏中的运行，即可运行当前手动业务流程，产生手动实例。
查看实例	单击操作栏中的查看实例，即可跳转至手动实例页面，查看手动业务流程的运行结果。
更多	单击更多 > 修改责任人，即可修改当前手动业务流程的节点责任人。
批量操作	在上图的模块3区域，您可以批量选择任务，进行修改责任人的操作。

手动业务流程DAG图

单击业务流程名称或操作栏中的DAG图，即可打开当前业务流程的DAG图。您可以在DAG图中，右键单击业务流程进行相关操作。



操作	描述
查看代码	查看当前节点的代码。
编辑节点	单击后可跳转至数据开发页面，对业务流程内容进行修改。
查看血缘	查看当前业务流程的血缘关系图。
修改调度资源组	修改当前业务流程所在的资源组。

1.5.2. 手动实例

手动实例是指手动任务产生的实例，手动任务的特点是没有调度依赖，只需要手动触发即可。

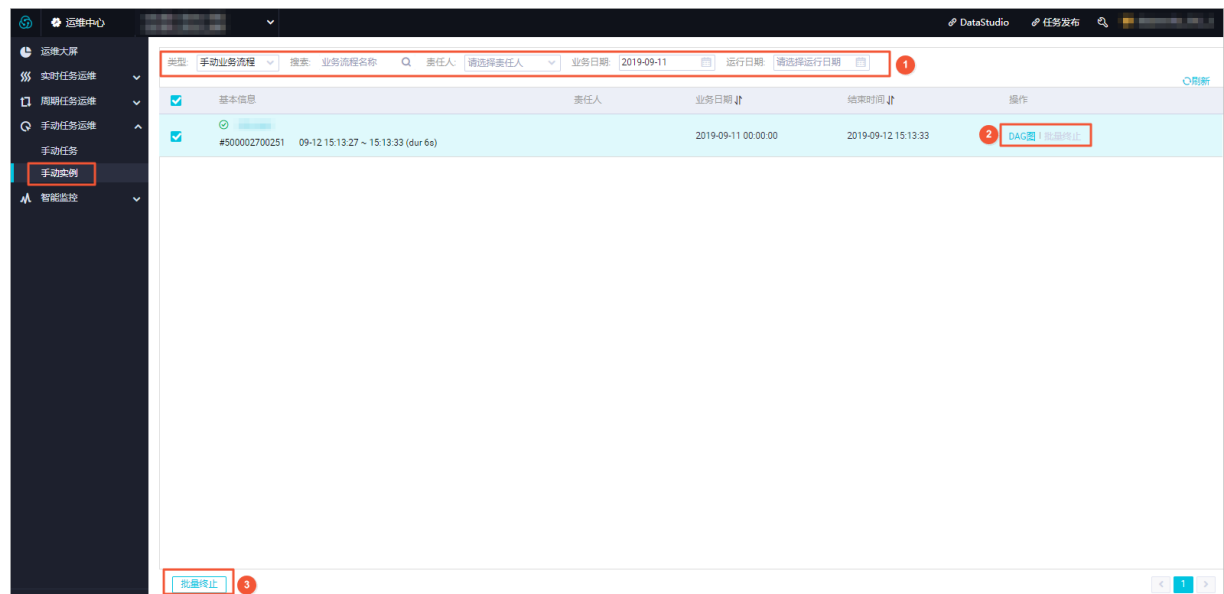
 注意

- 目前仅在周期实例任务失败时报警。手动实例、补数据实例和测试实例任务失败均无报警。
- 您需要购买DataWorks专业版及以上版本，手动实例即可支持运行诊断功能，详情请参见[运行诊断](#)。

进入手动实例页面

1. 登录[DataWorks控制台](#)。
2. 在左侧导航栏，单击[工作空间列表](#)。
3. 单击相应工作空间后的[进入数据开发](#)。
4. 单击左上角的图标，选择[全部产品 > 运维中心](#)。
5. 在左侧导航栏，单击[手动任务运维 > 手动实例](#)，查看手动业务流程或实例的列表、DAG图。

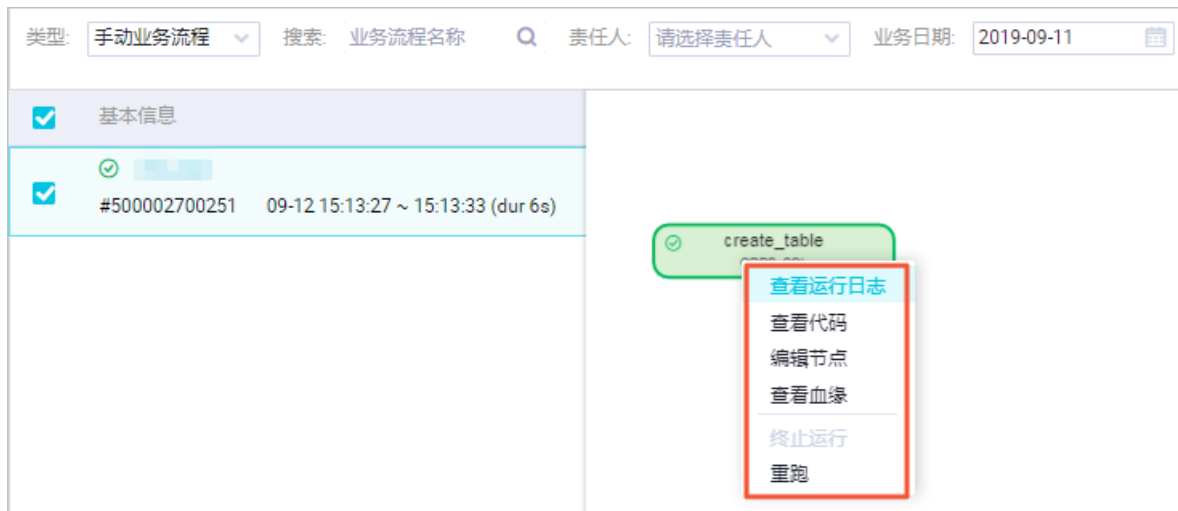
手动业务流程列表



操作	说明
筛选	在上图的模块1区域，您可以通过筛选条件过滤需要查询的任务。 您可以根据类型、业务流程名称、责任人、业务日期和运行日期等条件进行精确筛选。
DAG图	单击操作栏中的DAG图，即可打开当前业务流程的DAG图。查看实例运行的结果。
终止运行	如果实例在运行中，可以单击终止运行，停止任务运行。
重跑	重新调度当前实例。
批量操作	在上图的模块3区域，您可以批量选择任务，进行批量终止的操作。

手动业务流程DAG图

单击业务流程名称或操作栏中的DAG图，即可打开当前业务流程的DAG图。您可以在DAG图中，右键单击业务流程进行相关操作。



说明 手动业务流程没有依赖关系，所以DAG图中只会显示当前实例。

操作	说明
查看运行日志	查看当前实例正在运行、成功、失败等状态的运行日志。
查看代码	查看当前节点的代码。
编辑节点	单击后，即可跳转至数据开发页面，对当前业务流程的内容进行修改。
查看血缘	查看当前实例的血缘关系。
终止运行	终止任务的运行，仅对当前实例有效。
重跑	失败的任务或状态异常的任务重跑实例。

1.6. 智能监控

1.6.1. 智能监控概述

智能监控功能用于监控周期实例的运行情况，系统根据全局规则、自定义规则及智能基线，一旦检测到规则被触发，将根据您的设置，使用邮件、短信、电话、钉钉群消息、WebHook等形式通知您，方便您及时发现并处理异常问题。

背景信息

普通的监控系统只能实现基础的监控操作，无法满足DataWorks的如下需求：

- 准确地梳理出需要被监控的任务。

DataWorks的任务量较大，并且任务之间的依赖关系复杂。即使您知道最重要的任务是什么，也很难查找该类任务的所有上游并全部进行监控。基于该情况，如果您直接监控所有任务，会触发较多无用的报警，导致有用报警被忽略，监控效果不佳。

- 针对每个任务监控的报警方式不同。

例如，部分监控需要任务运行超过1个小时后报警，而有些监控需要任务运行超过2个小时后报警。如果单独对每个任务设置监控极为繁琐，并且很难预估每个任务应该设置的报警阈值。

- 针对每个任务监控的报警时间不同。

例如，重要性较低的任务可以放至上午上班后再报警，而重要性较高的任务则需要在出现故障后立刻报警。常用的监控系统无法区分每个任务的重要性。

- 及时关闭报警。

如果报警一直进行提醒，需要在您响应时提供关闭报警的入口。

智能监控拥有一整套的监控报警逻辑，您只需要提供所关注业务的重要任务名称，即可监控整体任务的产出过程，并生成对应的标准统一的报警机制。同时，智能监控还提供了轻量级的自助配置监控功能，方便您可以根据自己的需求定义报警规则。

智能监控的全路径监控功能保障了阿里巴巴集团所有重要业务的任务整体产出链路，其上下游路径分析功能可以及时发现风险并为业务部门提供相应运维信息。在智能监控的分析体系下，阿里巴巴集团业务保持了长期的高稳定性。

使用限制

仅DataWorks标准版及以上版本支持使用[基线实例](#)、[基线管理](#)及[事件管理](#)功能。

注意事项

RAM用户如果希望通过短信、电话等方式接收告警信息，则需要主账号登录[RAM访问控制](#)，完善RAM用户的手机号码、邮箱等个人信息。完善RAM用户的个人信息，详情请参见[修改RAM用户基本信息](#)。

监控方式

智能监控为您提供了系统内置的全局规则、您自定义的智能基线及自定义规则的监控方式，您可以使用[智能监控](#)功能监控周期调度任务状态、周期任务自动调度产生的[周期实例](#)运行状态、实时计算任务运行状态，平台一旦检测到规则被触发，将会根据您的自定义规则及智能基线的报警方式，通过邮件、短信、电话或钉钉群消息等形式通知您，方便您及时发现并处理异常问题。

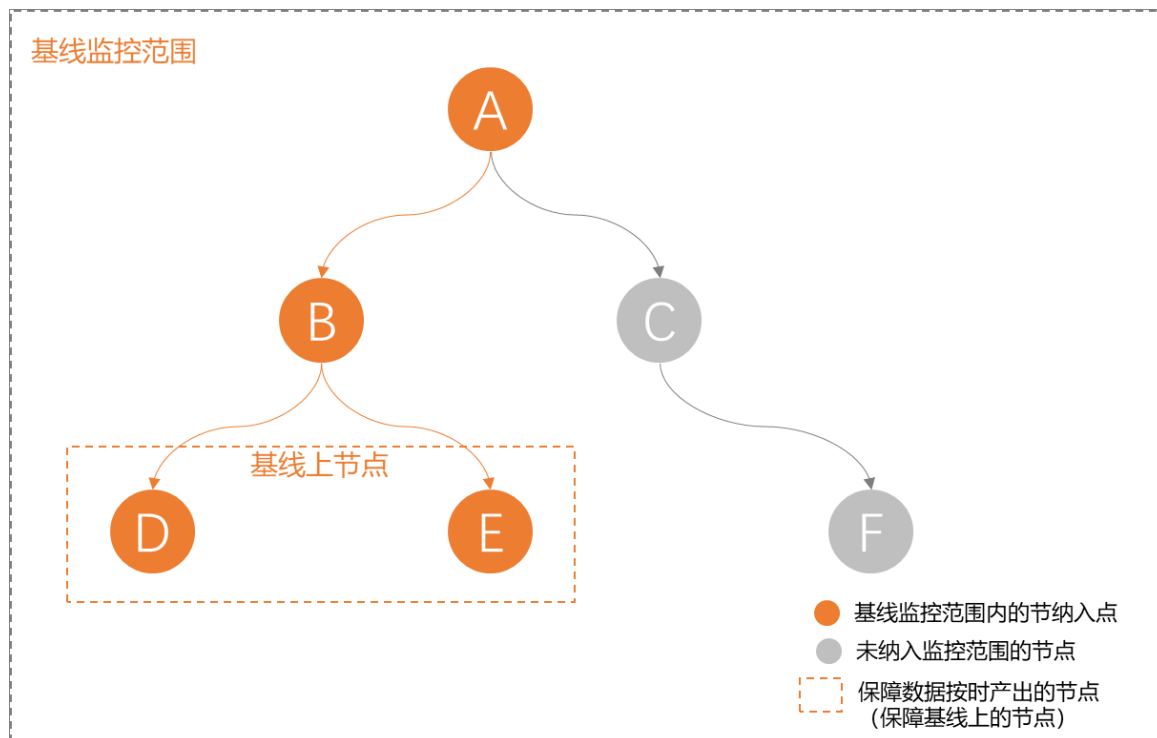
- 智能基线

在DataWorks上，周期任务周期性调度时将产生周期实例，当有任务需要重保并且上游任务依赖较为复杂时，可以使用基线功能将重要任务移到该基线上。当影响基线在预期时间内数据产出的情况发生时，基线将为您快速定位影响当前节点数据产出的关键路径，找到阻塞任务（即影响保障数据产出的关键实例）并第一时间发出报警，保障重要数据在预期时间内顺利完成。当影响基线的任务（基线上的任务及其上游任务）出错或者变慢，您将会收到相应的报警。其报警分为全局基线预警与全局事件报警（非空间级别报警规则）。支持短信、邮件、电话、钉钉机器人和webhook的方式发送报警给责任人（节点责任人、指定责任人或值班表中的排班等），关于基线使用详情请参见：[基线管理](#)。

○ 监控范围

基线是一组节点的管理单位，即节点分组，您可以将目标任务添加至相应基线进行智能监控。一条基线纳入监控后，该基线及基线上游的所有任务都会被监控。即影响基线上数据产出的上游任务才会被纳入基线监控范围。此时，智能监控不会默认监控所有任务。

- 当被监控的任务下游有被纳入监控基线的任务时，该下游任务才会被监控。
- 如果下游没有任务被纳入监控基线，即使该任务出错，智能监控也不会报警。



如上图所示，假设整个DataWorks有6个任务节点，任务D和任务E属于纳入基线上的节点，则任务D、任务E及它们所有的上游节点，均会被纳入监控范围。即上图中的任务A、任务B、任务D和任务E出现异常（出错或变慢），均会被智能监控察觉，而任务C和任务F不受智能监控所监控。

○ 捕获任务

监控范围确定后，当监控范围内的任务出现异常，智能监控会生成一个事件，基于对该事件的分析产生报警决策。任务的异常包括两种类型：出错、变慢，您可以进入[事件管理](#)查看影响基线上任务数据产出节点运行异常的事件详情。

- **出错**：任务运行失败。
- **变慢**：任务本次运行时间和过去一段时间内的平均运行时间相比，明显变长。

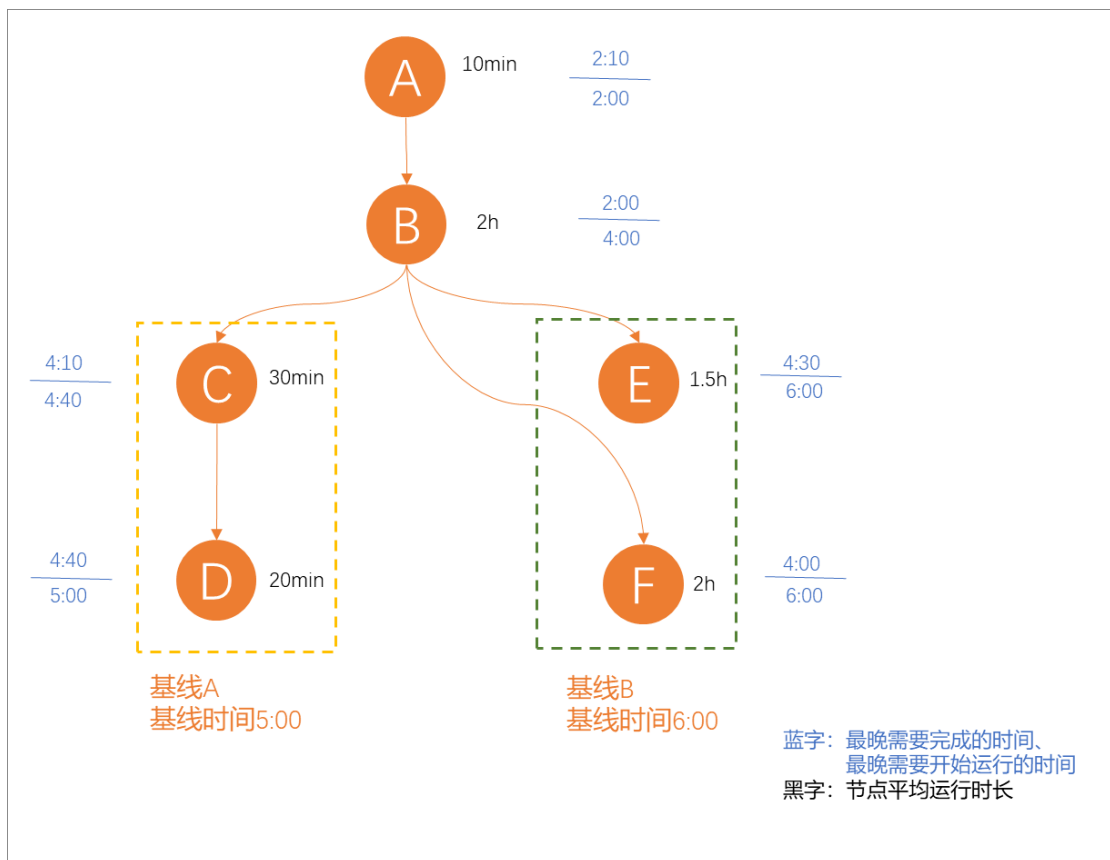
② 说明

- 影响重保基线上的任务数据产出的任务，即重保基线任务的上游任务出错或者变慢都会产生相应的事件，生成事件告警。
- 如果一个任务先超时、再出错，会生成两个事件。

○ 判断报警时机

a. 任务报警时机

余量是智能监控中目标任务可以被允许拖延开始执行的最长时间。 $\text{任务的最晚开始时间} = \text{基线时间} - \text{平均运行时间}$ 。



- 如上图所示，要满足基线A的基线时间5:00，依次推理：
 - A基线时间5:00 - 任务D运行时长20min - 任务C运行时长30min = 任务C最晚开始运行时间 4:10 = 任务B的最晚完成时间 4:10
 - 该时间也是任务B为满足A基线时间5:00数据产出的，任务B的最晚完成时间。（即上游B任务最晚完成时间为下游C任务最晚开始运行时间）
 - 如上图所示，要满足基线B的基线时间6:00，依次推理：
 - 任务B基线时间6:00 - 任务F运行时长2h = 任务B最晚完成时间 4:00
 - 综上所述：任务B的最晚完成时间为4:00时，方可同时满足基线A和基线B。
 - 4:00早于4:10，B任务满足B基线所需最晚完成时间早于B任务满足A基线时间所需最晚完成时间。

■ 任务异常警戒程度：余量

基线预警中通过余量来体现任务异常程度。余量是智能监控中目标任务可以被允许拖延开始执行的最长时间，是一个任务异常的警戒程度的体现。

余量计算：综上所述：

- 以A任务为例
 - 任务A最晚完成时间2:00 = 任务B最晚完成时间4:00 - 任务B任务运行时长2H。
 - 任务A最晚开始运行时间1:50 = 任务A最晚完成时间2:00 - 任务A的运行时长10min
- 如果任务A无法在1:50开始运行，则基线A容易破线。
 - 假设1点时任务A运行出错，则此时任务A的余量时间即为1:50和1:00之间的差值，即50分钟。由此可见余量是一个任务异常的警戒程度的体现。
 - 当任务A在1:50仍未开始运行或在最晚完成时间2:00仍未完成，将会触发事件报警。

○ 基线报警

基线报警是对开启基线监控的基线提供的一个报警提醒，每个基线都必须配置**预警余量**和**承诺时间**。当智能监控预测基线的完成时间超过预警余量时，会直接通知设置的报警对象3次，每次间隔30分钟。

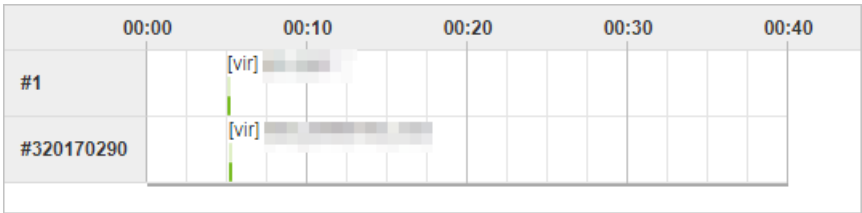
○ 报警方式

基线报警目前默认发送给基线责任人。您可以根据业务需求，在规则管理页面的**全局基线预警**，修改报警触发方式和报警行为。详情请参见[配置报警详情](#)

规则ID	规则名称	创建人	类型	对象类型	规则对象	触发条件	报警方式	接收人	操作
29681	节点孤立报警	-	全局规则	任务节点	所有节点	-	短信邮件	任务责任人	详情 关闭
29682	节点成环报警	-	全局规则	任务节点	所有节点	-	短信邮件	任务责任人	详情 关闭
29679	全局事件报警	-	全局规则	任务节点	所有事件	-	短信邮件	事件责任人	详情 关闭
29680	全局基线预警	-	全局规则	任务节点	所有基线实例	-	短信邮件	基线责任人	详情 关闭

○ 甘特图

甘特图属于智能监控**基线实例**的功能，用于反映目标任务的关键执行路径。



- 全局规则：包括节点孤立报警、节点成环报警、全局事件报警、全局基线预警，针对不同的对象触发不同的报警：

◦ 周期任务状态监控

DataWorks每晚根据周期任务生成第二天待自动调度运行的周期实例，所以为保障周期任务可以正常产生周期实例并且自动调度运行，DataWorks内置了全局报警规则来对周期任务进行定期监控扫描，如有异常便会自动报警。报警包括孤立节点与节点成环。

② 说明

- DataWorks每天定时9点、12点、16点对周期任务状态进行扫描，如有异常将会自动发送报警。但扫描时间点前10分钟内产生异常不会纳入本次扫描，该异常将会被纳入下一个周期的任务状态扫描中进行扫描。
- 全局规则为系统内置规则，您无须手动新建，默认以短信、邮件的方式报警给节点责任人。但您可以在规则管理页面中针对全局规则修改报警接收人。

■ 全局规则：节点孤立报警

孤立节点指节点在周期任务或周期实例中展开父节点时，没有依赖任何父节点，这类型节点在自动调度时不会被调度调起，所以孤立节点不会自动调度运行，如果该孤立节点下游依赖较多，则会造成严重的后果。孤立节点产生后会自动报警，如果收到孤立节点报警请及时处理。

② 说明 在DataWorks上，除了工作空间根节点外，自定义的每个周期调度的节点都需要有父节点依赖才可以被正常调度运行。

■ 全局规则：节点成环报警

节点成环指的是某节点为上游节点但又同时依赖了自己的下游节点。导致依赖关系成环，这类型节点在自动调度时不会被调度调起，节点成环后会自动报警，如果收到节点成环报警请及时处理。

◦ 周期实例运行状态监控

在DataWorks上，周期任务周期性调度时将产生周期实例，当有任务需要重保并且上游任务依赖较为复杂时，可以使用基线功能将重要任务移到该基线上。当影响基线在预期时间内数据产出的情况发生时，基线将为您快速定位影响当前节点数据产出的关键路径，找到阻塞任务（即影响保障数据产出的关键实例）并第一时间发出报警，保障重要数据在预期时间内顺利完成。当影响基线的任务（基线上的任务及其上游任务）出错或者变慢，您将会收到相应的报警。其报警分为全局基线预警与全局事件报警（非空间级别报警规则）。支持短信、邮件、电话、钉钉机器人和webhook的方式发送报警给责任人（节点责任人、指定责任人或值班表中的排班等），关于基线使用详情请参见：[基线管理](#)。

■ 全局基线预警

当基线根据基线的承诺时间、预警余量等预测无法保障基线上的任务在预警时间内完成时，便会触发基线提前预警功能。

■ 全局事件报警

当影响基线上数据产出的任务（基线上的任务及其上游任务）出错或者变慢时，将会收到相应的事件报警。

② 说明 全局规则为系统内置规则，您无须手动配置，但您可以在规则管理页面中针对全局规则修改报警接收人。

● 自定义规则

DataWork还提供了指定对象的自定义规则监控报警功能，实现对周期实例运行状态的监控。

◦ 自定义规则：指定对象监控

您可以对指定对象（任务节点、基线、工作空间、业务流程）上节点的状态（完成、未完成、出错、周期未完成、超时、自动重跑后仍出错等及实时计算任务的出错）设置监控报警。同时支持针对独享资源组利用率，及等待资源实例数超预期时告警，并通过您配置的告警方式来发送相应的告警信息，目前支持短信、邮件、电话、钉钉机器人和webhook的方式发送报警给责任人（节点责任人、指定责任人或值班表中的排班等），同时也支持疲劳度控制，即设置报警次数、报警间隔与报警接收时间。详情请参考[自定义规则](#)。

针对实例维度的监控

■ 自定义监控规则配置内容如下：

- 规则对象：包括节点、基线和工作空间。
- 触发条件：包括完成、未完成、出错、周期未完成、超时或自动重跑后仍出错。
- 报警方式：包括邮件、短信和电话。
- 最大报警次数：报警的最大次数，超过设置的次数后，不再产生报警。
- 最小报警间隔：两次报警之间的时间间隔。
- 免打扰时间：在设置的时间段内不会发送报警。
- 接收人：可以设置为责任人、值班表、其他接收人，支持钉钉群机器人告警。。

■ 自定义监控规则的触发条件如下：

■ 完成

可以设置针对节点、基线或业务流程完成的报警。一旦设置的对象上所有的任务均完成，则会发送报警。

② 说明

- 规则对象配置为基线时，当基线上所有的任务完成时，便会发送报警。
- 规则对象为业务流程时，当该业务流程下所有的任务完成时，便会发送报警。
- 规则对象为任务节点且添加了多个节点时，当所有节点执行完成时，才会发送告警。
- 不支持工作空间级别设置完成告警。
- 完成告警不支持任务责任人接收告警。

■ 未完成

针对节点、基线、工作空间或业务流程，从实例任务运行的起始时间点开始监控，到指定的目标时间点（[新建自定义规则](#)窗口配置的触发时刻）任务仍未结束运行，则系统发送报警。

② 说明 如果是小时分钟任务，这里设置的未完成时间告警指的是小时分钟任务当天所有的实例在设置的时间点未完成的话，便会告警。例如：

- 设置某条基线未完成时间为10点，则到10点时只要基线上有一个任务没有完成，便会发送报警，并推送没有完成的任务列表。
- 设置的小时任务每天00:00~23:59每小时调度一次，如果小时任务设置未完成时间为12点，那该规则每天都会触发告警，小时分钟任务当天某一周期未完成建议设置周期未完成告警。

■ 出错

可以设置针对节点、基线、工作空间或业务流程出错报警。任务一旦出错，则报警给设置的报警对象，并推送详细的任务出错信息。

② 说明

- 规则对象配置为基线、工作空间或业务流程时，当基线、工作空间或业务流程内的某个节点失败，便会触发告警。
- 此配置项监控的任务每次失败都会触发告警，如果任务重跑3次，3次均失败，将会触发三次该告警。

■ 周期未完成

表示在指定的周期内，实例任务仍未结束运行，则系统发送报警。通常用于监控以小时、分钟为周期单位的实例任务。例如，任务A每2小时调度一次，运行一次耗时25min。运行起始时间为每日0点0分，则该任务一天（24小时）共有12个任务周期，0点为第一个周期，2点为第二个周期，依次类推，22点为第12个周期。该任务正常运行时，会在每日0点25分、2点25分等时间节点执行完毕。如果在任意周期结束时间点该任务仍未结束运行，则发送报警。

② 说明 冻结的实例默认为完成状态。

■ 超时

可以设置针对节点、基线、工作空间或业务流程的超时报警。一旦设置的对象中，发现有被监控的任务执行时长超过指定的运行时长后仍处于未成功状态，则会发送报警。

② 说明 运行时长通过新建自定义规则窗口的开始运行起参数进行配置。

■ 自动重跑后仍出错

可以设置针对节点、基线、工作空间或业务流程告警，如果被监控的任务运行出错且自动重跑后仍出错，则系统发送报警。

针对资源组维度的监控：支持针对独享调度资源组与独享数据集成资源组维度的利用率监控与等待资源实例数监控。

○ 实时计算任务运行状态监控

您可以通过自定义规则，对实时计算任务的运行状态进行监控，当任务节点出错时通过配置的短信、邮件、电话、钉钉机器人和webhook等方式发送报警给责任人（节点责任人、指定责任人和值班表中的排班等）。详情请参考[自定义规则](#)

功能介绍

智能监控能够帮助您降低配置成本、杜绝无效报警自动覆盖所有重要任务，其主要的功能模块如下：

- **基线实例**：用于查看基线的相关信息，详情请参见[基线实例](#)。
- **基线管理**：用于创建和定义基线，详情请参见[基线管理](#)。
- **事件管理**：用于查看当前所有变慢和出错的事件，详情请参见[事件管理](#)。
- **规则管理**：用于自定义报警规则，以及添加钉钉群机器人获取Webhook地址。您可以根据实际业务需求，使用自定义监控规则定制监控内容，监控实例任务的运行状态，以便及时发现并处理异常。详情请参见[自定义规则](#)。
- **报警信息**：用于查看所有的报警信息，便于及时发现并处理异常。详情请参见[报警信息](#)。

- **值班表**：您可以使用DataWorks值班表功能设置实例运维报警的值班机制。当自定义规则的报警接收人设置为值班表后，DataWorks将按照排班将报警信息发送给对应的值班人，以便值班人员及时发现并处理问题。详情请参见[值班表](#)。

智能监控的相关常见问题，详情请参见[常见问题](#)。

1.6.2. 智能基线

1.6.2.1. 基线管理

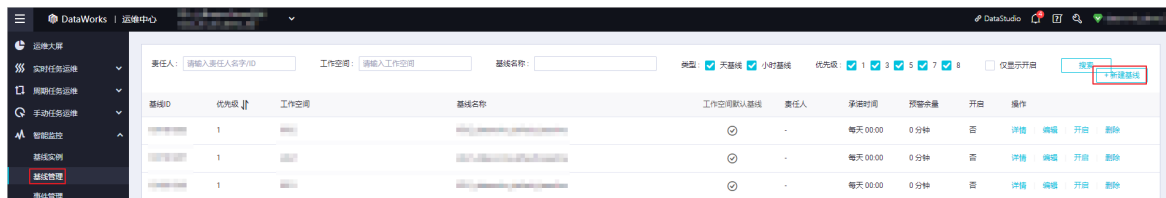
基线管理用于将原本优先级较低的任务节点或业务流程，添加至一个基线中，统一配置优先级和运行时间，保障任务优先运行的链路。本文为您介绍如何创建和定义基线。

使用限制

- 您需要购买DataWorks标准版及以上版本，才可以使用基线管理功能。
- 目前仅项目管理员可以创建基线。

创建基线

1. 进入数据开发页面。
 - i. 登录[DataWorks控制台](#)。
 - ii. 在左侧导航栏，单击工作空间列表。
 - iii. 选择工作空间所在地域后，单击相应工作空间后的进入数据开发。
2. 单击左上角的图标，选择全部产品 > 任务运维 > 运维中心（工作流）。
3. 单击左侧导航栏中的智能监控 > 基线管理。
4. 在基线管理页面，单击右上角的+新建基线，即可创建基线。



④ 说明 目前仅项目管理员可以创建基线。

5. 配置新建基线对话框中的各配参数。

新建基线

基线名称：

所属工作空间：

lzz_test001

责任人：

请输入责任人名字/ID

基线类型：

天基线

小时基线

保障任务：

序号

节点名称

责任人

工作空间

没有数据

任务节点

请输入任务节点名称/ID

优先级：

请选择

预计完成时间：

（历史数据不足，暂无法预估）

承诺时间：

每天

请选择时间

预警余量：

0

分钟

确定

取消

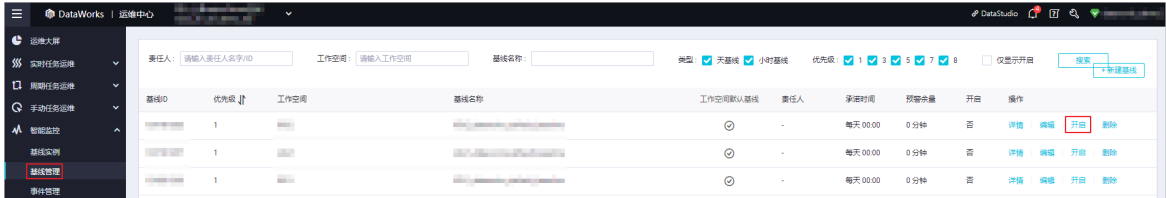
参数	说明
基线名称	输入基线的名称。
所属工作空间	基线关联的任务所属的工作空间。
责任人	您可以根据责任人名称或ID搜索创建的基线。
基线类型	包括天基线和小时基线，决定基线是按天还是按小时进行检测。 - 天基线：对应调度配置中的天调度任务。 - 小时基线：对应调度配置中的小时调度任务。
保障任务	- 任务节点：基线具体关联的任务节点，输入任务节点名称或ID后，单击右侧的图标进行添加，可以添加多个任务节点。 - 业务流程：输入业务流程名称或ID后，单击右侧的图标进行添加。 ? 说明 选择业务流程后，会添加业务流程中的所有任务。建议仅添加业务流程最下游的节点任务，不推荐选择业务流程。

81

> 文档版本：20211230

参数	说明
优先级	优先调度数值较高的基线，默认优先级为1。 说明 - 包年包月的工作空间才可以修改任务优先级，包括1、3、5、7和8。 - 数值越大优先级越高。
预计完成时间	根据任务节点之前完成周期调度的平均时间进行预估。如果没有历史数据，则系统会提示历史数据不足，暂无法预估。
承诺时间	如果实际完成时间晚于 承诺时间-预警余量时间 ，则会触发报警。
预警余量	例如，设置承诺完成时间为3:30，预警余量为10分钟，如果3:20任务没有完成，便会报警。假设此任务的平均运行时间是30分钟，如果2:50该任务仍未开始运行，便会告警。 说明 根据最近15天的平均值，可以推算该任务的平均运行时间为30分钟。

- 单击**确定**，完成基线的创建。
- 完成基线的创建后，单击相应基线操作栏中的**开启**，即可打开基线开关。



您可以单击相应基线后的**详情**、**编辑**、**开启**、**关闭**或**删除**，进行相关操作。

- 详情**：单击**详情**，即可查看基线任务的基本情况。
- 编辑**：单击**编辑**，即可直接修改基线任务。
- 开启或关闭**：控制基线任务的状态，开启后方可生成周期实例。
- 删除**：单击**删除**，即可直接删除基线任务。

添加任务

生产环境的任务默认都在**项目默认基线**上，添加基线实际是将任务从默认基线转移至您新添加的基线上。

说明 由于任务必须在基线上，所以不可以删除默认基线上的任务。删除自定义基线（您自行添加的基线）上的任务，实际是把您的任务从自定义基线转移至默认基线上。

您可以通过以下2种方式修改任务基线：

- 进入**基线管理**页面，单击右上角的**新建基线**进行添加。
- 进入**周期任务**页面，选择相应任务后的**更多 > 添加到基线**。



1.6.2.2. 基线实例

基线实例主要用于查看基线的相关信息。

说明

- 您需要购买DataWorks标准版及以上版本，才可以使用基线实例功能。
- 基线创建完成后，需要开启基线开关才会生成基线实例。

在基线实例页面，您可以通过业务日期、责任人、相关事件ID、工作空间和基线名称、类型、优先级和基线状态等搜索相应实例，并对其进行查看详情、处理和查看甘特图等操作。

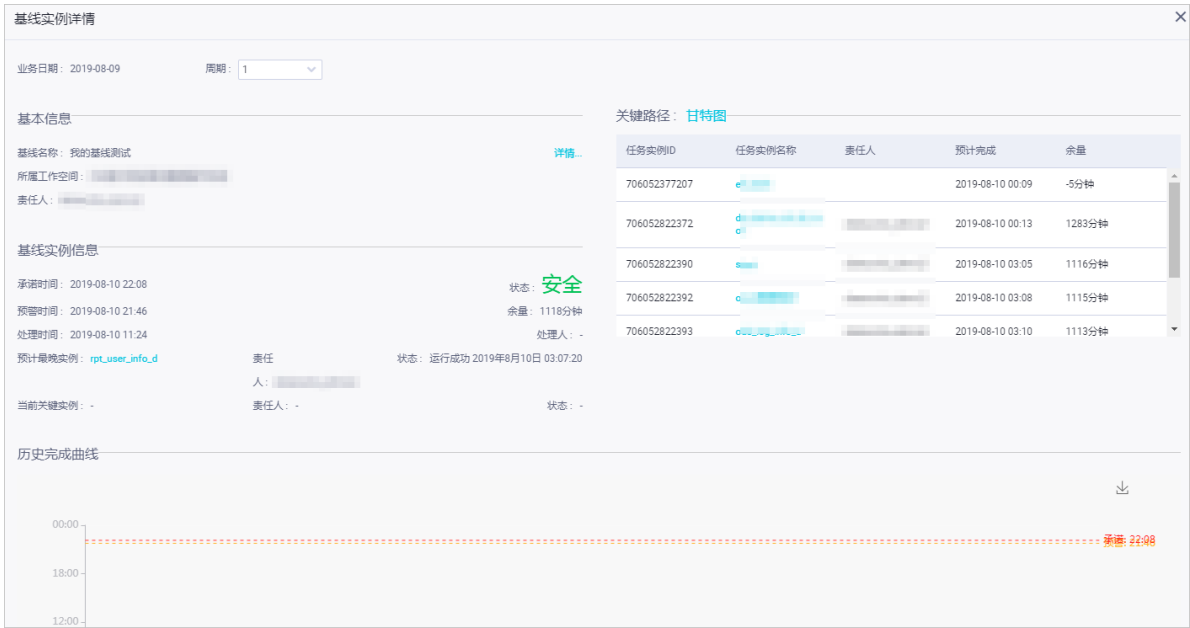


基线包括以下4种状态：

- **安全**：任务在预警时间之前完成。
- **预警**：任务在预警时间之后未完成，但还未到达承诺时间。
- **破线**：任务在承诺时间之后仍未完成。
- **其他**：基线所有任务处于暂停状态或基线没有关联任务。

基线的操作栏下包括详情、处理和甘特图。

- **详情**：单击详情，即可查看基线实例详情。

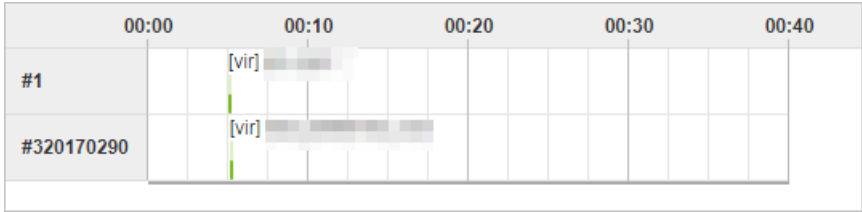


详情页面包括基本信息、关键路径、基线实例信息、历史完成曲线和相关事件。

说明

- 业务时间是系统时间-1天。
- 仅小时基线有周期。

- 处理：报警的基线在处理时间内停止报警。
- 甘特图：单击甘特图，即可查看任务的关键路径情况。



常见问题：为什么在基线实例页面会显示基线状态为空基线？

以下几个将会导致空基线产生：

- 场景一：节点只能存在在一条基线上。当任务从一条基线添加到另一条基线，实际上是节点在基线上的迁移，当一条基线上一个任务都没有的时候（原基线上的任务全都被移走），这条开启状态的基线会成为空基线。
- 场景二：基线新建当天会显示空基线，基线开启第二天正常产生基线实例。
- 场景三：小时基线对应的实例周期填写错误。

说明

假设小时任务定时每天6点，18点执行，那么此任务周期数为2，在小时基线配置时，任务第一个周期为6点，第二个周期为18点。

1.6.2.3. 事件管理

您可以在事件管理页面查看目前所有变慢和出错的事件。

说明 您需要购买DataWorks标准版及以上版本，才可以使用事件管理功能。

您可以进入运维中心 > 智能监控 > 事件管理页面，通过责任人、发现时间、事件状态、事件类型、任务节点或任务实例的名称/ID等条件进行搜索。



对于搜索出的结果，每一行代表一个事件（即关联到一个异常的任务）。最差基线代表该事件所影响到的基线中余量时间最少的基线。

- 单击相应事件操作栏中的详情，即可查看事件发生时间、告警时间、恢复事件、任务的过往运行记录及详细的任务日志。

其中，实际报警接收人为指派给的人，单击报警信息即可跳转至事件对应的报警详情页面。基线影响会显示该事件对应的任务影响的所有下游基线，通过观察对应的下游基线和破线程度，结合任务日志，可以判断该事件发生的具体原因。



- 单击处理后，事件的处理操作记录会被记录，并且在操作期间暂停报警。
- 单击忽略后，事件的忽略记录会被记录，并且永久停止报警。

1.6.3. 自定义规则

1.6.3.1. 自定义规则

自定义监控规则帮助您根据实际业务需求，定制监控内容，监控实例任务的运行状态，以便及时发现并处理异常。本文为您介绍如何在规则管理页面自定义报警规则，以及添加钉钉群机器人获取Webhook地址。

使用限制

- 自定义监控规则仅对周期实例生效。

- 电话告警功能可用版本：仅DataWorks专业版及以上版本，才可以使用电话告警功能。
- WebHook报警功能的使用限制如下：
 - 可用版本：仅DataWorks企业版及旗舰版支持使用WebHook功能。
 - 可用地域：仅西南1（成都）、华北3（张家口）、华北2（北京）、华东1（杭州）、华南1（深圳）、中国（香港）、欧洲中部1（法兰克福）、亚太东南1（新加坡）地域支持自定义规则和基线使用WebHook功能报警。
 - 可用报警方式：DataWorks仅支持使用WebHook功能推送报警信息至企业微信或飞书。

 **说明** 如果您需要使用其他的webhook，请工单联系我们。

创建自定义规则

1. 进入数据开发页面。
 - i. 登录DataWorks控制台。
 - ii. 在左侧导航栏，单击工作空间列表。
 - iii. 选择工作空间所在地域后，单击相应工作空间后的进入数据开发。
2. 单击左上角的图标，选择全部产品 > 任务运维 > 运维中心（工作流），进入运维中心页面。
3. 在左侧导航栏，单击智能监控 > 规则管理。
4. 在规则管理页面，单击右上角的新建自定义规则。
5. 在新建自定义规则对话框中，配置各项参数。

新建自定义规则

基本信息

规则名称:

请输入规则名称

对象类型:

任务节点

规则对象:

周期任务

请输入任务节点名称/ID

添加

序号	任务名称	责任人	工作空间
没有数据			

触发方式

触发条件:

选择触发条件

报警行为

请完善接收人的手机/邮箱信息以确保报警能被正常接收。

报警方式:

☒ 短信

☒ 邮件

☐ 电话

☐ 钉钉群机器人

☐ WebHook

接收人:

☐ 任务责任人

☐ 值班表

请选择值班表

☐ 其他

请输入接收人名字/ID

疲劳度控制

最大报警次数:

3

次

最小报警间隔:

30

分钟

免打扰时间:

00:00 至 00:00

确认

取消

分类	参数	描述
基本信息	规则名称	输入新建自定义规则的名称。
	对象类型	控制监控的粒度，包括任务节点、基线、工作空间、业务流程、独享调度资源组和独享数据集成资源组。
	规则对象	如果对象类型选择任务节点、基线、工作空间和业务流程，则需要填写规则对象。输入监控对象的名称或者ID后，在列表中选择需要添加的对象，单击  图标。
	任务白名单	当对象类型为基线、工作空间、业务流程时，支持您输入节点名称/ID，添加至白名单列表中。白名单中的任务将不受监控。

分类	参数	描述
	资源组名称	如果对象类型选择独享调度资源组和独享数据集成资源组，则需要选择资源组名称。
	触发条件	如果对象类型选择任务节点、基线、工作空间和业务流程，此时触发条件取值如下： 完成 表示从实例任务运行的起始时间点开始监控，在任务运行成功时系统发送报警。 未完成 表示从实例任务运行的起始时间点开始监控，到指定的目标时间点任务仍未结束运行，则系统发送报警。例如，实例任务的定时调度时间为1点，设置的未完成时间为2点，则1点时该任务开始运行，在2点时任务仍未结束运行，则发送报警。 出错 表示从实例任务运行的起始时间点开始监控，如果任务运行出错，则系统发送报警。 实例任务运行出错即在运维中心 > 周期任务运维 > 周期实例的基本信息列，目标实例显示状态。 周期未完成 表示在指定的周期内，实例任务仍未结束运行，则系统发送报警。通常用于监控以小时为周期单位的实例任务。 例如，任务A每2小时调度一次，运行一次耗时25min。运行起始时间为每日0点0分，则该任务一天（24小时）共有12个任务周期，0点为第一个周期，2点为第二个周期，依次类推，22点为第12个周期。该任务正常运行时，会在每日0点25分、2点25分等时间节点执行完毕。如果在任意周期结束时间点该任务仍未结束运行，则发送报警。

分类	参数	描述  说明 周期未完成 可用于监控业务流程等任务。
触发方式		当业务流程设置了周期未完成监控后，系统会根据您设置的周期N，对业务流程中的节点任务（例如，天任务、小时任务、分钟任务等），进行第N个周期任务的监控。如果任务实例数少于N时，则会忽略该任务的告警。 例如，设置的周期为3，业务流程中有如下两个节点任务，则告警监控情况如下： ■ 小时任务A：每2小时调度一次，运行一次耗时25min。运行起始时间为每日0点0分，则该任务一天（24小时）共有12个任务周期，0点为第一个周期，则第3个周期为4点。该任务正常运行时，第3个周期任务会在4点25分执行完毕。如果在该周期结束时间点该任务仍未结束运行，则发送报警。 ■ 分钟任务B：每10分钟调度一次，运行一次耗时2min。运行起始时间为每日0点0分，则该任务一小时共有6个任务周期，0点0分为第一个周期，则第3个周期为0点20分。该任务正常运行时，第3个周期任务会在0点22分执行完毕。如果在该周期结束时间点该任务仍未结束运行，则发送报警。
	触发条件	如对象类型选择独享调度资源组和独享数据集成资源组，触发条件取值为： ○ 表示从实例任务运行的起始时间点开始监控，到指定的运行时 ○ 利用率该任务个数值并持续多长时间发送报警。通常用于监控实例任务的运行时间 ○ 例如：利用率大于50%并持续15分钟。 ○ 自动重跑后仍出错 ○ 等资源实例数大于某个数值并持续多长时间。 ○ 表示从实例任务运行的起始时间点开始监控，如果任务运行出错且自动重跑后仍出错，则系统发送报警，

分类	参数	描述
报警行为	报警方式	包括邮件、短信、电话钉钉群机器人和WebHook。您可以添加钉钉群机器人接收报警，请参见下文的操作，将报警消息发送到钉钉群。如果您需要多个钉钉群接收报警信息，请添加多个Webhook地址。 - 报警方式为钉钉群机器人、webhook时，您可单击操作列的发送测试消息，测试该报警是否可以正常发送。如未收到告警信息，请参考智能监控进行排查。 - 报警方式为短信、邮件、电话时，您可以单击校验联系方式，校验手机号/邮箱是否填写正确。  注意 - 您需要购买DataWorks专业版及以上版本，才可以使使用电话告警功能。 - 如果您选择报警方式为电话，则需要选中为了避免短时间内产生大量报警电话，DataWorks会对报警电话进行过滤，同一个用户在20分钟内最多接受到一通报警电话，其余报警电话将被降级为短信，请知悉。 - 仅支持钉钉Webhook地址。
	接收人	报警的对象，包括任务责任人、值班表和其他。
疲劳度控制	最大报警次数	报警的最大次数，超过设置的次数后，不再产生报警。
	最小报警间隔	两次报警之间的最小时间间隔。
	免打扰时间	在设置的时间段内不会发送报警。

6. 单击**确定**，即可生成规则。
- 您可以单击**规则管理**页面相应规则的操作列下的**详情**，查看规则的具体内容。

发送报警消息至钉钉群

1. 打开需要告警的目标钉钉群，单击右上角的**群设置**图标。
2. 单击**智能群助手**。
3. 在**智能群助手**页面，单击**添加机器人**。
4. 在**群机器人**页面，单击图标。
5. 在**选择要添加的机器人**页面，单击**自定义**。
6. 在**自定义对话框**中，单击**添加**。
7. 在**添加机器人对话框**中，配置各项参数。

参数	描述
机器人名字	自定义机器人的名称。
添加到群组	添加机器人的群组，不可以修改。
自定义关键词	设定后，只有包含关键词的消息内容才会被正常发送。此处必须添加DataWorks为关键词，注意区分大小写。 ? 说明 最多可以设置10个关键词，消息中至少包含其中1个关键词才可以发送成功。

8. 选中我已阅读并同意《自定义机器人服务及免责条款》，单击完成。
9. 完成安全设置后，复制机器人的Webhook，单击完成。

 注意 请妥善保管Webhook地址，一旦泄露会有安全风险。

10. 进入规则管理页面，单击新建自定义规则，选择报警方式为钉钉群机器人，复制webhook地址至钉钉群机器人的输入框中。

新建自定义规则

触发方式

触发条件：

选择触发条件

报警行为

如未收到告警信息，请参考文档进行排查。

报警方式：

☐ 短信

☐ 邮件

☐ 电话

☒ 钉钉群机器人

钉钉群机器人：

@所有人	Webhook地址	操作
☐		发送测试消息 删除

疲劳度控制

最大报警次数：

3

次

最小报警间隔：

30

分钟

免打扰时间：00:00 至

00:00

确认

取消

DataWorks智能监控推送企业微信或飞书的报警信息样例如下。



1.6.3.2. 值班表

您可以使用DataWorks值班表功能设置实例运维报警的值班机制。当自定义规则的报警接收人设置为值班表后，DataWorks将按照排班将报警信息发送给对应的值班人，以便值班人员及时发现并处理问题。本文为您介绍如何新建并管理值班表。

使用限制

仅DataWorks专业版及以上版本，才可以使用值班表功能。如果您使用的是低版本的DataWorks，则需要升级至专业版及以上版本，才可以使用该功能。升级DataWorks版本，详情请参见[DataWorks增值版本](#)。

注意事项

- 当前DataWorks版本到期后，您将无法访问值班表页面，不能查看、新增、修改或删除值班表及对应的排班。
- 如果当前专业版或以上版本的DataWorks配置了值班表，当版本到期后未及时续费或降级了当前版本，则DataWorks会为您保留原有的值班表数据。当重新升级到专业版或更高版本后，原有的值班表功能可以正常使用。
- 当您创建的自定义规则中，报警接收人配置了值班表后，在产生报警时值班表对应的值班人员才会收到报警信息。配置自定义规则，详情请参见[自定义规则](#)。

新建值班表并配置值班安排

- 进入运维中心。
 - 登录[DataWorks控制台](#)。
 - 在左侧导航栏，单击工作空间列表。
 - 选择工作空间所在地域后，单击相应工作空间后的进入运维中心。
- 在运维中心页面的左侧导航栏，单击智能监控 > 值班表。
- 在值班表区域，单击图标，配置新建值班表的各项参数。

新建值班表

名称:

test

管理员:

x

值班提醒:

前1天

9

提醒方式:

☐ 短信

☐ 邮件

☐ 电话

☒ 钉钉群机器人

@所有人	Webhook地址	操作
没有数据		

新增

确认

取消

参数	描述
名称	输入值班表的名称。
管理员	默认为当前登录阿里云的账号，您可以根据需求设置DataWorks工作空间的主账号或RAM用户为 管理员 。
值班提醒	开启值班提醒后，将提前通知值班人下次值班的时间，默认配置为不提醒。您可以根据需求选择前1天、前2天或前一周提醒。 当选择前1天、前2天或前一周时，则可以配置具体的提醒时间以及提醒方式。
提醒方式	DataWorks支持使用短信、邮件、电话及钉钉群机器人（Webhook地址）等方式进行值班提醒。添加钉钉群机器人获取Webhook地址详情请参见发送报警消息至钉钉群。 仅当值班提醒配置为前1天、前2天或前一周时，需要配置该参数。

4. 单击**确认**。
5. 在新建的值班表页面，单击右上角的**新建值班**，创建值班信息。
- 您也可以直接单击指定日期的白色块，新建值班信息。

←

2021-05

📅

→

回到本月

全部值班人

+

新建值班

名称: test

管理员: 韩琳

当前值班人: 韩琳

值班提醒: 前1天14:00

提醒方式: 短信

使用情况: [详情](#)

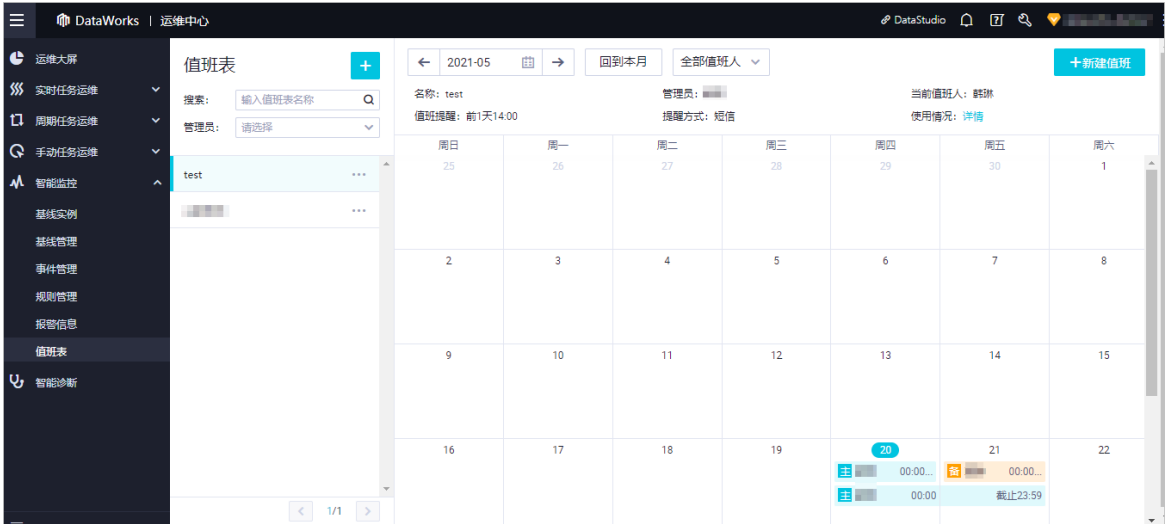
周日	周一	周二	周三	周四	周五	周六
9	10	11	12	13	14	15
16	17	18	19	20 主 00:00... 主 00:00 备 00:00... 截止23:59	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5

6. 配置新建值班对话框的各项参数。

参数	描述
值班人	当前值班表中配置的值班人员。您可以选择当前登录DataWorks工作空间的主账号或RAM用户。
类型	- 主：配置值班人为主值班人。 - 备：配置值班人为备用值班人。
全天	- 是：开始时间及结束时间，均为全天24小时值班。 - 否：可以自定义开始时间或结束时间的值班时间点。
开始时间	开始值班的时间。当全天配置为否时，可以自定义值班时间点。
结束时间	结束值班的时间。当全天配置为否时，可以自定义值班时间点。
重复	默认为不重复。您也可以选择重复的周期为每日、每周、每周一到周五、每月、每年以及自定义。 当配置了重复周期后，当前新增的值班信息会添加在所选的周期时间内。
时间间隔	时间间隔可以配置为当前值班时间区域前后的任意天数。 仅当重复配置为自定义时，需要配置该参数。

参数	描述
结束重复	用于指定重复周期的结束时间。默认为永不结束。 当重复配置为每日、每周、每周一到周五、每月、每年以及自定义时，需要配置该参数。

示例新建的值班信息如下图。



说明

- 新建的值班信息中，蓝色代表主值班人，黄色代表备用值班人。
- 新建值班时，您只能选择为当前日期及之后日期创建值班信息，当前日期之前的日期则无法创建值班信息。
- DataWorks支持在相同的值班时间配置多个主或备值班人。

管理值班表

- 在值班表区域，您可以执行如下操作：
 - 通过值班表名称或管理员，快速搜索符合条件的值班表。
 - 单击目标值班表后的图标，可以修改或删除已创建的值班表。

说明 仅值班表管理员可以编辑或删除值班表或已创建的值班安排。

- 双击目标值班表，则可进入值班表的详情页面，查看值班表的详细信息。



- 在目标值班表详情页面，您可以执行如下操作：
 - 在①区域，您可以查看值班表的基本信息以及当前的值班人。当您配置的自定义监控规则使用了目标值班表，则该值班表会按照排班情况发送相关报警信息，您可以单击使用情况后的详情查看具体使用信息。管理自定义规则，详情请参见[自定义规则](#)。
 - 在②区域，您可以通过值班人类型，筛选目标值班人的值班安排，也可以查看指定月份的值班安排。默认显示当前月份所有值班人的值班安排。

如果您选择了查看非当前月份的值班，在查看完毕后可以单击[回到本月](#)，快速返回至当前月份的值班界面。

- 在值班安排界面，您可以查看已创建值班的值班详情，包括值班人、值班日期、值班类型（属于主值班或备值班）等。单击已创建的值班，则可选择修改或删除当前值班。



1.6.4. 自动运维

您可以根据业务需求对独享资源组自定义运维规则，实现独享资源组的智能监控和自动化运维。本文为您介绍如何通过DataWorks的运维中心配置自动运维规则。

背景信息

DataWorks的自动运维功能可以实现根据已设置好的监控规则和运维规则来自动执行任务的运维操作。您可以根据具体业务需求，对运行在目标独享资源组上的实例任务定制监控指标，再结合自身业务逻辑自定义运维规则，并关联目标实例任务的监控规则。当实例任务满足过滤条件时，将自动触发并执行运维动作，以达到自动运维的目的。

使用限制

- 仅支持项目管理员设置、修改和删除自动运维规则。
- 仅支持对独享资源组上运行的实例任务配置自动运维规则。
- 目前仅支持查看30天内的自动运维规则执行记录。
- 目前仅支持当调度资源组使用率水位过高时终止运行指定任务的自动运维操作。
- 一条自动运维规则仅支持关联一条监控规则，但多条自动运维规则支持关联同一条监控规则。

进入自动运维

1. 进入数据开发页面。
 - i. 登录DataWorks控制台。
 - ii. 在左侧导航栏，单击工作空间列表。
 - iii. 选择工作空间所在地域后，单击相应工作空间后的进入数据开发。
2. 单击左上方的图标，选择全部产品 > 任务运维 > 运维中心（工作流）。
3. 在左侧导航栏，选择智能监控 > 自动运维，进入自动运维页面。

运维规则管理

自动运维页面可以为您展示已创建的运维规则列表及规则执行记录信息，您可以根据需求，执行如下操作：

 **说明** 仅支持项目管理员权限设置和修改规则，如需添加项目管理员权限，请参见[添加工作空间成员和角色](#)。

规则管理		执行记录			
search					
名称	关联监控规则	生效工作空间	最后修改时间	修改人	操作
xml_res_test4	xml_res_test2	springboot_标准模式2	2021-06-01 15:55:37	dataworks_3h1_1	查看 删除
wzp_test	xml_res_test2	springboot_标准模式2	2021-06-01 15:58:47	dataworks_3h1_1	查看 删除
test001	xml_res_test2	springboot_标准模式2	2021-06-03 10:42:59	dataworks_3h1_1	查看 删除

- 添加规则

规则管理页面，单击右上方的添加规则，按照要求填写如下信息：

新建规则

当关联的监控规则被触发，系统将对满足过滤条件的实例自动执行运维动作。目前只支持关联调度资源监控规则。

*

名称:

请输入规则名称

*

关联监控规则:

关联监控规则

去创建监控规则

*

运维动作:

请选择

过滤条件

*

资源组:

资源组

*

工作空间:

请选择

*

实例类型:

请选择

*

优先级:

请选择

*

状态:

请选择

白名单

序号	任务名称	责任人	工作空间	操作
没有数据				

请选择

添加

约束规则

*

生效时段:

请选择时间

至

请选择时间

*

最大生效次数:

次

*

最小生效间隔:

分钟

确定

取消

类别	参数	说明
基本信息	名称	输入新建的运维规则名称。
	关联监控规则	关联需要使用的监控规则。如需新建监控规则，请参见自定义规则。 ? 说明 目前仅支持关联调度资源监控规则。

类别	参数	说明
	运维动作	选择达到监控规则条件触发监控规则时，实例需要执行的运维操作。  说明 目前运维动作仅支持选择终止运行实例。
过滤条件	资源组	创建监控规则时设置的资源组名称。  说明 此处监控对象仅支持选择独享数据集成资源组和独享调度资源组。
	工作空间	选择该运维规则生效的工作空间。
	实例类型	选择使用该运维规则的实例类型。包括周期实例、补数据实例、测试实例、手动任务和手动业务流程
	调度周期	调度周期包括分钟、小时、天、周和月。
	优先级	选择运维规则执行优先级，数值越大优先级越高。取值如下： 1 3 5 7 8
	状态	选择实例运行状态。包括等待资源和运行中。
白名单	白名单	展示该规则可使用任务白名单，如需添加任务白名单，请在下拉列表选择需要使用的白名单，单击添加。  说明 自动运维规则对白名单任务不生效。
约束规则	生效时段	设置规则生效的起始时间和结束时间。
	最大生效次数	设置规则触发的最大次数。
	最小生效间隔	设置规则触发的最小时间间隔。

- 搜索规则

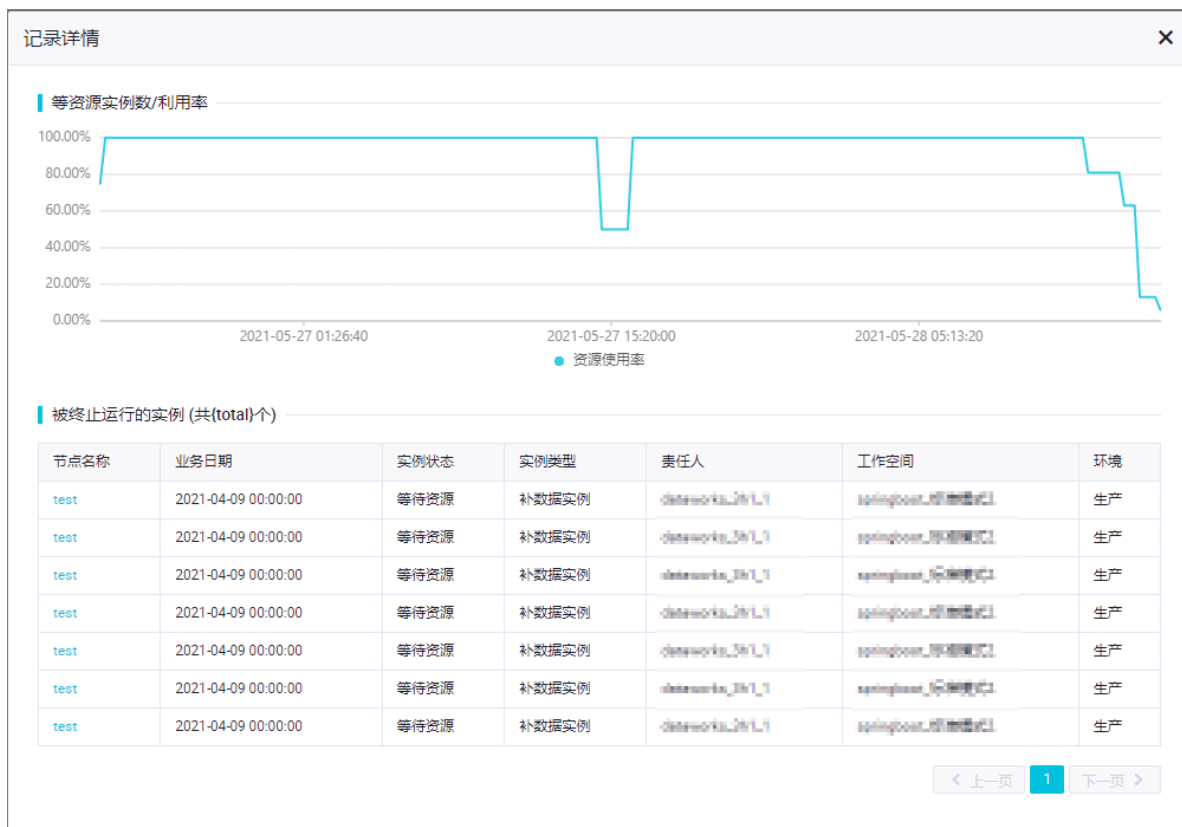
在规则管理页面左上方的搜索栏，可以通过输入规则名称进行模糊匹配来快速查找目标规则。

- 查看、修改和删除规则

- 如需查看规则信息，请在规则管理页面，单击目标规则所在行右侧的查看。
- 如需修改规则定义，请单击查看规则对话框右下方的修改。
- 如需删除目标规则，请单击目标所在行右侧的删除，在弹出的对话框单击确认。

● 查看规则执行记录

在执行记录页面，为您展示运维规则的执行情况，包括执行时间、规则责任人及涉及任务数等，如需查看详细信息，请单击目标执行记录右侧的查看详情，即可查看规则自生效时刻起24小时内的资源使用率水位曲线信息。



- 等资源实例数/利用率：以水位曲线形式为您展示等资源实例数和利用率情况，您还可以将鼠标悬停在水位线显示区域内的任意位置，即可显示此刻资源实例数和利用率情况。
- 被终止运行的实例：为您展示被终止运行的任务列表。

其他

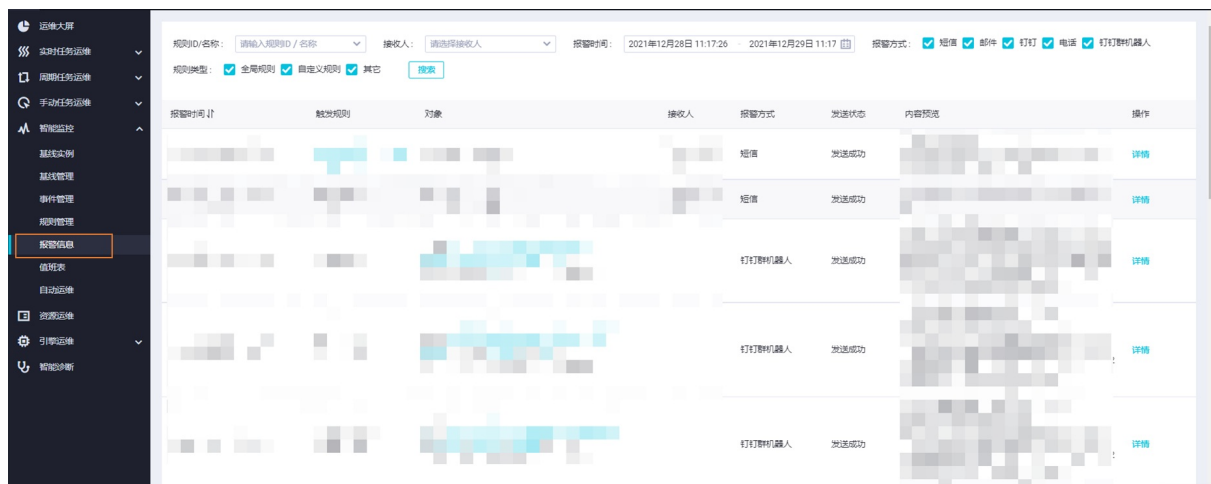
现在您已经学习了如何配置运维规则，配置好运维规则后，系统会自动根据运维规则对资源组使用情况进行监控，资源组监控详情请参见[资源运维](#)。

1.6.5. 报警信息

您可以在智能监控模块查看所有的报警信息。

您可以进入运维中心 > 智能监控 > 报警信息页面，通过规则ID/名称、接收人、报警时间、报警方式和规则类型等信息进行搜索。

② 说明 仅支持查看最近30天内的报警记录。



您可以查看相应报警的报警方式、发送状态等信息。单击操作栏中的详情，即可查看报警的详细内容。

1.6.6. 常见问题

本文为您介绍在运维中心设置了智能监控后，任务运行失败的常见报错及解决方案。

没有收到任务运行失败的告警短信/邮件

- 问题现象：在运维中心设置了智能监控后，任务运行失败了却没有收到告警的短信或邮件。
- 可能原因：

出现上述问题通常是因为设置的报警接收人的手机号码和邮箱信息不正确。

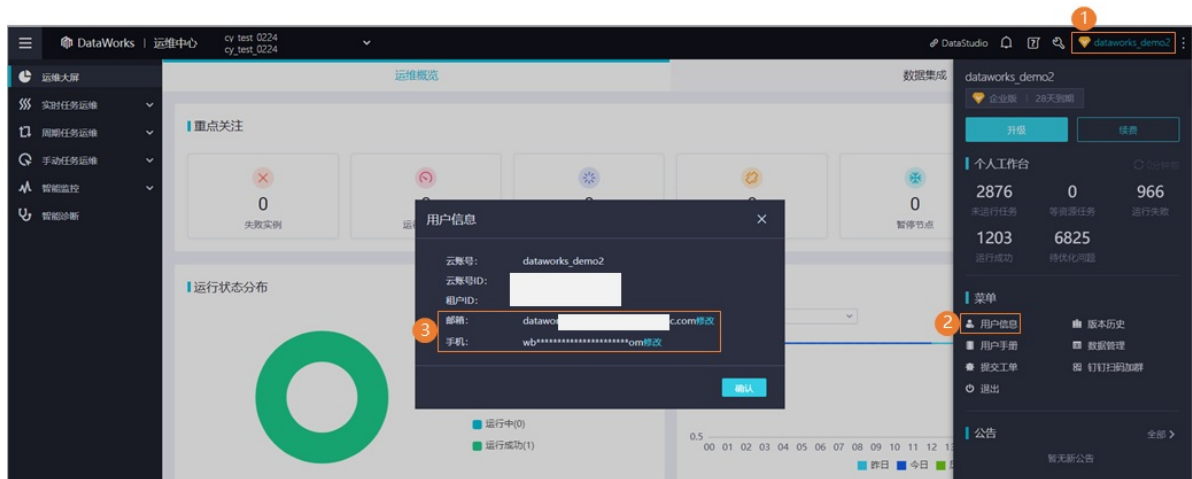
当DataWorks任务运行失败时，系统会向DataWorks智能监控中设置的报警接收人的手机号码和邮箱发送报警信息，如果手机号码和邮箱信息不正确，则系统会将报警信息发送到阿里云账号[基本接收管理](#)中设置的产品欠费、停服、即将释放等相关信息的信息接收人。此种情况下DataWorks的报警接收人将收不到报警信息。

- 解决方案：

您需要确认并修改DataWorks智能监控中设置的报警接收人的手机号码和邮箱信息，确保系统可以向其发送报警消息。

您可以单击右上角的账号名称，在用户信息中查看或修改邮箱和手机号码信息。

说明 RAM账号的邮箱和手机号信息需要对应的主账号进行修改。



1.7. 资源运维

资源运维功能用于监控DataWorks中任务运行所用资源组的使用情况，本文为您介绍如何使用资源运维功能。

背景信息

您可以根据具体业务需求对资源组上运行的实例任务定制监控指标和自动运维规则，资源运维页面会以可视化方式为您展示资源组使用情况及实例任务执行情况，实现资源组及实例任务的智能监控和自动化运维，减少繁杂的人工操作，提高运维管理效率。

使用限制

- 当前华东2（上海）、华东1（杭州）地域不支持使用DataWorks的资源运维功能。
- DataWorks的资源运维功能仅支持对独享调度资源组和独享数据集成资源组进行运维。

进入资源运维

1. 登录DataWorks控制台。
2. 在左侧导航栏，单击工作空间列表。
3. 选择工作空间所在地域后，单击相应工作空间后的进入数据开发。
4. 单击左上方的☰图标，选择全部产品 > 任务运维 > 运维中心（工作流）。
5. 在左侧导航栏，选择资源运维，默认显示当前DataWorks使用的独享资源组列表。

查看资源组列表

资源运维页面为您展示当前Dat aWorks任务运行正在使用的资源组（即运行状态的资源组），此处仅支持展示独享调度资源组和独享数据集成资源组。您可以根据需求，执行如下操作：

任务运维

实时任务运维

周期任务运维

手动任务运维

智能运维

资源运维

引擎运维

智能诊断

独享资源组

查看全部

资源组名称	类型	状态	到期时间	规格	数量	使用率
HZ_ZYJ_DL_DUXIANG	数据集成资源组	运行中	2021-07-10 00:00:00	24c48g	1	0.00%
HZ_ZYJ_DUXIANG	调度资源组	运行中	2021-07-10 00:00:00	24c48g	1	0.00%
cx_schedul	调度资源组	运行中	2021-07-09 00:00:00	8c16g	1	0.00%
zrm_cdi_060701	调度资源组	运行中	2021-07-08 00:00:00	16c32g	1	0.00%
dl_dzyx_d1	调度资源组	运行中	2021-06-19 00:00:00	8c16g	2	0.00%
cx_resource_new	数据集成资源组	运行中	2021-07-08 00:00:00	16c32g	1	0.00%
dagfood0604002	调度资源组	运行中	2021-07-05 00:00:00	8c16g	2	0.00%
dagfood0602001	调度资源组	运行中	2021-07-03 00:00:00	16c32g	1	0.00%
test_ujh	调度资源组	运行中	2021-06-28 00:00:00	12c24g	1	0.00%
dagfood0520085CH	调度资源组	运行中	2021-06-26 00:00:00	16c32g	1	0.00%
dagfood0520005	调度资源组	运行中	2021-06-26 00:00:00	8c16g	1	0.00%
dagfood0520004	调度资源组	运行中	2021-06-26 00:00:00	16c32g	1	0.00%
DIIDE_0b756353cde4f996ad55c844040b4	数据集成资源组	运行中	2021-06-20 00:00:00	4c8g	1	0.00%
es_test	数据集成资源组	运行中	2021-06-19 00:00:00	4c8g	2	0.00%
DIIDE_86c4b2f71c6d4cafa7b787944259f6e	数据集成资源组	运行中	2021-06-19 00:00:00	8c16g	2	0.00%
dagfood05140002	调度资源组	运行中	2021-06-15 00:00:00	12c24g	1	0.00%
dagfood05140001	调度资源组	运行中	2021-06-15 00:00:00	4c8g	1	0.00%
xm_pre_d1_test	数据集成资源组	运行中	2021-06-13 00:00:00	4c8g	1	0.00%
xm_pre_gateway2	调度资源组	运行中	2021-06-12 00:00:00	4c8g	1	0.00%
xm_pre_gateway	调度资源组	运行中	2021-06-12 00:00:00	4c8g	1	0.00%
DIIDE_c1e1a246fbc442c3a48c944557215c44	数据集成资源组	运行中	2021-07-09 00:00:00	4c8g	1	0.00%
DIIDE_3a23d410cbcd4778a31e855ea049f963b	数据集成资源组	运行中	2021-07-06 00:00:00	8c16g	2	0.00%
moln_procd_d1	数据集成资源组	运行中	2021-09-09 00:00:00	12c24g	2	0.00%

- 筛选目标资源组。

您可以在右上方的搜索栏输入所要查找的目标资源组名称，单击，即可进行模糊匹配来快速查找目标资源组信息。

- 查看其他资源组。

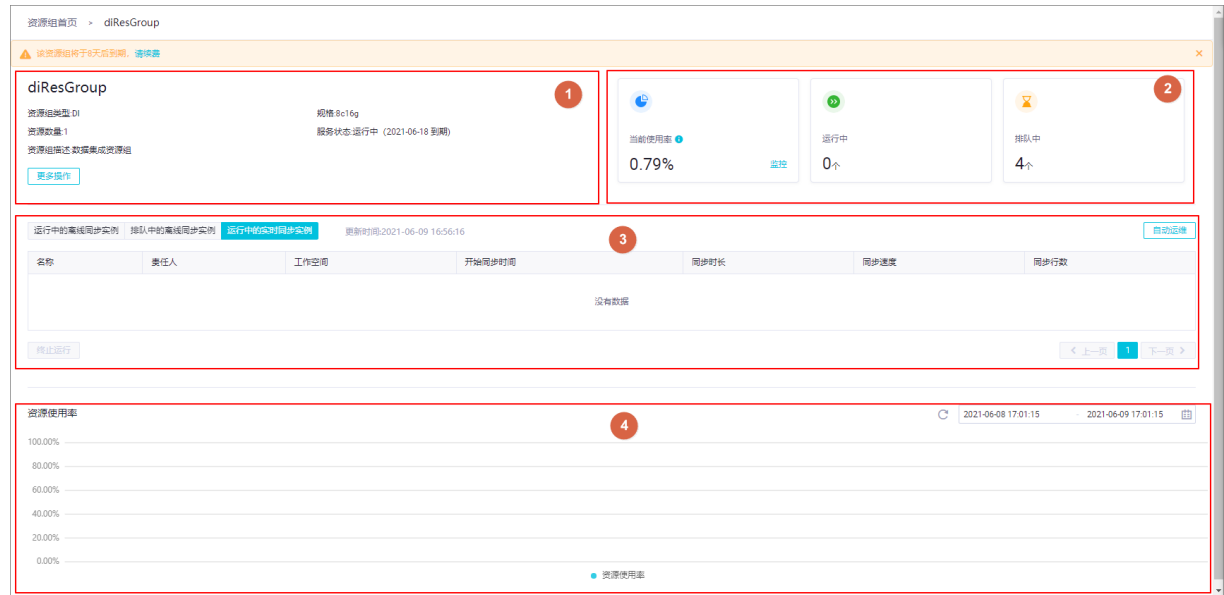
您可以单击独享资源组后的查看全部，即可跳转至Dat aWorks管理控制台页面下的资源组列表。

参数名	说明
资源组名称	创建资源组时自定义的资源组名称。单击资源组名称，可以跳转至资源组详情页面，具体介绍请参见 查看资源组详情 。
类型	创建资源组时选择的资源组类型，此处仅支持数据集成资源组和调度资源组。  说明 资源运维功能仅支持独享资源组。
状态	当前资源组的运行状态，包括： - 运行中 - 已到期 - 已释放 - 启动中 - 创建失败 - 更新中 - 更新失败 - 已释放 - 超时
到期时间	当前资源组可用的截止日期。

参数名	说明
规格	创建资源组时定义的资源组规格。
数量	创建资源时定义的资源数量。
使用率	资源组当前使用率，单击使用率后的  ，可以对整个资源组列表按照资源使用率从低到高或从高到低进行排序。

查看资源组详情

资源组详情页面为您展示目标资源组的具体使用情况，包括资源组的基本信息、资源使用率水位曲线和任务列表，并支持自定义监控规则，帮助您实现资源的智能监控和自动运维，查看和设置运维规则请参见[自动运维](#)。



区域	参数	说明
1	该区域为您展示资源组的基本信息。	如果您需要修改资源组信息，例如对资源组进行扩容、缩容等操作，您可以单击更多操作，即可跳转至DataWorks管理控制台的资源组列表。
2	该区域为您展示该资源组的使用情况，包括当前使用率、当前使用该资源组且处于运行中状态的实例数和等资源且处于排队中 状态的实例个数。	如果您需要查看为该资源组配置的监控规则，可以单击监控，即可跳转至运维中心>智能监控>规则管理。

区域	参数	说明
3	可以为您分类展示当前资源组上运行的所有实例列表。根据运行状态和同步方式进行实例分类，具体如下： - 运行中的离线同步实例 - 排队中的离线同步实例 - 运行中的实时同步实例	您可以单击自动运维，即可跳转至运维中心>智能监控>自动运维，有关自动运维功能具体介绍请参见自动运维。
4	以水位曲线形式为您直观展示当前资源组资源使用率。	你可以通过设置起始时间来筛选需要查看资源使用率情况的时间段。 您还可以将鼠标悬停至水位曲线上，查看此刻资源利用率。

1.8. 引擎运维

在DataWorks实例任务中，E-MapReduce计算引擎任务会被分为若干作业有序执行。您可以使用DataWorks的引擎运维功能，查看各个E-MapReduce作业的详细信息，及时查找并清理运行有误的作业，避免该类作业阻塞下游任务，影响实例任务的正常运行。

使用限制

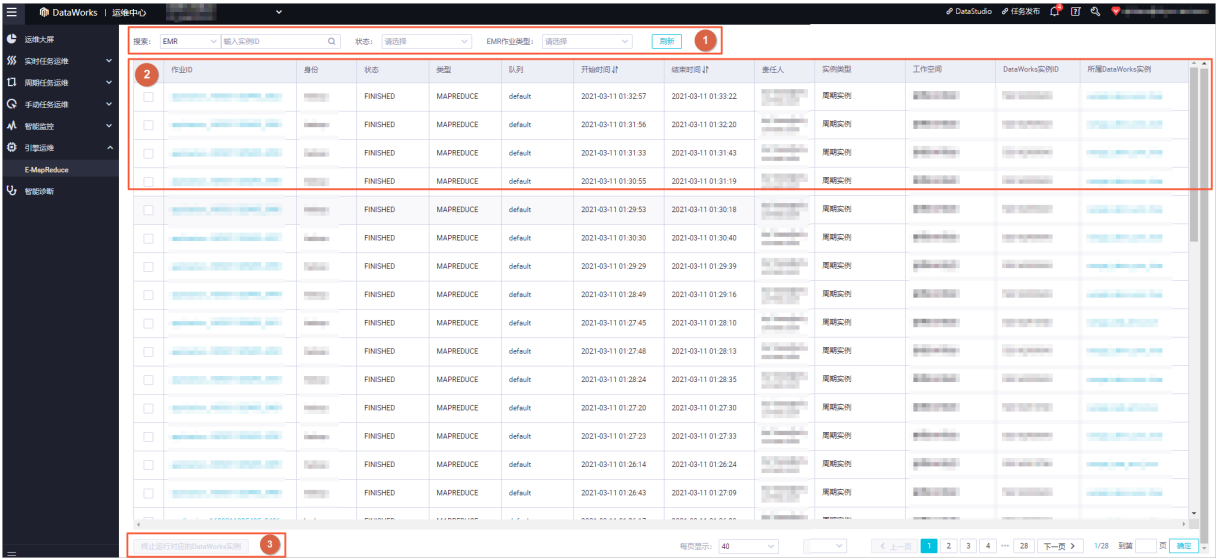
- DataWorks目前仅支持对E-MapReduce计算引擎进行运维，并且您需要升级E-MapReduce执行包的版本后才能正常获取运维数据。如果您当前需要使用引擎运维功能请[提交工单](#)升级E-MapReduce执行包的版本。
- 如果您的DataWorks项目空间没有添加E-MapReduce引擎，则[运维中心](#)左侧导航栏不会显示[引擎运维](#)。
- 目前仅华南1（深圳）地域支持使用DataWorks的引擎运维功能。

进入引擎运维

1. 进入数据开发页面。
 - i. 登录[DataWorks控制台](#)。
 - ii. 在左侧导航栏，单击[工作空间列表](#)。
 - iii. 选择工作空间所在地域后，单击相应工作空间后的[进入数据开发](#)。
2. 单击左上方的图标，选择[全部产品](#) > [任务运维](#) > [运维中心（工作流）](#)。
3. 在左侧导航栏，选择[引擎运维](#) > [E-MapReduce](#)，进入E-MapReduce引擎运维页面。

查看引擎作业

E-MapReduce引擎运维页面，为您展示了当前地域，与所选DataWorks项目空间绑定的E-MapReduce引擎，在DataWorks中下发的E-MapReduce作业的详细信息，您可以根据需求执行如下操作：



- 筛选目标E-MapReduce作业。

您可以在E-MapReduce引擎运维页面的顶部菜单栏，根据不同维度进行筛选，快速查看目标作业。

类别	描述
实例ID	您可以通过E-MapReduce的作业ID查找目标作业，或通过DataWorks的实例ID，查找该实例任务下的所有EMR作业。 说明 - 引擎运维页面默认为您展示最近三天的引擎实例数据。 - 通过E-MapReduce的作业ID或DataWorks的实例ID搜索时，仅支持展示最近30天的引擎实例数据。

类别	描述
状态	◦ NEW：表示E-MapReduce作业刚创建。 ◦ NEW_SAVING：表示正在保存E-MapReduce作业。 ◦ SUBMITTED：表示E-MapReduce作业已提交执行申请。 ◦ ACCEPTED：表示调度程序已接受E-MapReduce作业的执行申请。 ◦ RUNNING：表示E-MapReduce作业正在运行。  说明 如果目标作业长期处于运行状态，您可以手动终止运行目标作业所在的DataWorks实例任务，防止运行有误的作业长期占用资源，阻塞下游任务，详细内容请参见下文终止运行DataWorks实例任务。 ◦ FINISHED：表示E-MapReduce作业已运行结束。 ◦ SUCCEEDED：表示E-MapReduce作业运行成功。 ◦ FAILED：表示E-MapReduce作业运行失败。如果作业呈现该运行状态，则您需要及时识别并处理相关报错，以免阻塞下游作业，影响任务的正常运行。 ◦ KILLED：表示由作业的执行者或管理员终止了E-MapReduce作业的运行。
EMR作业类型	当前仅支持查看 MAPREDUCE 及 SPARK 类型的作业。

- 查看目标作业的详细信息。

- 您可以查看E-MapReduce作业的基本信息，包括作业ID、作业状态、开始运行时间、结束运行时间、所属DataWorks实例及实例ID等。

 **说明** 不同E-MapReduce作业所属的DataWorks实例名称可能相同，但如果不同E-MapReduce作业的运行开始时间不同时，则被认为属于不同的DataWorks实例任务。您可以通过作业所属的DataWorks实例ID判断任务是否属于同一个DataWorks实例任务。

- 通过作业的运行状态，识别运行出错的作业（例如，状态为**FAILED**的作业为运行失败的作业），单击**作业ID**或**所属DataWorks实例名称**，进入相应任务详情界面，确认并处理报错。
- 您可以对**开始时间**或**结束时间**，进行升序或降序排列，直观的查看作业的运行先后顺序及运行时长，了解目标作业的运行情况。

- 终止运行DataWorks实例任务。

如果E-MapReduce作业长期处于运行状态，即作业状态显示为**RUNNING**，则您可以勾选目标作业，单击**引擎运维**页面左下角的**终止运行**对应的**DataWorks实例**，终止运行目标作业所在的DataWorks实例任务。该类作业往往是由于运行内部出错，导致作业不能自动终止，防止该作业长时间占用资源，阻塞并影响其他作业的运行，则需要手动将其停止，及时处理相关报错。

注意

- 当多个E-MapReduce作业属于同一个DataWorks实例任务时，当终止运行任意E-MapReduce作业时，都会导致DataWorks实例任务的运行状态置为失败。
- 当前仅支持对运行中的DataWorks实例任务进行终止运行操作。
- 执行终止运行操作后，目标实例任务的状态将被置为失败。置为失败状态后，当前实例任务会阻塞该任务的下游节点运行。请知悉风险并谨慎操作。

1.9. 常见问题

1.9.1. 数据未正常产出或任务未运行如何处理

问题现象

在数据地图查看数据时，发现当天数据未正常产出，或者在运维中心的周期实例发现任务节点状态为灰色的未运行状态，如下图所示。



可能原因

未运行节点的上游节点未成功运行，上游节点实例可能被冻结，或还在等待运行。上游节点未成功运行，则下游节点不会启动运行。

解决方案

您需要点开任务DAG图，选中实例后右键查看父节点，查看上游节点实例状态。

- 上游节点实例状态：紫色

说明上游节点实例被冻结了，上游任务将不会运行，且阻塞下游任务。此种情况下，您可以在实例详情的操作记录中查看相关记录，了解上游节点冻结的操作记录。

- 上游节点实例状态：黄色

说明上游节点在等待运行中：

- 上游节点可能设置了定时运行，但是还未到定时时间，上游节点当前在等待任务定时时间。此种情况下您可以单击右下角的[查看实例详情](#)，看上游节点的定时时间，待到了定时时间后上游节点即开始运行，上游节点运行完成后下游节点即启动运行。
- 上游节点可能在等待调度资源，您可在运行日志中查看是否正在等待调度资源。若出现此种情况，说明当前项目下正在运行的任务占用的调度资源已达到可使用的资源上限，您可重新安排任务运行时间将各任务错峰运行，或购买独享调度资源组，独享调度资源组详情可参见[独享调度资源组](#)。
- 上游节点实例状态：灰色
说明上游节点也未运行，您需继续查看上游节点的父节点状态。

1.9.2. 找不到任务实例如何处理

问题现象

提交发布任务后，在运维中心的实例列表页面，找不到任务对应的实例。

可能原因

找不到任务对应的实例则说明实例未成功创建。DataWorks会根据任务节点中的调度配置中**生成实例方式**和**调度依赖**的配置结果，使用不同方式为您创建运行节点任务的实例。实例未成功创建的可能原因如下。

- 节点成为孤立节点时（即节点的**调度依赖**中，没有配置任何**依赖的上游节点**），则无法正常创建任务实例，在实例列表中找不到对应的实例。
此类场景通常发生在创建好节点和业务流程后，后期修改了节点的依赖配置，删除了节点的**依赖的上游节点**配置结果后提交发布，导致节点成为孤立节点，无法生成对应的实例。
- 生成实例方式配置为**T+1次日生成时**：
 - 如果您提交发布的时间早于23:30，则DataWorks会在当天的23:30为您创建实例，提交发布后的第二天可以在运维中心的实例列表页面找到对应实例。
 - 如果您提交发布的时间晚于23:30，则DataWorks会在提交发布后的第二天的23:30为您创建实例，第三天可以在运维中心的实例列表页面找到对应实例。

如果您在实例创建成功前在运维中心的实例列表中查看实例，则找不到对应的实例。

- 生成实例方式配置为**发布后即时生成时**：

❓ 说明 旧版工作流不支持将生成实例方式配置为发布后即时生成。

- 提交发布的时间晚于23:30时，则DataWorks不会创建实例，此种情况下实例无法成功创建，您在实例列表中找不到对应的实例。
- 提交发布的时间早于23:30时，则DataWorks会即刻为您创建任务实例，您可以在实例列表个找到对应的实例。
 - 任务定时运行时间在提交发布时间的十分钟之后（例如提交发布时间是18:00，定时运行时间是18:30），则DataWorks会正常创建实例运行任务，您可以在实例列表中找到对应的实例。
 - 任务定时运行时间距离提交发布时间不足十分钟（例如提交发布时间是18:00，定时运行时间是18:05），则DataWorks会生成一个已完成的实例，节点实例属性为实时生成的过期实例。

此种场景下，如果您提交发布的时间晚于23:30，则找不到对应的实例。但是需要注意找得到实例时，实例是否是过期实例。

解决方案

1. 检查节点的调度依赖配置，确保节点不是孤立节点。

在数据开发页面的调度依赖配置中，查看各个节点的依赖的上游节点是否有配置内容，确保节点不是孤立节点。

2. 检查节点的提交发布时间，在23：30之前提交发布节点。

在发布中心的发布包列表页面，查看发布时间，如果发布时间晚于23：30，您可以在第二天的23：30之前重新提交发布节点。

② 说明 如果节点的生成实例方式配置为发布后即时生成，您还可以再检查一下任务定时运行时间与提交发布的时间间隔是否超过十分钟，如果没有的话，则需要重新提交发布，确保定时运行时间在提交发布时间的十分钟之后，否则生成的实例为过期实例，无法正常运行节点任务。

2.任务发布

2.1. 发布任务

本文为您介绍如何发布标准模式工作空间的节点，以及如何通过跨项目克隆发布简单模式工作空间的节点。

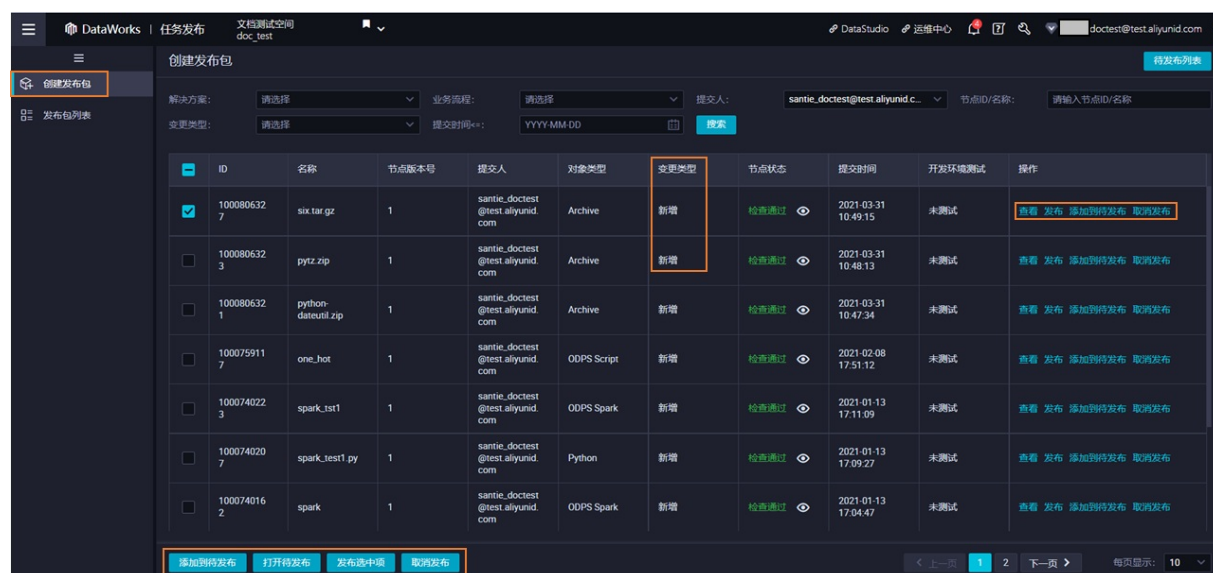
背景信息

在严谨的数据开发流程下，开发者通常会在用于开发的项目内，完成代码开发、流程调试、依赖属性和周期调度属性配置后，再提交任务至生产环境调度运行。

DataWorks的标准模式为您提供在一个项目内，完成从开发到生产的全链路能力，建议您通过该模式完成数据开发与生产发布。详情请参见[简单模式和标准模式的区别](#)。

标准模式工作空间下，提交的节点会默认添加至创建发布包页面，该页面为您展示已提交的新增、更新、下线节点、资源和函数等操作。

在创建发布包页面发布的任务会生成发布包，您可以在发布包列表页面查看相关节点的发布记录和状态。



如果您在数据开发页面新增、更新、删除的节点、资源和函数，同样在生产环境生效。您需要在创建发布包页面发布相关操作至生产环境。您可以在创建发布包页面添加单个或多个节点至待发布列表，进行批量发布。

创建发布包页面支持修改每页显示的条数。

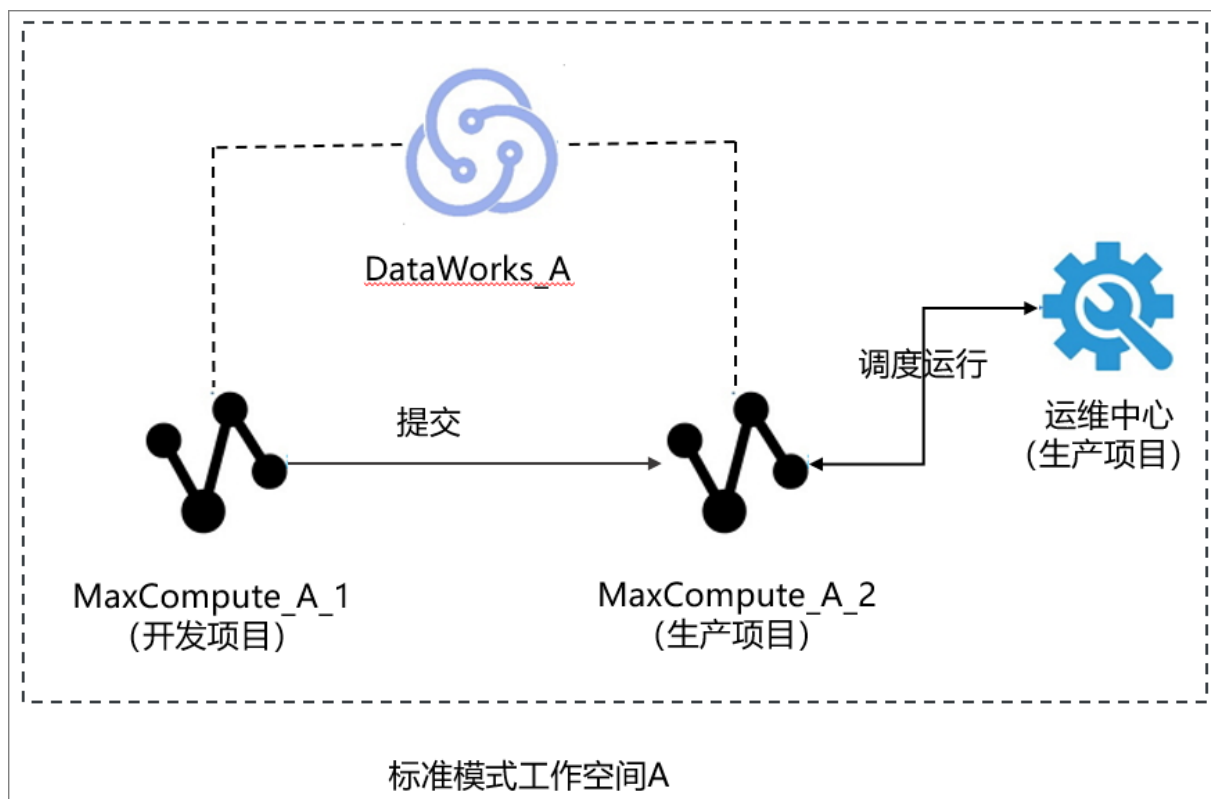
单击相应节点后的查看，即可查看当前版本的代码内容变更。

不同实例生成方式对实例生效时间的影响如下：

- T+1次生成实例的节点：在23:30前发布变更操作，周期节点运维在第二天生效。
- 发布后即时生成实例的节点：如果是新增的节点，定时时间在发布时间点十分钟后的实例会正常转出。如果是修改的节点，定时时间在发布时间点十分钟后的实例，会根据最新的调度配置替换修改操作之前的实例。详情请参见[实例生成方式：发布后即时生成实例](#)。
- 当天23:30后发布新增或修改的节点，会在第三天生效。
- 当天23:30后发布的即时生成的实例，不会生效。

发布标准模式工作空间的节点

如果您使用的是标准模式的工作空间，系统默认一个DataWorks工作空间对应两个相互绑定的MaxCompute项目（开发环境与生产环境），您可以直接从开发环境提交并发布任务至生产环境。



1. 进入数据开发页面。

- i. 登录DataWorks控制台。
- ii. 在左侧导航栏，单击工作空间列表。
- iii. 单击相应工作空间后的进入数据开发。

2. 提交节点。

在标准模式的工作空间，仅开发角色可以提交节点。

- i. 双击打开已配置完成的业务流程，单击工具栏中的图标。
- ii. 在提交对话框，选择需要提交的节点名称，输入备注，并选中忽略输入输出不一致的告警。

 **说明** 如果您的节点已经被提交，且没有修改节点内容，只是修改了业务流程或节点属性。您可以不选择节点，输入备注后直接提交业务流程。相关改动会正常被提交。

如果节点已经被提交过，在不改变节点内容的情况下，无法再次选择该节点。

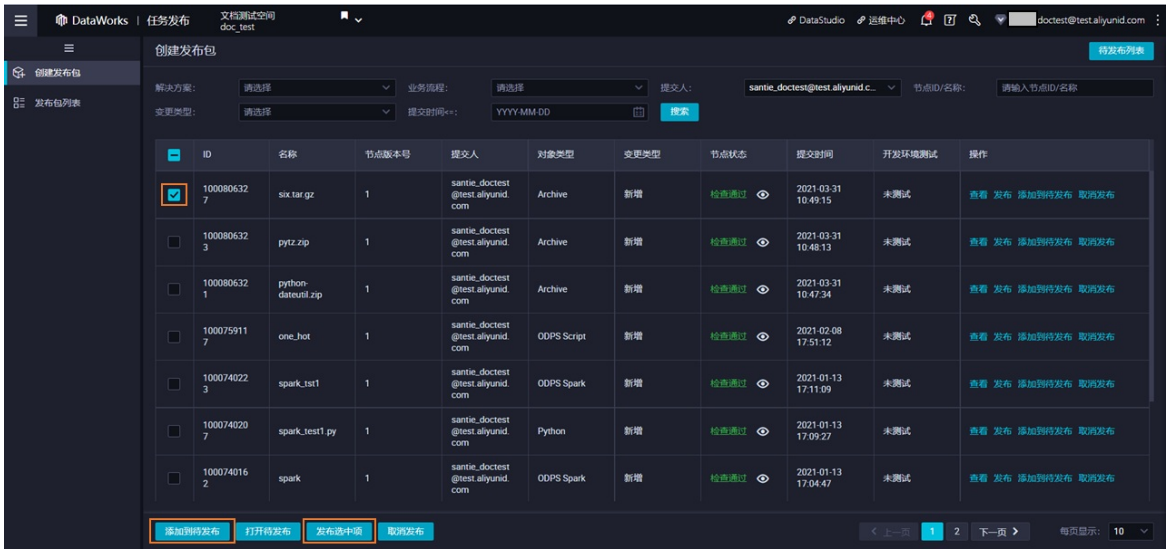
iii. 单击提交。

3. 提交成功后，单击右上角的任务发布。

在标准模式的工作空间，仅运维、部署及管理员角色可以发布节点任务。

4. 在创建发布包页面，批量选中需要发布的节点，单击添加到待发布。

您可以根据提交人、节点ID、节点类型和变更类型等条件过滤和搜索任务。如果单击发布选中项，可以立即发布至生产环境调度运行。



5. 单击打开待发布，确认待发布列表中的信息无误后，单击全部打包发布，即可发布列表中的节点至生产环境。

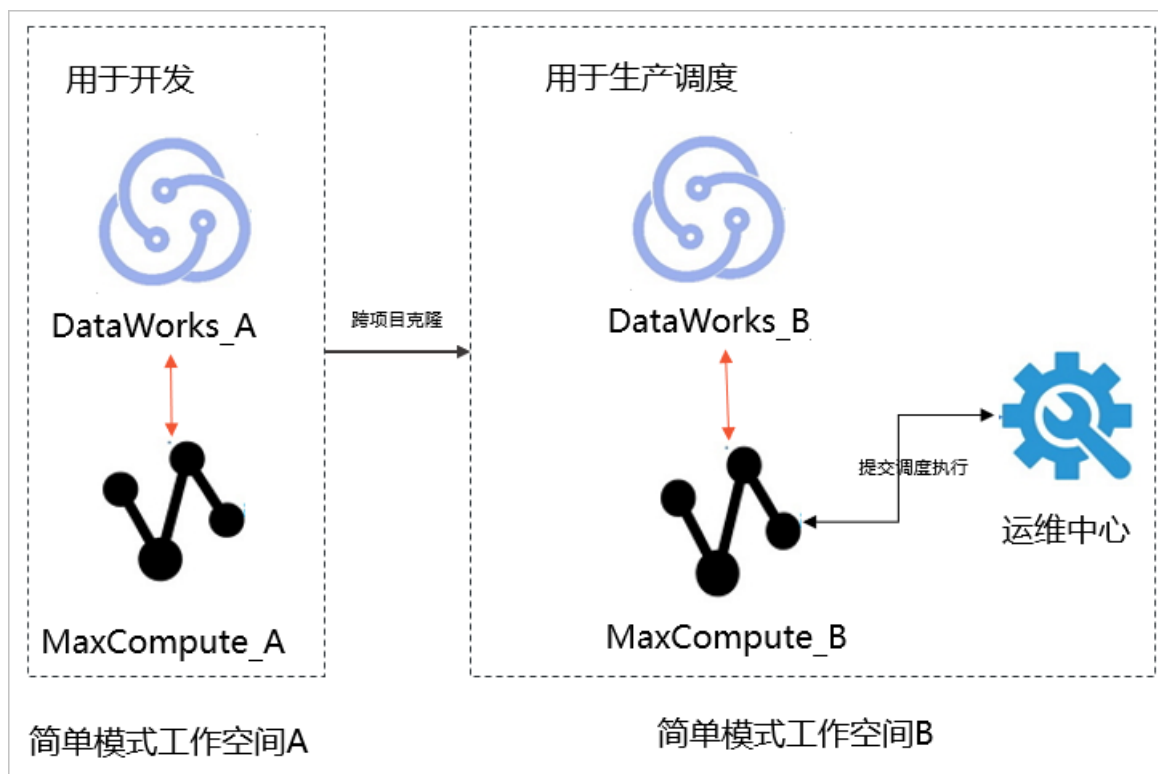


① 说明 简单模式的工作空间严禁直接操作生产环境内的表数据。您可以通过标准模式的工作空间，获得始终稳定、安全、可靠的生产环境。因此，建议您使用标准模式工作空间进行任务的发布与调度。

发布简单模式工作空间的节点

您可以克隆并提交任务至用于生产的工作空间，即通过简单模式工作空间（用于开发）结合简单模式工作空间（用于生产），实现简单模式工作空间内开发环境和生产环境隔离。

例如，您创建两个简单模式的工作空间，分别用于开发和生产。您可以先使用跨项目克隆功能，克隆A工作空间中的任务至B工作空间，再在B工作空间内提交克隆的任务至调度引擎进行调度。



说明

- 权限要求：除项目管理员外，执行操作的子账号需要具有运维角色的权限（创建克隆包、发布克隆任务），才能独立完成该流程。
- 支持的工作空间类型：仅简单模式工作空间支持克隆任务至其它工作空间。
- 准备工作：创建简单模式的源工作空间A和标准模式的目标工作空间B。

1. 进入数据开发页面。

- i. 登录DataWorks控制台。
- ii. 在左侧导航栏，单击工作空间列表。
- iii. 单击相应工作空间后的进入数据开发。

2. 提交节点。

- i. 双击打开已配置完成的业务流程，单击工具栏中的图标。
- ii. 在提交对话框，选择需要提交的节点名称，输入备注，并选中忽略输入输出不一致的告警。
- iii. 单击提交。

3. 单击页面右上角的跨项目克隆。

4. 在创建克隆包页面，选择需要克隆的节点和克隆目标工作空间。

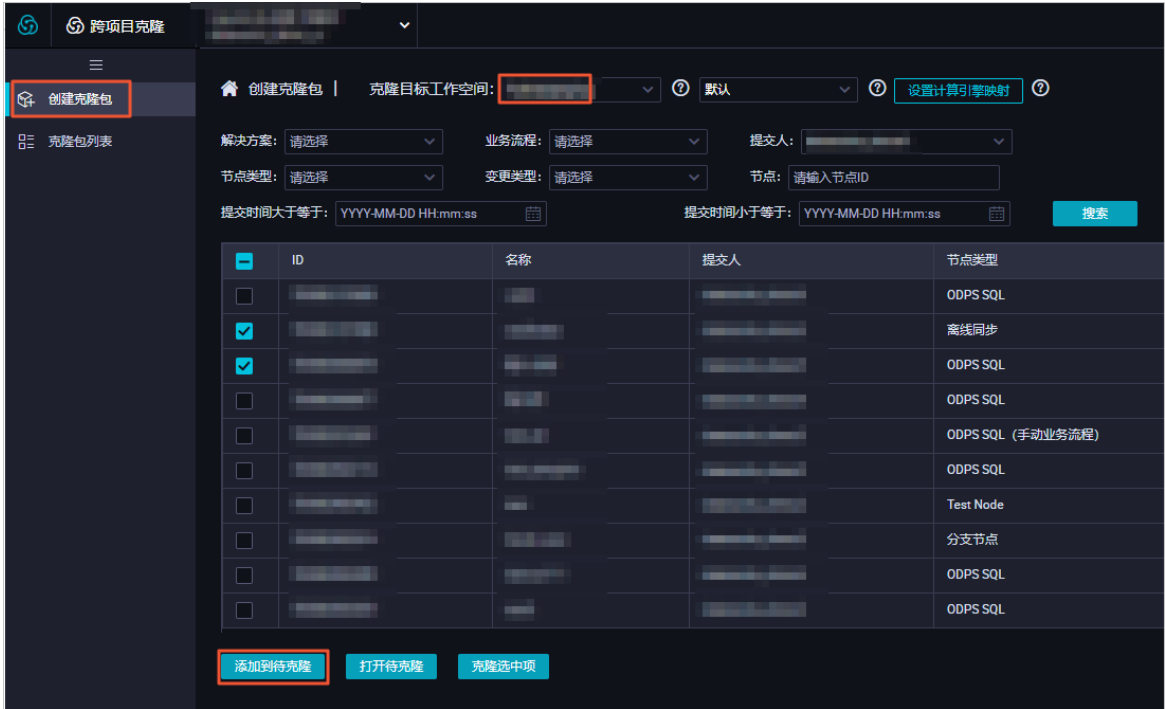
5. 单击设置计算引擎映射，设置当前工作空间与目标工作空间计算引擎的映射关系。

目标工作空间存在多个计算引擎，因此需要设置当前工作空间与目标工作空间计算引擎的映射关系，才可以进行克隆。如果不设置，则克隆至目标工作空间的默认计算引擎中。

说明

- 如果目标工作空间不存在克隆的节点所属的引擎类型，计算引擎映射信息对话框会进行提示。您可以通过选择跳过目标引擎实例为空的节点，来过滤无法克隆的节点，否则在克隆过程中会报错。
- 在源工作空间和目标工作空间某种引擎类型存在两个以上引擎实例的情况下，会显示设置计算引擎映射按钮。

6. 单击添加到待克隆，添加需要克隆的节点至待克隆列表。



- 7. 打开右上角的待克隆列表，单击全部克隆。
- 8. 确认引擎映射预检后的计算引擎映射信息，单击确定。
- 9. 待页面提示克隆成功后，即可进入目标工作空间查看克隆结果，通常会克隆业务流程的整体目录结构。

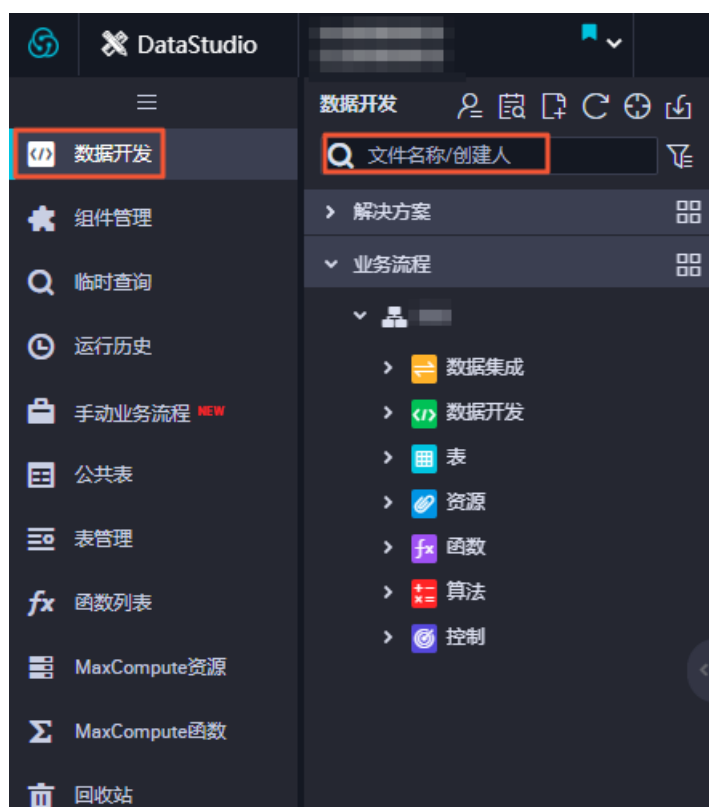


2.2. 下线任务

任务下线是指在某些情况下，需要将任务永久删除，包括开发环境的任务下线和生产环境的任务下线两种场景。

开发环境的任务下线

1. 登录DataWorks控制台，进入数据开发页面。
2. 通过任务节点类型、关键字来搜索需要删除的任务。



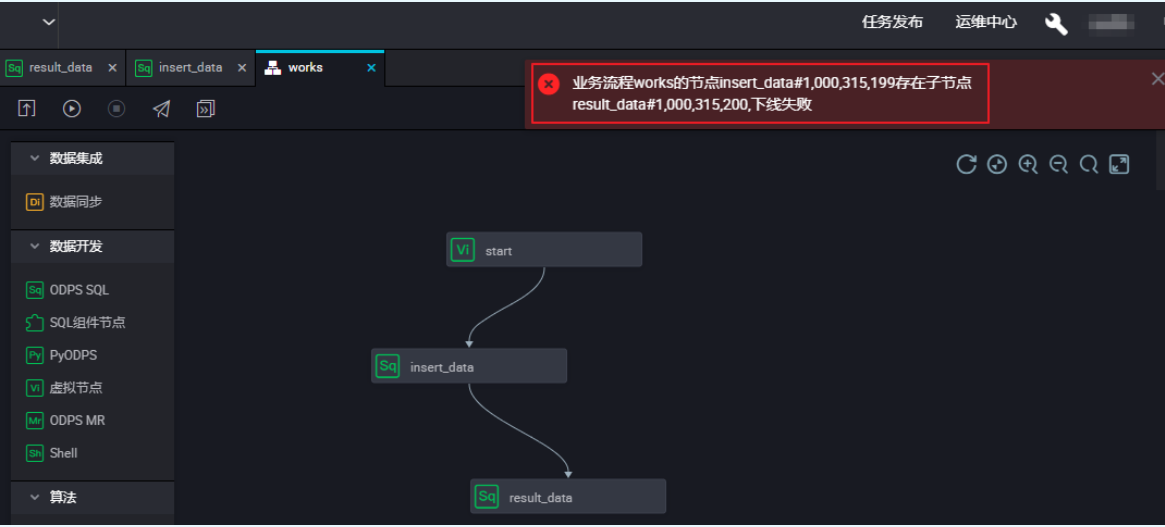
3. 右键单击要删除的任务，选择删除，则开发环境任务下线完成。



生产环境的任务下线

当已发布到生产环境的任务需要删除时，需按照删除任务>发布下线任务>执行发布的流程进行下线。

说明 子节点依赖关系的处理：由于生产环境的任务下线涉及到子节点的依赖，因此在您删除任务前，请先一层层往下处理好子节点的依赖关系后，再进行删除。否则会提示存在子节点，下线失败。



处理步骤：

1. 查找此节点的下游节点，可在工作流管理查看生产调度依赖关系。
2. 在数据开发页面重新编辑此子节点的父节点，或者直接删除此子节点。

如果提示子节点下还有子节点，请参见上述步骤逐层向下处理。

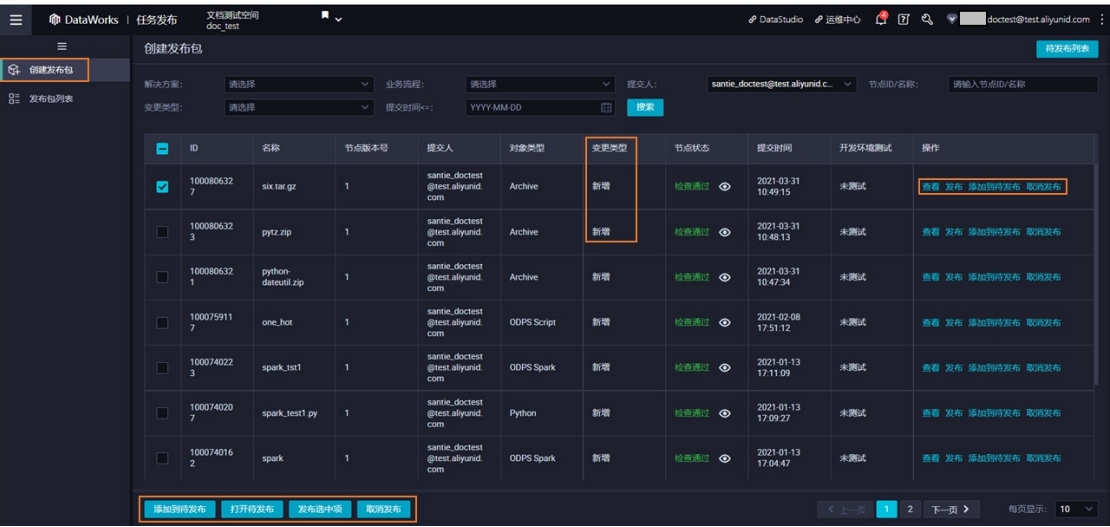
1. 删除任务。

可以参见前文开发环境的任务下线操作，删除需要下线的任务。

2. 发布下线任务。

说明 仅管理员及运维角色具有发布权限。如果是其他角色，建议通知运维人员进行发布。

- i. 删除需要下线的任务后，单击右上角的任务发布。
- ii. 在创建发布包页面，勾选需要下线的任务。



iii. 单击发布选中项。

您也可以单击添加到待发布，进入待发布列表进行发布。

3. 执行发布。

单击**确认执行**对话框中的**发布**，完成下线任务的发布。



3. 跨项目克隆

3.1. 跨项目克隆说明

跨项目克隆主要用于隔离同租户（阿里云账号）简单模式工作空间下的开发环境和生产环境，您也可以利用跨项目克隆功能实现计算、同步等类型的任务在工作空间之间的克隆迁移。本文为您介绍如何处理跨项目克隆时任务间的依赖关系。

通过跨项目克隆功能进行克隆任务后，系统为区分同租户（阿里云帐号）下不同工作空间之间任务的输出名称，会自动对每个任务输出名称作出一系列命名更改，目的是为了平滑复制依赖关系或保持原有依赖关系不变。

说明

- 跨项目克隆不支持跨地域发布。
- 目前不支持克隆旧版工作流至新的工作空间，请迁移源端旧版工作流中的任务至业务流程下的某个目录后，再克隆该业务流程至目标端工作空间。

克隆责任人分为默认和克隆包创建者：

- 当克隆责任人为默认的项目管理员时，克隆至目标工作空间后，您可以选择克隆后任务责任人为默认或克隆包创建者。



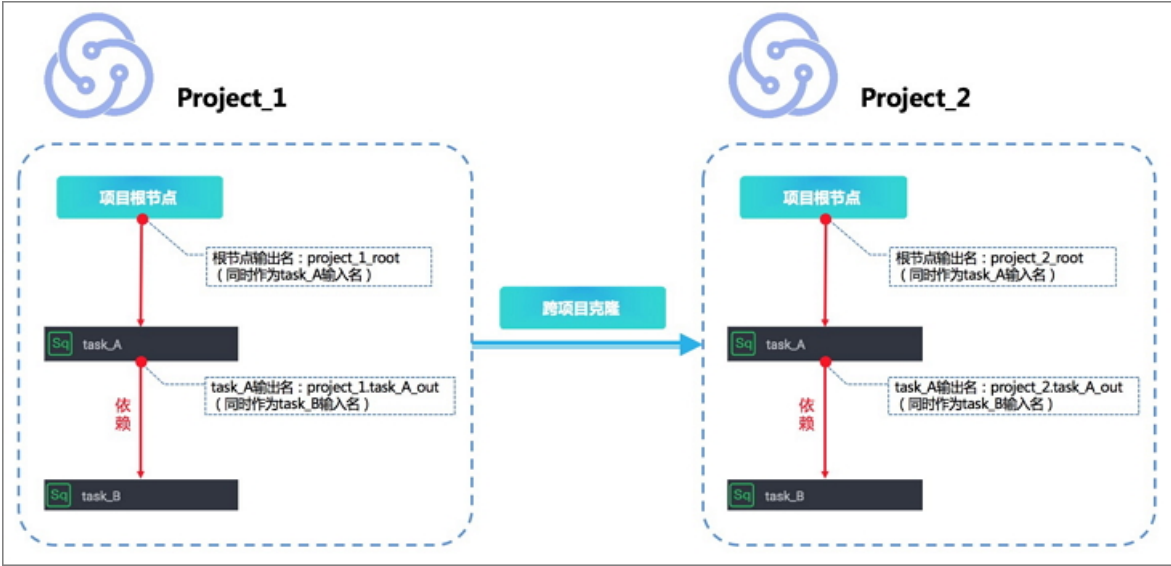
克隆成功后，责任人将第一优先级被置为原责任人。如果原责任人不在目标工作空间，则置为克隆包创建者。

- 当克隆责任人为克隆包创建者时，克隆至目标工作空间后，您可以选择克隆后任务责任人为默认或克隆包创建者。

克隆成功后，责任人将第一优先级被置为原责任人。如果原责任人不在目标工作空间，会询问是否变更责任人。如果确认变更，则任务克隆成功且责任人变更为克隆包创建者。如果不变更责任人，则克隆任务取消。

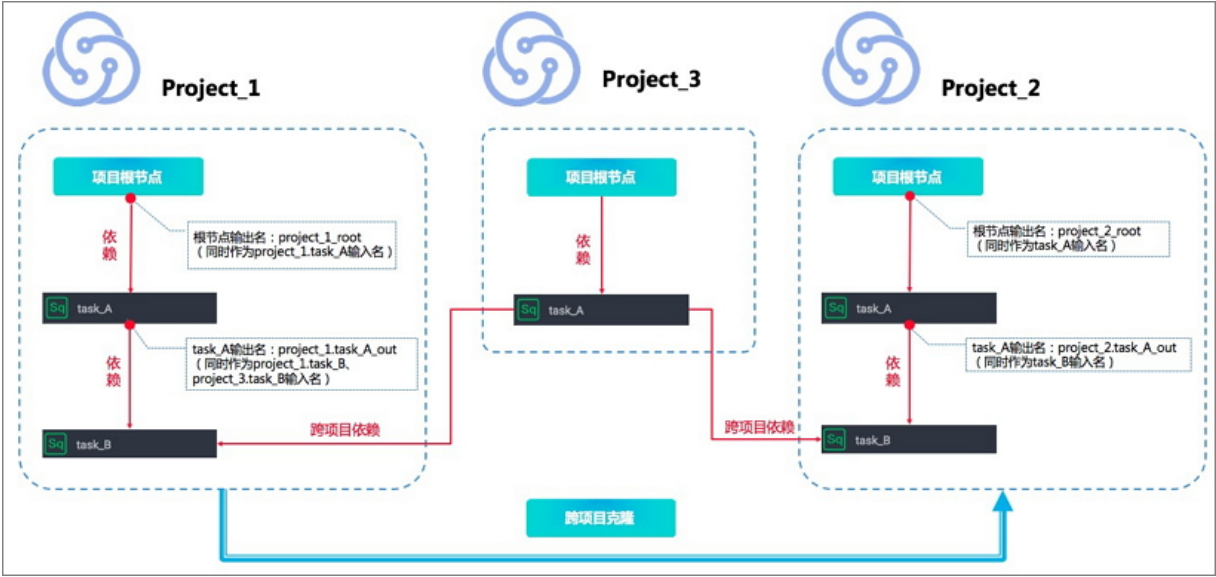
完整的业务流程克隆

用户使用task_A任务的输出点在project_1中为project_1.task_A_out，克隆至project_2之后输出点名为project_2.task_A_out。



跨项目依赖任务克隆

project_1中的任务task_B依赖了project_3中的任务task_A，在将project_1.task_B克隆为project_2.task_B之后，依赖关系将一同克隆，即project_2.task_B仍然依赖project_3.task_A。



3.2. 跨项目克隆实践

本文将为您介绍跨项目克隆的操作实践。

支持的场景

跨项目克隆支持以下两种场景：

- 从一个简单模式的工作空间克隆至另一个简单模式的工作空间。
- 从一个简单模式的工作空间克隆至另一个标准模式的工作空间。

说明

- 跨项目克隆以节点为单位，文件夹和业务流程会跟随节点一起克隆至目标工作空间。
- 跨项目克隆不能在上游节点没有克隆至目标工作空间的情况下，进行克隆下游节点的操作。

操作步骤

- 进入DataStudio（数据开发）页面，打开相应的业务流程。
- 单击右上角的跨项目克隆，跳转至相应的克隆页面，过滤出相应的节点，并将任务克隆到目标工作空间。



- 筛选需要克隆的节点，并选择克隆目标工作空间。



- 单击设置计算引擎映射，设置当前工作空间与目标工作空间计算引擎的映射关系。



目标工作空间存在多个计算引擎，因此需要设置当前工作空间与目标工作空间计算引擎的映射关系，方可进行克隆操作。如果不设置，则克隆至目标工作空间的默认计算引擎中。

说明

- 当克隆的节点中存在部分节点所属的引擎类型在目标工作空间中不存在的情况，计算引擎映射信息对话框会进行提示，您可以通过勾选计算引擎确认对话框会给出提醒，可以通过勾选跳过目标引擎实例为空的节点，来过滤无法克隆的节点，否则在克隆过程中会报错。
- 在源工作空间和目标工作空间某种引擎类型存在两个以上引擎实例的情况下，会显示设置计算引擎映射按钮。

5. 单击添加到克隆，添加需要克隆的节点至待克隆列表。



6. 打开右上角的待克隆列表，单击全部克隆。



7. 确认引擎映射预检后的计算引擎映射信息，单击确定。



8. 待页面提示克隆成功后，即可进入目标工作空间查看克隆结果，通常会克隆业务流程的整体目录结构。

