

数据治理中心

产品介绍

文档版本 01
发布日期 2022-08-01



版权所有 © 华为云计算技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目录

1 图解数据治理中心 DataArts Studio.....	1
2 什么是数据治理中心 DataArts Studio.....	3
3 基本概念.....	6
4 产品功能.....	12
5 产品优势.....	19
6 应用场景.....	21
7 如何选择版本.....	25
8 计费说明.....	28
9 安全.....	34
9.1 责任共担.....	34
9.2 资产识别与管理.....	35
9.3 身份认证与访问控制.....	36
9.4 数据保护技术.....	37
9.5 审计与日志.....	38
9.6 服务韧性.....	38
9.7 监控安全风险.....	38
9.8 故障恢复.....	38
9.9 更新管理.....	38
9.10 认证证书.....	39
10 权限管理.....	41
11 权限列表.....	57
12 约束与限制.....	71
13 与其他云服务的关系.....	74

1 图解数据治理中心 DataArts Studio

数据治理中心 DataArts Studio
一站式数据集成、开发、治理、开放平台

企业数字化转型面临的挑战有哪些？

- 数据治理的挑战**
 - 数据语言不统一
 - 数据标准不一
 - 数据质量不佳
 - 数据不可信
- 数据运营的挑战**
 - 数据开发周期长，效率低
 - 数据口径不一致
- 数据创新的挑战**
 - 数据不共享，不流通
 - 数据无法精准业务

什么是DataArts Studio？

数据治理中心(DataArts Studio) 是为了应对企业数字化转型的挑战，针对企业数字化运营诉求提供的端到端全生命周期管理，具有智能数据管理能力的**一站式数据集成、开发、治理、开放平台**。

DataArts Studio构建于数据湖底座之上，提供数据集成、数据开发、数据治理、数据运营、数据安全、数据服务、数据安全等能力。DataArts Studio支持对接所有华为云数据湖与大数据服务作为数据湖底座，如大数据服务MRS、数据湖探索服务DLI (Data Lake Insight)、数据仓库服务DWS等，也支持对接外部的企业数据湖，例IBCloud、MyC3等。

数据治理中心DataArts Studio

数据集成 | 数据开发 | 数据治理 | 数据运营 | 数据安全

数据湖底座 | 数据集成 | 数据开发 | 数据治理 | 数据运营 | 数据安全

数据集成 | 数据开发 | 数据治理 | 数据运营 | 数据安全

1 一站式数据治理平台

提供数据治理的一站式治理平台，帮助企业实现端到端的数据治理。

2 智能数据治理管控

提供智能数据治理管控，帮助企业实现端到端的数据治理。

3 丰富的数据开发类型

支持多人在线协同开发，支持脚本开发与可视化开发，提供丰富的数据开发类型。

4 统一数据治理

提供统一的数据治理，帮助企业实现端到端的数据治理。

5 数据运营全生命周期

提供数据运营全生命周期管理，帮助企业实现端到端的数据运营。

6 全方位的安全保障

提供全方位的安全保障，帮助企业实现端到端的数据安全。

DataArts Studio功能介绍

数据集成
提供企业数据集成能力，帮助企业实现端到端的数据集成。

数据开发
提供企业数据开发能力，帮助企业实现端到端的数据开发。

数据治理
提供企业数据治理能力，帮助企业实现端到端的数据治理。

数据运营
提供企业数据运营能力，帮助企业实现端到端的数据运营。

数据安全
提供企业数据安全保障能力，帮助企业实现端到端的数据安全。

2 什么是数据治理中心 DataArts Studio

企业数字化转型面临的挑战

企业在进行数据管理时，通常会遇到下列挑战。

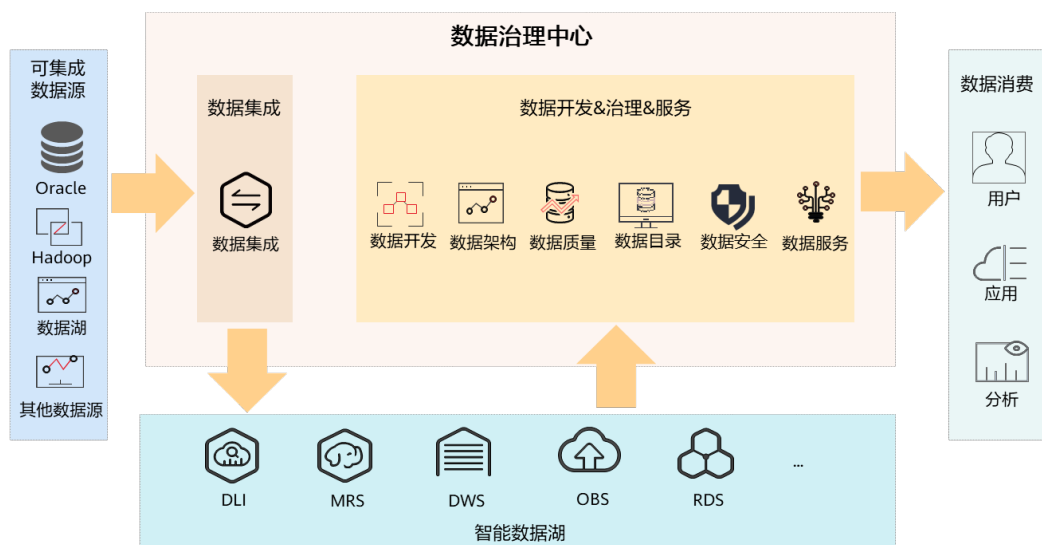
- 数据治理的挑战
 - 缺乏企业数据体系标准和数据规范定义的方法论，数据语言不统一。
 - 缺乏面向普通业务人员的高效、准确的数据搜索工具，数据找不到。
 - 缺乏技术元数据与业务元数据的关联，数据读不懂。
 - 缺乏数据的质量管控和评估手段，数据不可信。
- 数据运营的挑战
 - 数据运营效率低，业务环境的快速变化带来大量多样化的数据分析报表需求，因为缺乏高效的数据运营工具平台，数据开发周期长、效率低，不能满足业务运营决策人员的诉求。
 - 数据运营成本高，数据未服务化，导致数据复制多、数据口径不一致，同时数据重复开发，造成资源浪费。
- 数据创新的挑战
 - 企业内部存在大量数据孤岛，导致数据不共享、不流通，无法实现跨领域的数据分析与数据创新。
 - 数据的应用还停留在数据分析报表阶段，缺乏基于数据反哺业务推动业务创新的解决方案。

什么是 DataArts Studio?

数据治理中心DataArts Studio是为了应对上述挑战，针对企业数字化运营诉求提供的具有数据全生命周期管理和智能数据管理能力的一站式治理运营平台，包含数据集成、数据开发、数据架构、数据质量监控、数据资产管理、数据服务、数据安全等功能，支持行业知识库智能化建设，支持大数据存储、大数据计算分析引擎等数据底座，帮助企业快速构建从数据接入到数据分析的端到端智能数据系统，消除数据孤岛，统一数据标准，加快数据变现，实现数字化转型。

产品架构如[图2-1](#)所示。

图 2-1 产品架构



如图所示，DataArts Studio 基于数据湖底座，提供数据集成、开发、治理等能力。DataArts Studio 支持对接所有华为云的数据湖与数据库云服务作为数据湖底座，例如 MRS Hive、数据仓库服务 DWS 等，也支持对接企业传统数据仓库，例如 Oracle、MySQL 等。

DataArts Studio 包含如下功能组件：

- 管理中心**

提供 DataArts Studio 数据连接管理的能力，将 DataArts Studio 与数据湖底座进行对接，用于数据开发与数据治理等活动。
- 数据集成**

数据集成提供 20+ 简单易用的迁移能力和多种数据源到数据湖的集成能力，全向导式配置和管理，支持单表、整库、增量、周期性数据集成。
- 数据架构**

作为数据治理的一个核心模块，承担数据治理过程中的数据加工并业务化的功能，提供智能数据规划、自定义主题数据模型、统一数据标准、可视化数据建模、标注数据标签等功能，有利于改善数据质量，有效支撑经营决策。
- 数据开发**

大数据开发环境，降低用户使用大数据的门槛，帮助用户快速构建大数据处理中心。支持数据建模、数据集成、脚本开发、工作流编排等操作，轻松完成整个数据的处理分析流程。
- 数据质量**

数据全生命周期管控，数据处理全流程质量监控，异常事件实时通知。
- 数据目录**

提供企业级的元数据管理，厘清信息资产。通过数据地图，实现数据血缘和数据全景可视，提供数据智能搜索和运营监控。
- 数据服务**

数据服务定位于标准化的数据服务平台，提供一站式数据服务开发、测试部署能力，实现数据服务敏捷响应，降低数据获取难度，提升数据消费体验和效率，最终实现数据资产的变现。

- **数据安全**

数据安全为数据治理中心提供数据生命周期内统一的数据使用保护能力。通过访问权限管理、敏感数据识别、隐私保护管理等措施，帮助用户建立安全预警机制，增强整体安全防护能力，让数据可用不可得和安全合规。

3 基本概念

数据库、数据仓库、数据湖与华为智能数据湖方案

请参考[数据库、数据仓库、数据湖与华为智能数据湖方案](#)。

DataArts Studio 实例

DataArts Studio实例是数据治理中心给用户提供的最小计算资源单位。数据治理中心以DataArts Studio实例的方式提供给用户，用户可以同时创建多个DataArts Studio实例，并分别管理和访问每个DataArts Studio实例。每个DataArts Studio实例具有用户指定的基础计算资源，包含管理中心、数据架构、数据集成、数据开发、数据质量、数据目录等组件。用户可根据业务需要申请相应规格的DataArts Studio实例。

工作空间

工作空间是从系统层面为管理者提供对使用DataArts Studio的用户（成员）权限、资源、DataArts Studio底层计算引擎配置的管理能力。

工作空间作为成员管理、角色和权限分配的基本单元，每个团队都可具有独立的工作空间。

您只有在加入工作空间并被分配权限后，才可具备管理中心、数据目录、数据质量、数据架构、数据服务、数据开发和数据集成等组件的操作权限。

成员和角色

成员是被授予工作空间访问或使用权限的华为账号。在添加工作空间成员时，您需要同时为添加的成员设置相应的角色。

角色是一组操作权限的集合。不同的角色拥有不同的操作权限，把角色授予成员后，成员即具有了角色的所有权限。每位成员至少要拥有一个角色，并且可以同时拥有多种角色。

数据集成集群

一个数据集成集群运行在一个弹性云服务器之上，用户可以在集群中创建数据迁移作业，在云上和云下的同构/异构数据源之间批量迁移数据。

数据源

即数据的来源，本质是讲存储或处理数据的媒介，比如：关系型数据库、数据仓库、数据湖等。每一种数据源不同，其数据的存储、传输、处理和应用的模式、场景、技术和工具也不相同。

源数据

源数据强调数据状态是“创建”之后的“原始状态”，也就是没有被加工处理的数据。在数据管理的过程中，源数据一般是指直接来自源文件（业务系统数据库、线下文件、IoT等）的数据，或者直接复制源文件的“副本数据”。

数据连接

定义访问数据实体存储（计算）空间所需的信息的集合，包括连接类型、名称和登录信息等。

并发数

并发数是数据集成作业中，可以从源端并行读取的最大线程数。

脏数据

脏数据是对于业务没有意义或者格式非法的数据。例如，源端是VARCHAR类型的数据写到INT类型的目标列中，导致因为转换不合理而无法写入的数据。

作业（数据开发）

在数据开发中，作业由一个或多个节点组成，共同执行以完成对数据的一系列操作。

节点

节点用于定义对数据执行的操作。例如，使用“MRS Spark”节点可以实现在MRS中执行预先定义的Spark作业。

解决方案

解决方案定位于为用户提供便捷的、系统的方式管理作业，更好地实现业务需求和目标。每个解决方案可以包含一个或多个业务相关的作业，一个作业可以被多个解决方案复用。

资源

用户可以上传自定义的代码或文本文件作为资源，并在节点运行时调用。

表达式

数据开发作业中的节点参数可以使用表达式语言（Expression Language，简称EL），根据运行环境动态生成参数值。数据开发 EL表达式使用简单的算术和逻辑计算，引用内嵌对象，包括作业对象和一些工具类对象。

环境变量

环境变量是在操作系统中一个具有特定名字的对象，它包含了一个或者多个应用程序所将使用到的信息。

补数据

手工触发周期方式调度的作业任务，生成过去某时间段内的实例。

数据治理

数据资源及其应用过程中相关管控活动、绩效和风险管理的集合。

数据调研

基于现有业务数据、行业现状进行数据调查、需求梳理、业务调研，输出企业业务流程以及数据主题划分。

主题设计

通过分层架构表达对数据的分类和定义，帮助厘清数据资产，明确业务领域和业务对象的关联关系。

主题域分组

基于业务场景对主题域分组。

主题域

互不重叠数据的高层面的数据分类，用于管理其下一级的业务对象。

业务对象

指企业运作和管理中不可缺少的重要人、事、物信息。

流程设计

流程设计是针对流程的一个结构化的整体框架，描述了企业流程的分类、层级以及边界、范围、输入/输出关系等，反映了企业的商业模式及业务特点。

数据标准

数据标准用于描述公司层面需共同遵守的数据含义和业务规则。其描述了公司层面对某个数据的共同理解，这些理解一旦确定下来，就应作为企业层面的标准在企业内被共同遵守。

码表

通常只包括一系列允许的值和附加文本描述，与数据标准关联用于生成值域校验质量监控。

SDI

Source Data Integration (SDI) 又称贴源数据层。SDI是源系统数据的简单落地。

DWI

Data Warehouse Integration (DWI) 又称数据整合层。DWI整合多个源系统数据，源系统进来数据会有整合、清洗，基于三范式关系建模。

DWR

Data Warehouse Report (DWR) 又称数据报告层。DWR基于多维模型，和DWI层数据粒度保持一致。

DM

Data Mart (DM) 又称数据集市。DM面向展现层，数据有多级汇总。

关系建模

关系建模是用实体关系 (Entity Relationship, ER) 模型描述企业业务，它在范式理论上符合3NF，出发点是整合数据，将各个系统中的数据以整个企业角度按主题进行相似性组合和合并，并进行一致性处理，为数据分析决策服务，但是并不能直接用于分析决策。

维度建模

维度建模是从分析决策的需求出发构建多维模型，它主要是为分析需求服务，因此它重点关注用户如何更快速地完成需求分析，同时具有较好的大规模复杂查询的响应性能。

多维模型是由数字型度量值组成的一张事实表连接到一组包含描述属性的多张维度表，事实表与维度表通过主/外键实现关联。

在DataArts Studio数据架构中，维度建模是以维度建模理论为基础，构建总线矩阵、抽象出事实和维度，构建维度模型和事实模型，同时对报表需求进行抽象整理出相关指标体系，构建出汇总模型。

指标 (数据架构)

指标是衡量目标总体特征的统计数值，是能表征企业某一业务活动中业务状况的数值指示器。指标一般由指标名称和指标数值两部分组成，指标名称及其涵义体现了指标质的规定性和量的规定性两个方面的特点，指标数值反映了指标在具体时间、地点、条件下的数量表现。

度量

度量是用于衡量业务状况的可量化的数值表现，通常为数字，如：金额、数量、周期等。指标与度量的关系：度量是一个数值型数据，其本身不带有业务含义，只有将度量放在业务语境下，方能体现出业务含义，才能成为指标。

维度

维度是用于观察和分析业务数据的视角，支撑对数据汇聚、钻取、切片分析，用于SQL中的GROUP BY条件。多数维度具有层级结构，如：地理维度(其中包括国家、地

区、省以及城市等级别的内容)、时间维度(其中包括年度、季度、月度等级别的内容)。

原子指标

原子指标中的度量和属性来源于多维模型中的维度表和事实表，与多维模型所属的业务对象保持一致，与多维模型中的最细数据粒度保持一致。原子指标中仅含有唯一度量，所含其它所有与该度量、该业务对象相关的属性，旨在用于支撑衍生指标的敏捷自助消费。如：零售门店数量(包含门店名称、门店等级等属性)。

衍生指标

衍生指标是原子指标通过添加口径/修饰词、维度卷积而成，口径/修饰词、维度均来源于原子指标中的属性。例如：促销员门店覆盖率。

复合指标

复合指标由一个或多个衍生指标叠加计算而成，其中的维度、口径/修饰词均继承于衍生指标（不能脱离衍生指标维度和口径/修饰词的范围，去产生新的维度和口径/修饰词）。

数据质量规则

不符合业务实质、不满足数据应用要求的数据判断逻辑。

数据资产

由企业拥有或者控制的，能够为企业带来未来经济利益的，以物理或电子的方式记录的数据资源。在企业中并非所有的数据都构成数据资产，数据资产是能够为企业产生价值的的数据资源。

数据地图

以数据搜索为核心，通过可视化方式，综合反映有关数据来源、数量、分布、标准、流向、关联关系、数据质量。让用户找到数据、读懂数据、消费数据，致力于为用户提供高效率的数据消费产品。

元数据

元数据是关于数据的组织、数据域及其关系的信息，简言之，元数据就是关于数据的数据。元数据包括元数据实体和元数据元素。元数据元素是元数据的基本单元，若干个相关的元数据元素构成了元数据实体。

在DataArts Studio中，元数据是数据的描述数据，可以为数据说明其属性（数据连接、类型、名称、大小等），或其相关数据（位于拥有者、标签、分类、密级等）。

元数据采集

支持创建自定义策略的采集任务，采集数据源中的技术元数据。

数据资产报告

数据资产总览与统计信息展示。

数据服务

数据服务是基于数据分发、发布的框架，将数据作为一种服务产品提供，满足客户的实时数据需求，能复用并符合企业和工业标准，兼顾数据共享和安全。

API 网关

API网关（API Gateway）提供API托管服务，涵盖API发布、管理、运维、售卖的全生命周期管理。帮助您简单、快速、低成本、低风险地实现微服务聚合、前后端分离、系统集成，向合作伙伴、开发者开放功能和数据。

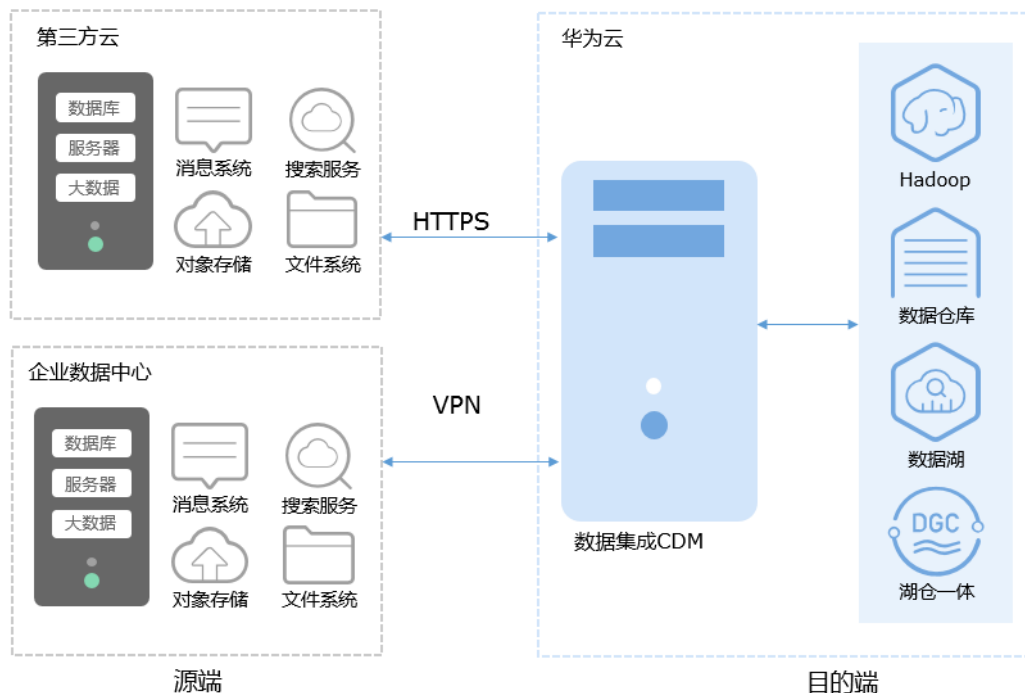
4 产品功能

数据集成：多种方式异构数据源高效接入

数据集成提供30+同构/异构数据源之间数据集成的功能，帮助您实现数据自由流动。支持自建和云上的文件系统，关系数据库，数据仓库，NoSQL，大数据云服务，对象存储等数据源。

数据集成基于分布式计算框架，利用并行化处理技术，支持用户稳定高效地对海量数据进行移动，实现不停服数据迁移，快速构建所需的数据架构。

图 4-1 数据集成



数据集成提供全向导式任务管理界面，帮助用户在几分钟内完成数据迁移任务的创建，轻松应对复杂迁移场景。数据集成支持的功能主要有：

- 表/文件/整库迁移

支持批量迁移表或者文件，还支持同构/异构数据库之间整库迁移，一个作业即可迁移几百张表。

- **增量数据迁移**

支持文件增量迁移、关系型数据库增量迁移、HBase增量迁移，以及使用Where条件配合时间变量函数实现增量数据迁移。

- **事务模式迁移**

支持当迁移作业执行失败时，将数据回滚到作业开始之前的状态，自动清理目的表中的数据。

- **字段转换**

支持去隐私、字符串操作、日期操作等常用字段的数据转换功能。

- **文件加密**

在迁移文件到文件系统时，数据集成支持对写入云端的文件进行加密。

- **MD5校验一致性**

支持使用MD5校验，检查端到端文件的一致性，并输出校验结果。

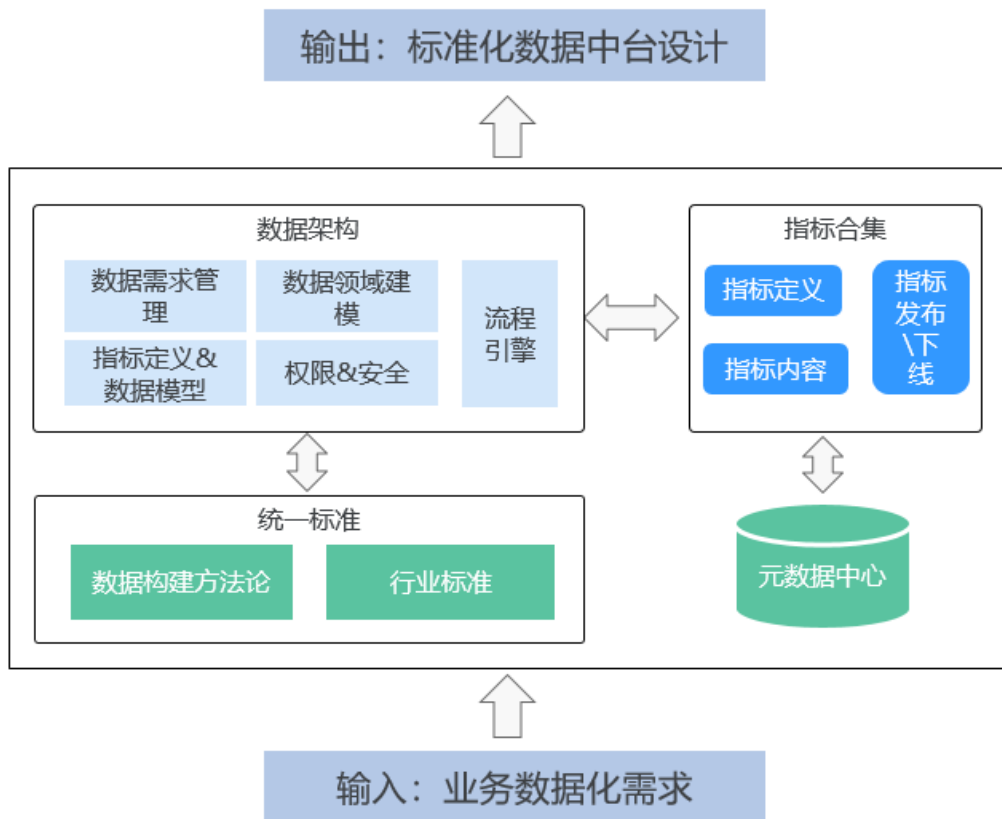
- **脏数据归档**

支持将迁移过程中处理失败的、被清洗过滤掉的、不符合字段转换或者不符合清洗规则的数据自动归档到脏数据日志中，方便用户分析异常数据。并支持设置脏数据比例阈值，来决定任务是否成功。

数据架构：数据建模可视化、自动化、智能化

DataArts Studio数据架构践行数据治理方法论，将数据治理行为可视化，打通数据基础层到汇总层、集市层的数据处理链路，落地数据标准和数据资产，通过关系建模、维度建模实现数据标准化，通过统一指标平台建设，实现规范化指标体系，消除歧义、统一口径、统一计算逻辑，对外提供主题式数据查询与挖掘服务。

图 4-2 数据架构



DataArts Studio数据架构主要包括以下三个部分：

- **主题设计**

构建统一的数据分类体系，用于目录化管理所有业务数据，便于数据的归类，查找，评价，使用。通过分层架构对数据分类和定义，可帮助用户厘清数据资产，明确业务领域和业务对象的关联关系。

- **数据标准**

构建统一的数据标准体系，数据标准流程化、系统化。用户可基于国家标准或行业标准，对每一行数据、每一个字段的具体取值进行标准化，从而提升数据质量和易用性。

- **数据建模**

构建统一的数据模型体系，通过规范定义和数据建模，自顶向下构建企业数据分层体系，沉淀企业数据公共层和主题库，便于数据的流通、共享、创造、创新，提升数据使用效率，极大的减少数据冗余，混乱，隔离，不一致以及谬误等。

DataArts Studio数据架构支持的数据建模方法有：

- **关系建模**

关系建模是用实体关系（Entity Relationship, ER）模型描述企业业务，它在范式理论上符合3NF，出发点是整合数据，将各个系统中的数据以整个企业角度按主题进行相似性组合和合并，并进行一致性处理，为数据分析决策服务，但是并不能直接用于分析决策。

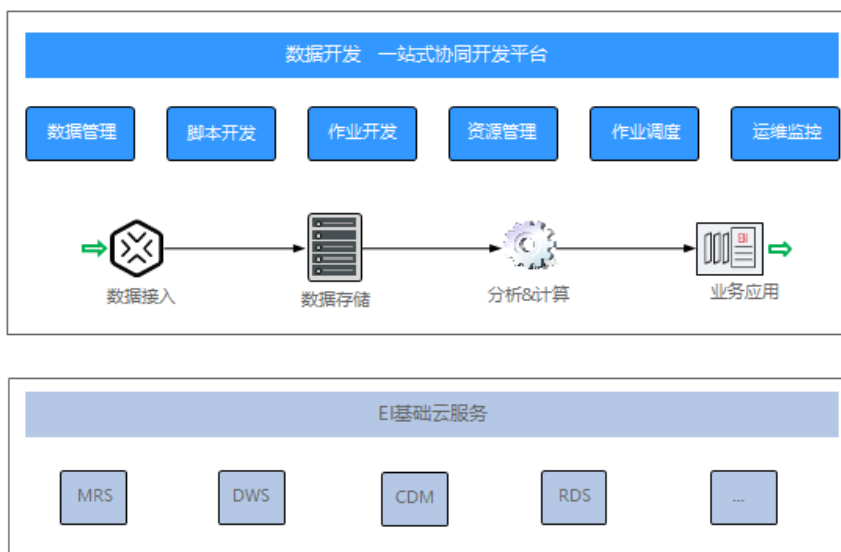
- **维度建模**

维度建模是以维度建模理论为基础，构建总线矩阵、抽象出事实和维度，构建维度模型和事实模型，同时对报表需求进行抽象整理出相关指标体系，构建出汇总模型。

数据开发：一站式协同开发平台

DataArts Studio数据开发是一个一站式敏捷大数据开发平台，提供可视化的图形开发界面、丰富的数据开发类型（脚本开发和作业开发）、全托管的作业调度和运维监控能力，内置行业数据处理pipeline，一键式开发，全流程可视化，支持多人在线协同开发，支持管理多种大数据云服务，极大地降低了用户使用大数据的门槛，帮助用户快速构建大数据处理中心。

图 4-3 数据开发模块架构



数据开发支持数据管理、脚本开发、作业开发、资源管理、作业调度、运维监控等操作，帮助用户轻松完成整个数据的处理分析流程。

- **数据管理**
 - 支持管理DWS、DLI、MRS Hive等多种数据仓库。
 - 支持可视化和DDL方式管理数据库表。
- **脚本开发**
 - 提供在线脚本编辑器，支持多人协作进行SQL、Shell、Python脚本在线代码开发和调测。
 - 支持使用变量。
- **作业开发**
 - 提供图形化设计器，支持拖拽式 workflow 开发，快速构建数据处理业务流水线。
 - 预设数据集成、SQL、Spark、Shell、机器学习等多种任务类型，通过任务间依赖完成复杂数据分析处理。
 - 支持导入和导出作业。
- **资源管理**
 - 支持统一管理在脚本开发和作业开发使用到的file、jar、archive类型的资源。

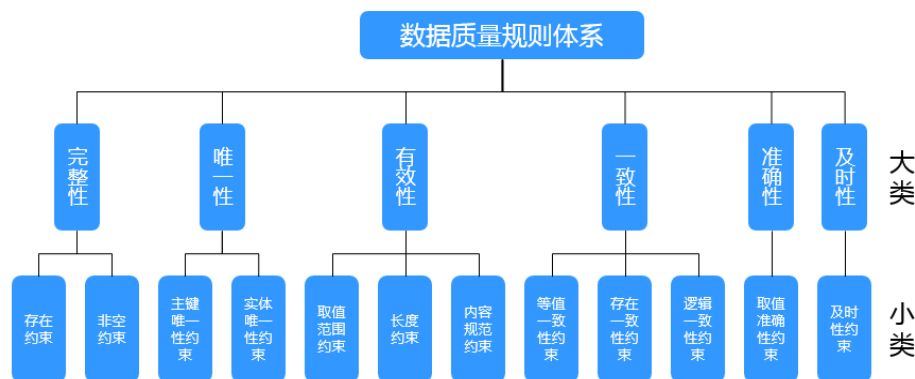
- **作业调度**
 - 支持单次调度、周期调度和事件驱动调度，周期调度支持分钟、小时、天、周、月多种调度周期。
 - 作业调度支持多种云服务的多种类型的任务混合编排，高性能的调度引擎已经经过几百个应用的检验。
- **运维监控**
 - 支持对作业进行运行、暂停、恢复、终止等多种操作。
 - 支持查看作业和其内各任务节点的运行详情。
 - 支持配置多种方式报警，作业和任务发生错误时可及时通知相关人，保证业务正常运行。

数据质量：可控可检验

数据质量模块支持对业务指标和数据质量进行监控，数据质量可检验，帮助用户及时发现数据质量问题。

- **业务指标监控**
业务指标监控是对业务指标数据进行质量管理的有效工具，可以灵活的创建业务指标、业务规则和业务场景，实时、周期性进行调度，满足业务的数据质量监控需求。
- **数据质量监控**
数据质量监控是对数据库里的数据质量进行质量管理的工具，您可以配置数据质量检查规则，在线监控数据准确性。
数据质量可以从完整性、有效性、及时性、一致性、准确性、唯一性六个维度进行单列、跨列、跨行和跨表的分析，也支持数据的标准化，能够根据数据标准自动生成标准化的质量规则，支持周期性的监控。

图 4-4 数据质量规则体系



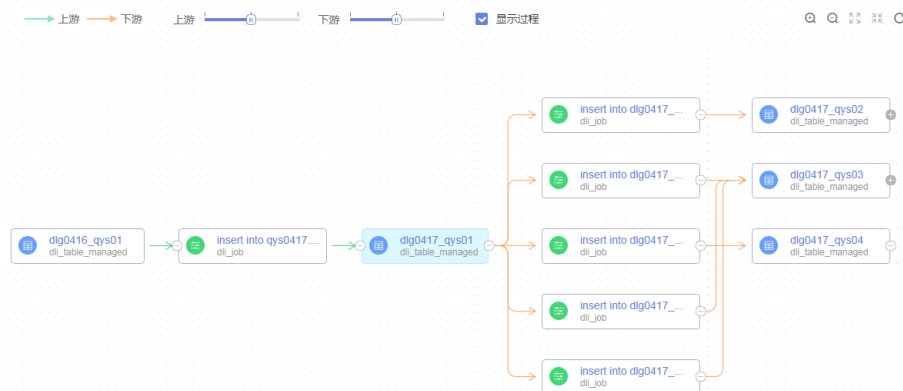
数据资产管理：360 度全链路数据资产可视化

DataArts Studio提供企业级的元数据管理，厘清信息资产。数据资产管理可视，支持钻取、溯源等。通过数据地图，实现数据资产的数据血缘和数据全景可视，提供数据智能搜索和运营监控。

- **元数据管理**

元数据管理模块是数据湖治理的基石，支持创建自定义策略的采集任务，可采集数据源中的技术元数据。支持自定义业务元模型，批量导入业务元数据，关联业务和技术元数据、全链路的血缘管理和应用。

图 4-5 全链路数据血缘



● **数据地图**

数据地图围绕数据搜索，服务于数据分析、数据开发、数据挖掘、数据运营等数据表的使用者和拥有者，提供方便快捷的数据搜索服务，拥有功能强大的血缘信息及影响分析。

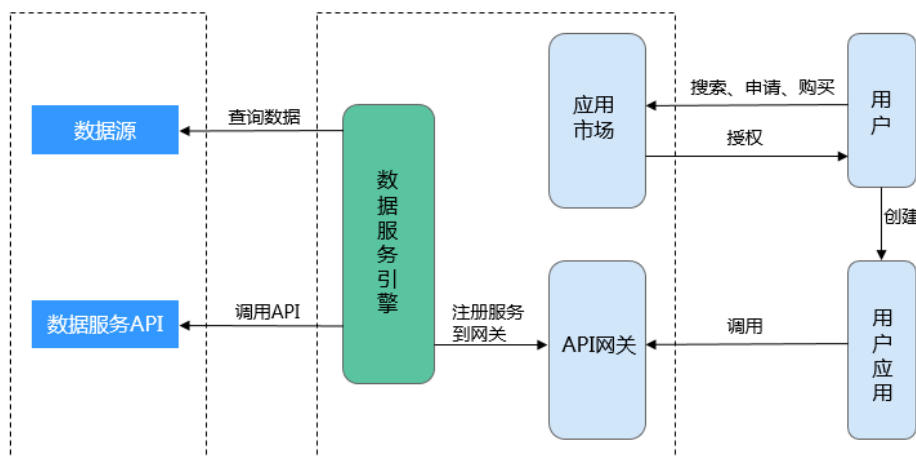
- 在数据地图中，可通过关键词搜索数据资产，支持模糊搜索，快速检索，定位数据。
- 使用数据地图根据表名直接查看表详情，快速查阅明细信息，掌握使用规则。获得数据详细信息后，可添加额外描述。
- 通过数据地图的血缘分析可以查看每个数据表的来源、去向，并查看每个表及字段的加工逻辑。
- 对数据资产，可以从业务角度定义分类或标签。

数据服务：提升访问查询检索效率

DataArts Studio数据服务旨在为企业搭建统一的数据服务总线，帮助企业统一管理对内对外的API服务，支撑业务主题/画像/指标的访问、查询和检索，提升数据消费体验和效率，最终实现数据资产的变现。数据服务为您提供快速将数据表生成数据API的能力，同时支持您将现有的API快速注册到数据服务平台以统一管理和发布。

数据服务采用Serverless架构，您只需关注API本身的查询逻辑，无需关心运行环境等基础设施，数据服务会为您准备好计算资源，并支持弹性扩展，零运维成本。

图 4-6 数据服务架构图



数据安全：全方位安全保障

- 网络安全**
 基于网络隔离、安全组规则以及一系列安全加固项，实现租户隔离和访问权限控制，保护系统和用户的隐私及数据安全。
- 用户权限策略**
 基于角色的访问控制，用户通过角色与权限进行关联，并支持细粒度权限策略，可满足不同的授权需求。针对不同的用户，DataArts Studio提供了管理者、开发者、部署者、运维者、访客五种不同的角色，各个角色拥有不同的权限。
- 数据安全**
 针对数据架构、数据服务等关键流程，DataArts Studio提供了审核流程。数据的分级分类管理，数据的全生命周期管理，保证数据的隐私合规、可回溯。

5 产品优势

一站式数据运营平台

贯穿数据全流程的一站式治理运营平台，提供全域数据集成、标准数据架构、连接并萃取数据价值、全流程数据质量监控、统一数据资产管理、数据开发服务等，帮助企业构建完整的数据中台解决方案。

全链路数据治理管控

数据全生命周期管控，提供数据架构定义及可视化的模型设计，智能化的帮助用户生成数据处理代码，数据处理全流程质量监控，异常事件实时通知。

丰富的数据开发类型

支持多人在线协作开发，脚本开发可支持SQL、Shell在线编辑、实时查询；作业开发可支持CDM、SQL、MRS、Shell、Spark等多种数据处理节点，提供丰富的调度配置策略与海量的作业调度能力。

统一调度和运维

全面托管的调度，支持按时间、事件触发的任务触发机制，支持分钟、小时、天、周和月等多种调度周期。

可视化的任务运维中心，监控所有任务的运行，支持配置各类报警通知，便于责任人实时获取任务的情况，保证业务正常运行。

可复用行业知识库

提供垂直行业可复用的领域知识库，涵盖行业数据标准、行业领域模型、行业数据主题库、行业算法库和行业指标库等，支持智慧政务、智慧税务、智慧园区等行业，帮助企业快速定制数据运营端到端解决方案。

统一数据资产管理

全局资产视图、快速查看、智能管理、数据溯源和数据开放共享，从业务视角管理和查看数据，定义业务架构、业务分类和业务术语，统一管理资产访问权限。

数据运营全场景可视

数据治理运营过程可视，拖拉拽配置，无需编码；处理结果可视，更直观，便于交互和探索；数据资产管理可视，支持钻取、溯源等。

全方位的安全保障

统一的安全认证，租户隔离，数据的分级分类管理，数据的全生命周期管理，保证数据的隐私合规、可审计、可回溯。

基于角色的访问控制，用户通过角色与权限进行关联，并支持细粒度权限策略，可满足不同的授权需求。

6 应用场景

一站式的数据运营治理平台

从数据采集->数据架构->质量监控->数据清洗->数据建模->数据联接->数据整合->数据消费->智能分析，一站式数据智能运营平台，帮助企业快速构建数据运营能力。

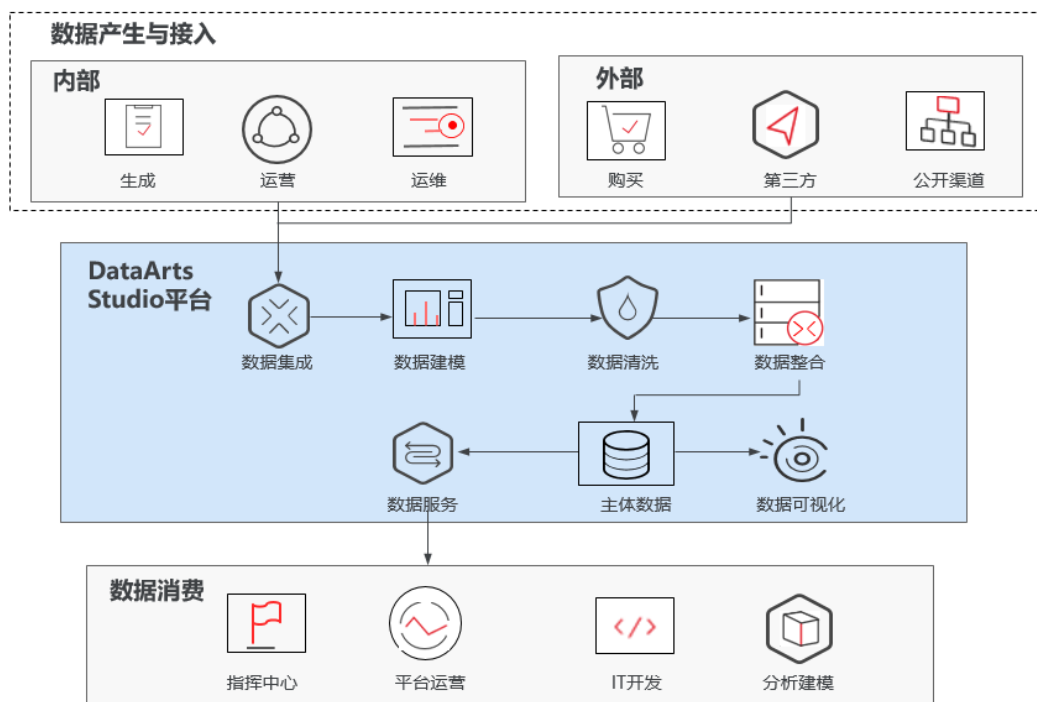
优势

- 多种云服务作业编排
- 全链路数据治理管控
- 丰富数据引擎支持

支持对接所有华为云的数据湖与数据库云服务，也支持对接企业传统数据仓库，比如Oracle等。

- 简单易用
图形化编排，即开即用，轻松上手。

图 6-1 一站式数据运营治理平台



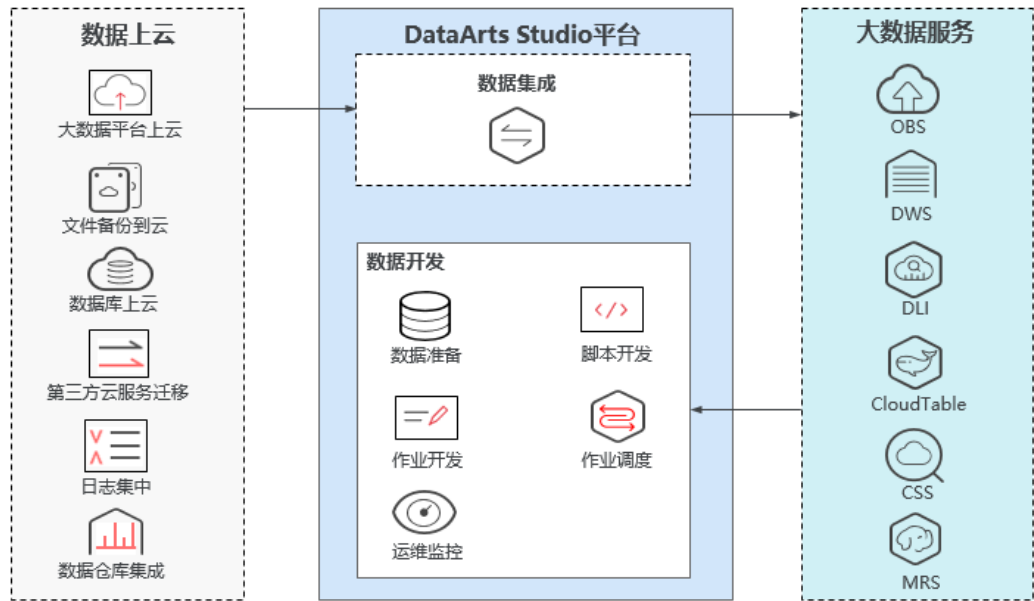
云上数据平台快速搭建

快速将线下数据迁移上云，将数据集成到云上大数据服务中，并在DataArts Studio的界面中就可以进行快速的数据开发工作，让企业数据体系的建设变得如此简单。

优势

- **数据集成一键式操作**
通过服务界面配置化操作，可实现线上线下数据快速集成到云数据仓库。
- **支持多种数仓服务类型**
根据需求，可以灵活选择数据服务类型，可以选择DWS服务建数仓，也可以选择MRS服务等数据平台。
- **安全稳定、降低成本**
一站式的服务能力和稳定的数仓服务，让云上数据万无一失；免自建大数据集群、免运维，极大降低企业建设数仓成本。

图 6-2 云上数据平台



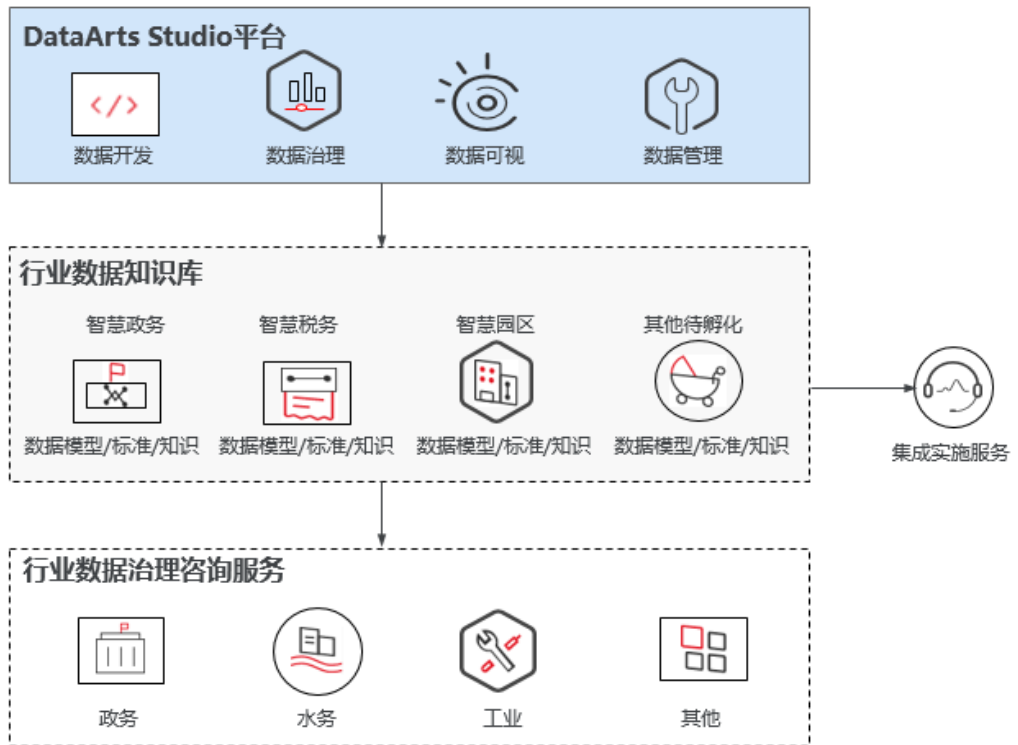
基于行业领域知识库快速构建数据中台

通过应用华为在企业业务领域积累的丰富的行业领域模型和算法，帮助企业构建数据中台，快速提升数据运营能力。

优势

- 多行业支持
覆盖政务/税务/城市/交通/园区等各行业。
- 标准规范支持
支持分层结构的行业数据标准。
- 领域模型丰富
支持包含人员/组织/事件/时空/车辆/资产/设备/资源等八大类数据以及相互之间关系的行业领域模型。
- 快速应用行业库
支持快速应用的行业主题库、行业算法库、行业指标库。

图 6-3 数据中台



7 如何选择版本

请您根据您所需的功能和规格，慎重选择DataArts Studio版本。

- 购买版本后，无法直接降级版本。如已购买企业版，无法直接降级到初级版，您需要采取先备份数据，然后退订原有实例，购买新实例，最后迁移备份数据到新实例的方式来实现版本降级。
- 如果您的业务量不断增长，已购版本无法满足您的业务需求，您可以选择升级版本。登录DataArts Studio控制台，找到需要升级的DataArts Studio实例，单击“升级”，然后根据页面提示购买更高规格的套餐。

如何选择 DataArts Studio 版本

表 7-1 DataArts Studio 各版本建议使用场景

版本	建议使用场景
初级版	建设初数据湖项目，无全职数据开发人员、不涉及数据治理。
基础版	拥有1~2名全职数据开发人员，原始数据规模小于1000张表。
高级版	拥有5~10名全职数据开发人员，有明确的数据标准、数据质量管理、数据规模小于2000张表。
专业版	有完善的数据管理团队和体系，拥有10~30名全职数据开发人员，中大型企业规模。
企业版	大型企业，多分支机构企业。

DataArts Studio 版本规格说明

表 7-2 DataArts Studio 版本支持的组件

DataArts Studio组件	初级版	基础版	高级版	专业版	企业版
数据集成	√	√	√	√	√

DataArts Studio组件	初级版	基础版	高级版	专业版	企业版
管理中心	√	√	√	√	√
数据架构	x	√	√	√	√
数据开发	√	√	√	√	√
数据质量	x	√	√	√	√
数据目录	x	√	√	√	√
数据服务	x	√	√	√	√
数据安全	x	√	√	√	√

表 7-3 DataArts Studio 版本规格说明（单实例规格）

DataArts Studio规格	初级版	基础版	高级版	专业版	企业版
DataArts Studio数据集成集群 [1]	集群数量： 1个 规格名称： cdm.medium vCPUs/内存： 4核 8GB	集群数量： 1个 规格名称： cdm.medium vCPUs/内存： 4核 8GB	集群数量： 1个 规格名称： cdm.large vCPUs/内存： 8核 16GB	集群数量： 1个 规格名称： cdm.xlarge vCPUs/内存： 16核 32GB	集群数量： 1个 规格名称： cdm.xlarge vCPUs/内存： 16核 32GB
作业节点调度次数/天 [2]	5,000次/天	20,000次/天	40,000次/天	80,000次/天	200,000次/天
技术资产数量 [3]	不支持	1,000	2,000	4,000	10,000
数据模型数量 [4]	不支持	1,000	2,000	4,000	10,000

注释：

[1] DataArts Studio数据集成集群： DataArts Studio实例赠送的数据集成集群，推荐作为管理中心数据连接的Agent代理使用，不建议同时作为Agent代理和运行数据迁移作业使用。用于运行数据迁移作业的其他规格数据集成集群可以通过“批量数据迁移增量包”进行购买，如何购买请参考[购买功能增量包](#)。

[2] 作业节点调度次数/天： 该规格是以每天执行的数据开发作业、质量作业、对账作业、业务场景和元数据采集作业的调度次数之和计算的。其中数据开发作业的每天调度次数，是以节点（包含Dummy节点）为粒度进行度量的，另外补数据任务也会计入度量次数，但测试运行、失败重试不会计入。例如某作业包含2个DWS SQL节点和1个

Dummy节点，每天00:00开始执行，隔10小时调度一次，当天额外进行了过去10天的补数据，则该作业当天调度次数为 $2*3+2*3*10=66$ ，后续每天调度次数为 $2*3=6$ 。

另外，当作业节点调度的已使用次数+运行中次数+本日将运行次数之和大于此版本规格时，就会提示作业节点调度次数/天超过配额。

[3] 技术资产数量：该规格指的是数据目录中表和OBS文件的数量。

[4] 数据模型数量：该规格指的是数据架构中逻辑模型、物理模型、维度表、事实表和汇总表的数量。

8 计费说明

数据治理中心DataArts Studio采用基础包+增量包的计费项。为了便于您便捷的下单购买，在控制台购买界面中系统会为您计算好所购买的套餐包价格，您可一键完成整个配置在购买。

须知

在您使用DataArts Studio的过程中，可能还会产生以下相关服务的费用，敬请知悉：

- OBS服务：在数据集成或数据开发时，DataArts Studio会将脏数据或日志文件写入到OBS服务中，则会产生对象存储服务费用，收费标准请参见[OBS价格详情](#)。
- EIP服务：如果您为数据集成集群或数据服务专享版集群开通了公网IP，则会产生弹性公网IP服务费用，收费标准请参见[EIP价格详情](#)。
- SMN服务：如果您在使用DataArts Studio各组件过程中开启了消息通知功能，则会产生消息通知服务费用，收费标准请参见[SMN价格详情](#)。
- DEW服务：在数据集成或创建管理中心连接时，如果启用了KMS，则会产生密钥管理费用，收费标准请参见[DEW价格详情](#)。

计费项

DataArts Studio采用基础包+增量包的计费项，具体说明如[表8-1](#)所示：

表 8-1 DataArts Studio 计费项

计费项	计费项子类	计费项说明	计费模式	计费说明
DataArts Studio基础包	初级版 基础版 高级版 专业版 企业版	DataArts Studio基础包对应DataArts Studio实例，如果需要使用DataArts Studio服务，则必须先购买DataArts Studio实例。 DataArts Studio实例根据不同版本，包含了不同的组件和规格，具体请参见 如何选择DataArts Studio版本 。	包年包月	DataArts Studio基础包不含其他相关云服务所产生的费用，例如弹性公网IP、对象存储服务等等。 DataArts Studio基础包的计费详情，请参见 数据治理中心价格详情 。
DataArts Studio增量包（可选）	批量数据迁移增量包	批量数据迁移增量包对应数据集成CDM集群。 数据集成CDM集群提供数据上云和数据入湖的集成能力，全向导式配置和管理，支持单表、整库、增量、周期性数据集成。	<ul style="list-style-type: none"> ● 按需计费 ● 套餐包（按需资源包） 	计费详情请参见 批量数据迁移增量包价格详情 。
	数据服务专享集群增量包	数据服务专享集群增量包对应数据服务专享版集群。 数据服务定位于标准化的数据服务平台，提供一站式数据服务开发、测试部署能力，实现数据服务敏捷响应，降低数据获取难度，提升数据消费体验和效率，最终实现数据资产的变现。数据服务专享版需要通过数据服务专享集群增量包创建集群后使用。	包年包月	套餐包计费情况以实际费用为准。 说明 数据服务专享版在每个DataArts Studio实例下具有创建10个专享版API免费试用额度，超出试用配额后会产生数据服务专享版API的费用，所创建的超出试用配额API按每天每个进行收费。

说明

套餐包（按需资源包）方式购买批量数据迁移增量包时，需注意如下几点：

- 套餐包（按需资源包）方式购买批量数据迁移增量包后，系统不自动创建CDM集群，而是和具体的区域和实例规格绑定，在生效期内的每个计费月内按月提供745小时/月的使用时长，在绑定区域给对应实例规格的CDM集群使用。
- 套餐包（按需资源包）仅支持给DataArts Studio控制台购买的CDM集群使用；在CDM控制台购买的CDM集群，不支持使用DataArts Studio增量包形式购买的套餐包（按需资源包），仅支持使用在云数据迁移CDM服务控制台购买的折扣套餐（按需资源包）。
- 如果当前绑定区域有1个或多个对应实例规格的CDM集群，则扣费方式是先扣除已购买资源包内的时长额度，超出部分以按需计费的方式进行结算（资源包对应多个集群时，会出现每月订购周期内可使用时长不足的情况）。

例如购买了1个月的套餐包（745小时/月），按区域和实例规格匹配到两个CDM集群后，从当前开始的1个月订购有效期内，两个集群同时使用只能使用 $745/2=372.5$ 小时，约15.5天，剩余时间内两个集群按照按需计费的方式结算费用。

- 如果当前绑定区域没有对应实例规格的CDM集群，购买套餐包后不会消耗所购买的时长；但在生效期内，若未使用CDM集群，套餐包也不会延期。建议您先安排好服务使用计划，再购买套餐包。
- 如果您希望享受套餐包的优惠价格，需要先购买一个“套餐包”增量包，再购买一个和套餐包具有相同区域和规格的“按需计费”增量包。
- 如果您先购买一个“按需计费”增量包，再购买一个相同区域和规格的“套餐包”增量包，则在购买套餐包之前已经产生的费用按“按需计费”计费，购买套餐包之后的费用按“套餐包”计费。

计费模式

DataArts Studio不同计费项所支持的计费模式，如表8-2所示。

表 8-2 DataArts Studio 支持的计费模式

计费模式	模式介绍	适用场景	支持的计费项
包年包月	<ul style="list-style-type: none"> 通过预付费，以包年或包月模式购买云资源。 购买成功后，系统分配云资源给账户使用。 使用中的云资源支持退订。 使用期限到期后，如未续费则进入宽限期以及保留期，直到资源被回收。 	适用于业务量稳定、需长期使用固定云资源的场景，可以享受更低的包周期资费。	<ul style="list-style-type: none"> DataArts Studio基础包 数据服务专享集群增量包 作业节点调度次数/天增量包 技术资产数量增量包 数据模型数量增量包
按需计费	<ul style="list-style-type: none"> 不需要预付费，可直接创建、使用按需资源。 按照资源使用量或使用时间进行计费。 	适用于业务量激增、短期或间断性使用云资源的场景，使用时长不固定，资费较高。	批量数据迁移增量包

计费模式	模式介绍	适用场景	支持的计费项
套餐包 (按需资源包)	<ul style="list-style-type: none">通过预付费，购买一定时间期限内的云资源使用量配额。购买后，系统不会自动分配云资源。使用中的按需资源包不支持退订，详情可查看不可退订。使用期限或使用量到期后，还可以继续使用云资源，超出部分按需计费。	适用于使用期限不固定、但需要长期使用云资源的客户，可以享受套餐包的低资费。	批量数据迁移增量包

变更配置

在购买DataArts Studio时，有多种基础包或增量包规格供您选择，您可根据业务需要选择合适的基础包或增量包。当您完成DataArts Studio基础包或增量包的购买后，DataArts Studio提供如下变更配置的方式。

- DataArts Studio基础包升级：**DataArts Studio基础包套餐允许升级，升级时，已经产生的费用按天结算。升级成功后，按新订购套餐进行计费。

如需升级，请登录DataArts Studio控制台，找到需要升级的DataArts Studio实例，单击“升级”，然后根据页面提示购买更高规格的套餐。

套餐升级后，系统会根据升级后的套餐创建一个新的数据集成集群。原基础包中的数据集成集群仍为您保留，但是不会再产生费用。您需要将原集群的数据连接、作业等信息手动迁移到新集群，具体操作请参见[批量管理作业](#)。

若DataArts Studio提供的变更配置方式不满足您的要求，您也可以重新购买DataArts Studio基础包或增量包，退订当前基础包或增量包。

续费

- DataArts Studio基础包**

DataArts Studio基础包套餐采用包年包月计费模式，套餐到期后相继进入宽限期和保留期（详情请参见[到期与欠费](#)），建议请在所购买的套餐包时长用完前进行续费。

DataArts Studio基础包支持自动续费，自动续费的默认续费周期为：

- 按月购买：自动续费周期为1个月。
- 按年购买：自动续费周期为1年。

您可以通过以下两种方式开通自动续费：

- 登录DataArts Studio控制台，在购买DataArts Studio实例的页面中，勾选“自动续费”选项。
- 如果您已购买DataArts Studio实例，请进入续费管理页面，在实例列表中查找所需续费的数据集成实例，单击其所在行的“开通自动续费”，然后请根据页面提示完成自动续费的开通。

您也可以进行手动续费，请进入[续费管理](#)页面，在列表中查找所需续费的DataArts Studio实例，单击其所在行的“续费”，进行手动续费操作。有关续费的更多信息，请参见[续费管理](#)。

- **DataArts Studio增量包**
 - **“包年包月”计费的DataArts Studio增量包**

包年包月计费的DataArts Studio增量包，当时长到期后相继进入宽限期和保留期（详情请参见[到期与欠费](#)），因此在欠费前请及时充值。

如需充值，请进入[续费管理](#)页面，在左侧导航树单击“总览”，在总览页面单击“充值”对账户进行充值。
 - **“按需计费”的批量数据迁移增量包**

按需计费是按每小时扣费，当余额不足后会相继进入宽限期和保留期（详情请参见[到期与欠费](#)），因此在欠费前请及时充值。

如需充值，请进入[续费管理](#)页面，在左侧导航树单击“总览”，在总览页面单击“充值”对账户进行充值。
 - **“套餐包（按需资源包）”计费的批量数据迁移增量包**

套餐包（按需资源包）计费的批量数据迁移增量包，所购买的时长到期后套餐结束，会自动转为“按需计费”方式。如果希望继续享受“套餐包”的优惠方式，请在所购买的套餐包时长用完前进行再次购买。

如需再次购买，请进入[续费管理](#)页面，找到所需购买的DataArts Studio增量包，然后单击“再次购买”。

到期与欠费

- **DataArts Studio基础包**

DataArts Studio基础包套餐采用包年包月计费模式，套餐到期后进入宽限期，宽限期内您可正常访问及使用DataArts Studio，且正常计费。您需要在宽限期内购买套餐时长，否则会进入保留期。

保留期内，数据仍予以保留，但是您将无法访问DataArts Studio实例，您无法在DataArts Studio管理控制台进行操作，相关接口也无法调用。如果在保留期结束时您没有续费，DataArts Studio将终止服务，系统中的数据也将被永久删除。

系统会根据用户等级赋予不同的宽限期和保留期，详见[宽限期保留期](#)。
- **DataArts Studio增量包**
 - **“包年包月”计费的DataArts Studio增量包**

包年包月计费模式的增量包到期后进入宽限期，宽限期内您可正常使用增量包，且正常计费。您需要在宽限期内购买套餐时长，否则会进入保留期。

保留期内，数据仍予以保留，但是您将无法使用增量包功能。如果在保留期结束时您没有续费，增量包将终止服务，系统中的数据也将被永久删除。

系统会根据用户等级赋予不同的宽限期和保留期，详见[宽限期保留期](#)。
 - **“按需计费”的批量数据迁移增量包**

按每小时扣费，没有到期的概念，只要账户中有余额就可以继续使用。当余额不足，无法对上一个小时的费用进行扣费，就会导致DataArts Studio增量包欠费，欠费后进入宽限期，宽限期内您可正常访问及使用DataArts Studio增量包的服务，但处于宽限期的资源是继续计费的。您需要在宽限期内续费，否则会进入保留期。

保留期内，数据仍予以保留，但是您将无法访问DataArts Studio增量包的服务，也无法进行相关操作，相关接口无法调用。如果在保留期结束时您没有续费，DataArts Studio将终止服务，系统中的数据也将被永久删除。

系统会根据用户等级赋予不同的宽限期和保留期，详见[宽限期保留期](#)。

- **“套餐包（按需资源包）”计费的批量数据迁移增量包**
所购买的时长到期后套餐结束，会自动转为“按需计费”方式继续计费。

退订

- **DataArts Studio基础包**

DataArts Studio基础包套餐生效期间，您可以根据需要，参考[云服务退订退订DataArts Studio包年包月套餐](#)。

另外在退订DataArts Studio后，对于在DataArts Studio使用过程中可能会产生费用的以下相关服务，请分别退订其资源，避免其依然计费。

- OBS服务：在数据集成或数据开发时，DataArts Studio会将脏数据或日志文件写入到OBS服务中，则会产生对象存储服务费用，收费标准请参见[OBS价格详情](#)。
- EIP服务：如果您为数据集成集群或数据服务专享版集群开通了公网IP，则会产生弹性公网IP服务费用，收费标准请参见[EIP价格详情](#)。
- SMN服务：如果您在使用DataArts Studio各组件过程中开启了消息通知功能，则会产生消息通知服务费用，收费标准请参见[SMN价格详情](#)。
- DEW服务：在数据集成或创建管理中心连接时，如果启用了KMS，则会产生密钥管理费用，收费标准请参见[DEW价格详情](#)。

- **DataArts Studio增量包**

- **“包年包月”计费的DataArts Studio增量包**
需参考[云服务退订退订套餐包](#)。
- **“按需计费”的批量数据迁移增量包**
“按需计费”的批量数据迁移增量包，在删除数据集成集群后，则不再按需计费。
- **“套餐包（按需资源包）”计费的批量数据迁移增量包**
按需资源包当前不支持退订，具体详情可查看[不可退订](#)。

9 安全

9.1 责任共担

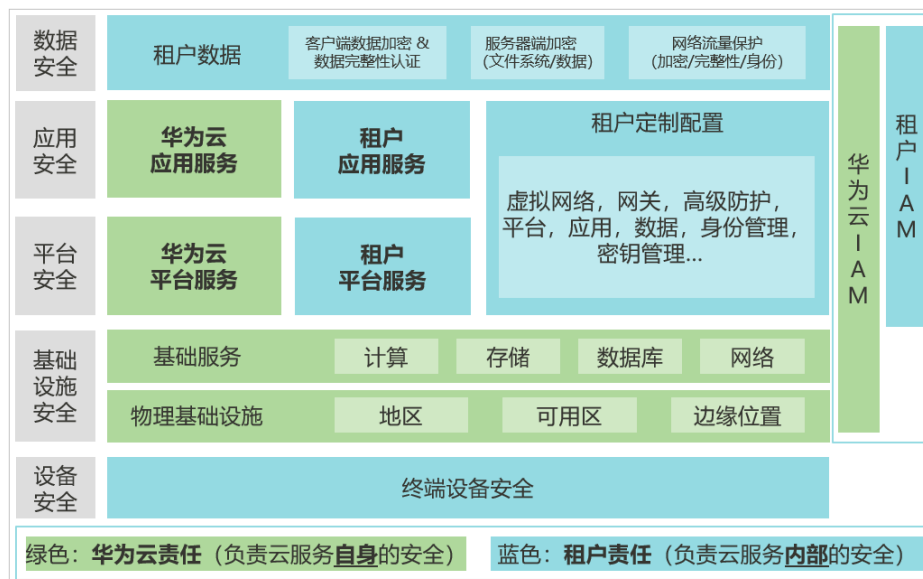
华为云秉承“将对网络和业务安全性保障的责任置于公司的商业利益之上”。针对层出不穷的云安全挑战和无孔不入的云安全威胁与攻击，华为云在遵从法律法规业界标准的基础上，以安全生态圈为护城河，依托华为独有的软硬件优势，构建面向不同区域和行业的完善云服务安全保障体系。

安全性是华为云与您的共同责任，如图9-1所示。

- **华为云**：负责云服务自身的安全，提供安全的云。华为云的安全责任在于保障其所提供的 IaaS、PaaS 和 SaaS 类云服务自身的安全，涵盖华为云数据中心的物理环境设施和运行其上的基础服务、平台服务、应用服务等。这不仅包括华为云基础设施和各项云服务技术的安全功能和性能本身，也包括运维运营安全，以及更广义的安全合规遵从。
- **租户**：负责云服务内部的安全，安全地使用云。华为云租户的安全责任在于对使用的 IaaS、PaaS 和 SaaS 类云服务内部的安全以及对租户定制配置进行安全有效的管理，包括但不限于虚拟网络、虚拟主机和访客虚拟机的操作系统，虚拟防火墙、API 网关和高级安全服务，各项云服务，租户数据，以及身份账号和密钥管理等方面的安全配置。

《[华为云安全白皮书](#)》详细介绍华为云安全性的构建思路与措施，包括云安全战略、责任共担模型、合规与隐私、安全组织与人员、基础设施安全、租户服务与租户安全、工程安全、运维运营安全、生态安全。

图 9-1 华为云安全责任共担模型



9.2 资产识别与管理

云资源的标识与管理可通过标签实现。

使用场景

通常您的业务系统可能使用了华为云的多种云服务，您可以为这些云服务下不同的资源实例分别设置标签（对于DataArts Studio而言，标签作用于其实例上），各服务产生的计费详单中都会体现这些资源实例和实例上设置的标签。如果您的业务系统是由多个不同的应用构成，为同一种应用拥有的资源实例设置统一的标签将很容易帮助您对不同的应用进行使用量分析和成本核算。

对DataArts Studio来说，标签用于标识购买的实例，以此来达到对购买的DataArts Studio实例进行分类的目的。当为实例添加标签时，该实例上所有请求产生的计费话单里都会带上这些标签，从而可以针对话单报表做分类筛选，进行更详细的成本分析。例如：某个实例作用于A部门，我们可以用该部门名称做为标签，设置到被使用的实例上。在分析话单时，就可以通过该部门名称的标签来分析此部门的开发使用成本。

DataArts Studio以键值对的形式来描述标签。一个实例默认最大拥有20个标签。每个标签有且只有一对键值。键和值可以任意顺序出现在标签中。同一个实例标签的键不能重复，但是值可以重复，并且可以为空。

使用方式

DataArts Studio支持通过控制台方式创建实例标签，详情请参见[创建实例标签](#)。

9.3 身份认证与访问控制

身份认证

用户访问DataArts Studio的方式有多种，包括控制台、开放API等，无论访问方式封装成何种形式，其本质都是通过DataArts Studio提供的REST风格的API接口进行请求。

DataArts Studio的接口均需要进行认证鉴权以此来判断是否通过身份认证。通过控制台发出的请求需要通过Token认证鉴权，调用API接口[认证鉴权](#)支持Token认证和AK/SK认证两种方式。

访问控制

您可以使用统一身份认证服务（Identity and Access Management，简称IAM）进行精细的权限管理。该服务提供用户身份认证、权限分配、访问控制等功能，可以帮助您安全的控制华为云资源的访问。关于IAM的详细介绍，请参见[IAM产品介绍](#)。

DataArts Studio**仅支持基于系统角色的授权，不支持策略授权**。为了实现精细的权限管控，DataArts Studio提供了**系统角色+工作空间角色**授权的能力，由工作空间角色授权具体的操作权限，并支持自定义不同权限点的工作空间角色。

如[图9-2](#)和[表9-1](#)所示，DataArts Studio的IAM系统角色包括DAYU Administrator和DAYU User；工作空间角色是基于IAM角色DAYU User进一步授予的，[权限列表](#)列出了DataArts Studio常用操作与工作空间角色的授权关系，您可以参照这些权限列表选择合适的角色。

图 9-2 权限体系

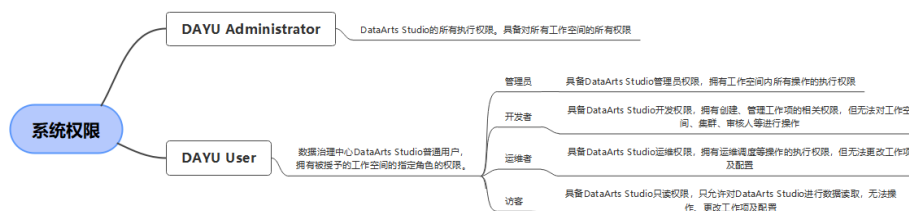


表 9-1 DataArts Studio 系统角色

系统角色名称	描述	类别
DAYU Administrator	实例管理员，拥有对DataArts Studio实例及工作空间的所有管理权限、依赖服务权限，以及所有工作空间内的所有业务操作权限。 说明 Tenant Administrator具有除统一身份认证服务外，其他所有服务的所有执行权限。即Tenant Administrator权限的用户也拥有对DataArts Studio的所有执行权限。	系统角色

系统角色名称	描述	类别
DAYU User	<p>普通用户，具备DataArts Studio实例及工作空间的查看权限，以及依赖服务权限。普通用户需要被授予任一工作空间角色后，才能拥有对应角色的业务操作权限。</p> <p>工作空间有管理员、开发者、部署者、运维者和访客五种预置角色，每种角色的介绍如下，具体操作权限请参见权限列表。</p> <ul style="list-style-type: none"> • 管理员：工作空间管理员，拥有工作空间内所有的业务操作权限。建议将项目负责人、开发责任人、运维管理员设置为管理员角色。 • 开发者：开发者拥有工作空间内创建、管理工作项的业务操作权限。建议将任务开发、任务处理的用户设置为开发者。 • 运维者：运维者具备工作空间内运维调度等业务的操作权限，但无法更改工作项及配置。建议将运维管理、状态监控的用户设置为运维者。 • 访客：访客可以查看工作空间内的数据，但无法操作业务。建议将只查看空间内容、不进行操作的用户设置为访客。 • 部署者：企业模式独有，具备工作空间内任务包发布的相关操作权限。在企业模式中，开发者提交脚本或作业版本后，系统会对应产生发布任务。开发者确认发包后，需要部署者审批通过，才能将修改后的作业同步到生产环境。 	系统角色

9.4 数据保护技术

收集范围

为了确保您的个人数据（例如用户名、密码、手机号码等）不被未经过认证、授权的实体或者个人获取，DataArts Studio对用户数据的存储和传输进行加密保护，以防止个人数据泄露，保证您的个人数据安全。

- 用户名：用于审批，消息通知，作为用户资产的属性存在，获取委托token的key值以及国际化。
- 用户IP：用于作为用户资产的唯一标识的必要信息存在。
- 手机号：用于审批及消息通知。
- 邮箱：用于审批及消息通知。
- 创建数据库连接的相关信息：包括用于创建数据库类型数据源所需的数据库IP、端口、用户名、密码和密钥对密码。
- 创建DIS以及对象存储类型连接的相关信息：包括用于创建DIS/对象存储类型数据连接所需的Access key、Secret Access Key。

数据存储安全

DataArts Studio通过加密算法对用户个人敏感数据加密后进行存储。对于用户敏感信息，在服务购买周期到期后14天后将会删除。

数据传输安全

用户个人敏感数据（包括密码）将通过TLS 1.2、TLS1.3进行传输中加密，所有DataArts Studio的API调用都支持 HTTPS 来对传输中的数据进行加密。

9.5 审计与日志

云审计服务（Cloud Trace Service，以下简称CTS），是华为云安全解决方案中专业的日志审计服务，提供对各种云资源操作记录的收集、存储和查询功能，可用于支撑安全分析、合规审计、资源跟踪和问题定位等常见应用场景。

CTS可记录的DataArts Studio操作列表详见[如何查看审计日志](#)。用户[开通云审计服务](#)并创建和配置追踪器后，CTS开始记录操作事件用于审计，用户可查看CTS保存最近7天的审计日志。

CTS支持[配置关键操作通知](#)。例如用户可将与IAM相关的高危敏感操作，作为关键操作加入到CTS的实时监控列表中进行监控跟踪。当用户使用DataArts Studio服务时，如果触发了监控列表中的关键操作，那么CTS会在记录操作日志的同时，向相关订阅者实时发送通知。

9.6 服务韧性

DataArts Studio通过AZ故障容错、过载保护、备份恢复等技术方案，保障数据的持久性和可靠性。

9.7 监控安全风险

DataArts Studio提供基于云监控服务CES的资源监控能力，帮助用户监控账号下的DataArts Studio数据集成模块中的CDM集群，执行自动实时监控、告警和通知操作。用户可以实时掌握集群运行中所产生的网络流入速率、网络流出速率、CPU使用率、内存使用率、磁盘利用率、失败作业率等信息。

关于DataArts Studio支持的监控指标，以及如何创建监控告警规则等内容，请参见[查看监控指标](#)。

9.8 故障恢复

DataArts Studio云服务每天会定时全量备份数据库数据，当服务故障后，通过备份恢复等技术方案支持数据恢复。

9.9 更新管理

更新漏洞

DataArts Studio云服务通过华为云安全公告密切跟踪漏洞，如Apache Log4j2 远程代码执行漏洞（CVE-2021-44228）、Fastjson存在反序列化漏洞

(CNVD-2022-40233)等，如发现服务模块涉及漏洞影响，会迅速通过官方解决方案升级现网更新漏洞。

更新配置

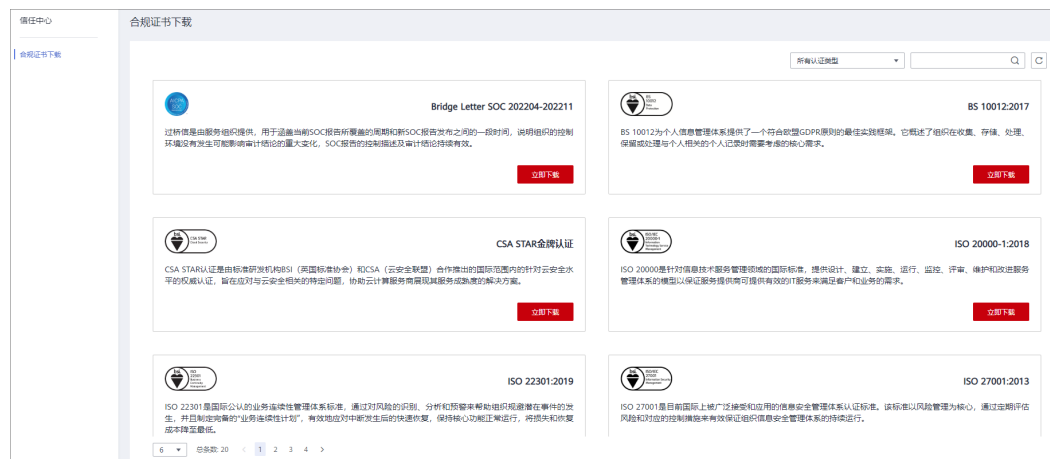
DataArts Studio云服务通过版本更新升级更新配置。

9.10 认证证书

合规证书

华为云服务及平台通过了多项国内外权威机构（ISO/SOC/PCI等）的安全合规认证，用户可自行[申请下载](#)合规资质证书。

图 9-3 合规证书下载



资源中心

华为云还提供以下资源来帮助用户满足合规性要求，具体请查看[资源中心](#)。

图 9-4 资源中心



10 权限管理

如果您需要对购买的DataArts Studio资源，给企业中的员工设置不同的访问权限，以达到不同员工之间的权限隔离，您可以使用统一身份认证服务（Identity and Access Management，简称IAM）进行精细的权限管理。该服务提供用户身份认证、权限分配、访问控制等功能，可以帮助您安全的控制华为云资源的访问。

通过IAM，您可以在华为云账号中给员工创建IAM用户，并授权来控制他们对华为云资源的访问范围。例如您的员工中有负责软件开发的人员，您希望他们拥有DataArts Studio的使用权限，但是不希望他们拥有删除工作空间等高危操作的权限，那么您可以使用IAM为开发人员创建用户，通过授予仅能使用DataArts Studio服务，但是不允许删除工作空间的权限，控制他们对DataArts Studio资源的使用范围。

IAM是华为云提供权限管理的基础服务，无需付费即可使用，您只需要为您账号中的资源进行付费。关于IAM的详细介绍，请参见[IAM产品介绍](#)。

DataArts Studio 权限

默认情况下，管理员创建的IAM用户没有任何权限，需要将其加入用户组，并给用户组授予策略或角色，才能使得用户组中的用户获得对应的权限，这一过程称为授权。授权后，用户就可以基于被授予的权限对云服务进行操作。

DataArts Studio部署时通过物理区域划分，为项目级服务。授权时，“作用范围”需要选择“区域级项目”，然后在指定区域对应的项目中设置相关权限，并且该权限仅对此项目生效；如果在“所有项目”中设置权限，则该权限在所有区域项目中都生效。访问DataArts Studio时，需要先切换至授权区域。

DataArts Studio**仅支持基于系统角色的授权，不支持策略授权**。为了实现精细的权限管控，DataArts Studio提供了**系统角色+工作空间角色**授权的能力，由工作空间角色授权具体的操作权限，并支持自定义不同权限点的工作空间角色。

说明

IAM提供了以下两种授权机制。注意，DataArts Studio仅支持其中的IAM角色方式，不支持IAM策略。

- **IAM角色**：IAM最初提供的一种根据用户的工作职能定义权限的粗粒度授权机制。该机制以服务为粒度，提供有限的服务相关角色用于授权。传统的IAM角色并不能满足用户对精细化授权的要求，无法完全达到企业对权限最小化的安全管控要求。
- **IAM策略**：IAM最新提供的一种细粒度授权的能力，可以精确到具体服务的操作、资源以及请求条件等。基于策略的授权是一种更加灵活的授权方式，能够满足企业对权限最小化的安全管控要求。

如图10-1和表10-1所示，DataArts Studio的IAM系统角色包括DAYU Administrator和DAYU User；工作空间角色是基于IAM角色DAYU User进一步授予的，**权限列表**列出了DataArts Studio常用操作与工作空间角色的授权关系，您可以参照这些权限列表选择合适的角色。

图 10-1 权限体系

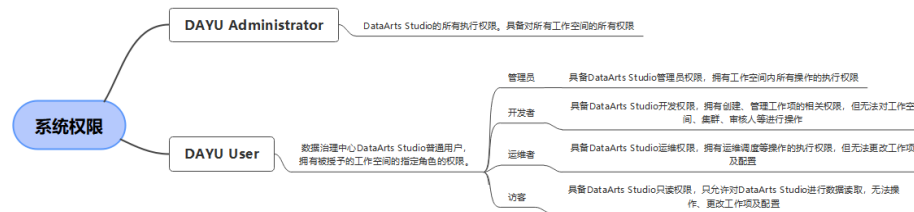


表 10-1 DataArts Studio 系统角色

系统角色名称	描述	类别
DAYU Administrator	<p>实例管理员，拥有对DataArts Studio实例及工作空间的所有管理权限、依赖服务权限，以及所有工作空间内的所有业务操作权限。</p> <p>说明 Tenant Administrator具有除统一身份认证服务外，其他所有服务的所有执行权限。即Tenant Administrator权限的用户也拥有对DataArts Studio的所有执行权限。</p>	系统角色
DAYU User	<p>普通用户，具备DataArts Studio实例及工作空间的查看权限，以及依赖服务权限。普通用户需要被授予任一工作空间角色后，才能拥有对应角色的业务操作权限。</p> <p>工作空间有管理员、开发者、部署者、运维者和访客五种预置角色，每种角色的介绍如下，具体操作权限请参见权限列表。</p> <ul style="list-style-type: none"> ● 管理员：工作空间管理员，拥有工作空间内所有的业务操作权限。建议将项目负责人、开发责任人、运维管理员设置为管理员角色。 ● 开发者：开发者拥有工作空间内创建、管理工作项的业务操作权限。建议将任务开发、任务处理的用户设置为开发者。 ● 运维者：运维者具备工作空间内运维调度等业务的操作权限，但无法更改工作项及配置。建议将运维管理、状态监控的用户设置为运维者。 ● 访客：访客可以查看工作空间内的数据，但无法操作业务。建议将只查看空间内容、不进行操作的用户设置为访客。 ● 部署者：企业模式独有，具备工作空间内任务包发布的相关操作权限。在企业模式中，开发者提交脚本或作业版本后，系统会对应产生发布任务。开发者确认发包后，需要部署者审批通过，才能将修改后的作业同步到生产环境。 	系统角色

DataArts Studio 控制台功能依赖的角色或策略

DataArts Studio服务各组件功能所需依赖服务的权限如表10-2所示。在实际授权场景中，推荐为开发者用户配置DataArts Studio服务级别的依赖服务最小权限（可参考[如何最小化授权IAM用户使用DataArts Studio](#)，为用户配置最小权限），开发者用户的最小依赖服务权限如表10-3所示。

说明

在实际授权场景中，DAYU Administrator和DAYU User系统角色已经预置了依赖服务的管理员权限。为了避免普通用户/用户组被授予DAYU User系统角色导致其拥有的依赖服务权限过大的风险，您可以在为用户组授权DAYU User系统角色后，手动删除用户组的周边依赖权限，再为用户组授予所需依赖服务的最小权限合集。

表 10-2 DataArts Studio 控制台依赖服务的角色或策略

控制台功能	依赖服务	需配置角色/策略	具体功能
管理中心	BSS	bss:coupon:view bss:renewal:update bss:discount:view bss:order:view bss:order:pay bss:order:update	创建增量包或DataArts Studio实例
	KMS	kms:cmk:get kms:cmk:list kms:cmk:create kms:cmk:decrypt kms:cmk:encrypt kms:dek:create kms:dek:encrypt kms:dek:decrypt	创建数据连接时，使用KMS加解密
	DWS	dws:cluster:list dws:cluster:getDetail dws:openAPICluster:getDetail	创建DWS数据连接
	MRS	mrs:cluster:get mrs:cluster:list	创建MRS数据连接
	VPC	vpc:publicips:get vpc:publicips:list vpc:vpcs:get vpc:port:get vpc:subnets:get	创建MRS数据连接

控制台功能	依赖服务	需配置角色/策略	具体功能
	RDS	rds:*.get rds:*.list	创建RDS数据连接
数据集成	VPC	vpc:publicIps:get vpc:publicIps:list vpc:vpcs:get vpc:vpcs:list vpc:port:get vpc:subnets:get vpc:securityGroups:get vpc:firewalls:list vpc:routeTables:list vpc:subNetworkInterfaces:list	创建CDM集群或DataArts Studio实例
	ECS	ecs:flavors:get ecs:cloudServerFlavors:get ecs:availabilityZones:list	创建CDM集群或DataArts Studio实例
	CDM	cdm:cluster:create	创建CDM集群
	KMS	kms:cmk:get kms:cmk:list kms:cmk:create kms:cmk:decrypt kms:cmk:encrypt kms:dek:create kms:dek:encrypt kms:dek:decrypt	创建数据连接时，使用KMS加解密
	MRS	mrs:cluster:get mrs:cluster:list mrs:job:get mrs:job:list	创建MRS数据连接
	DWS	dws:cluster:list dws:cluster:getDetail dws:openAPICluster:getDetail	创建DWS数据连接

控制台功能	依赖服务	需配置角色/策略	具体功能
	CDM	cdm:cluster:get cdm:cluster:list cdm:link:operate cdm:job:operate	通过CDM控制台操作时，需要CDM服务权限
	CES	ces:*.get ces:*.list	查看CES监控
	CSS	css:*.get css:*.list	创建CSS连接
	CloudTable	cloudtable:*.get cloudtable:*.list	创建CloudTable连接
	RDS	rds:*.get rds:*.list	创建RDS连接
	RMS	rms:resources:list	创建CDM集群
数据开发	OBS	obs:object:GetObject obs:object:PutObject obs:bucket:GetBucketLocation obs:bucket:ListAllMyBuckets obs:bucket:ListBucket obs:bucket:CreateBucket	运行脚本、运行作业以及备份作业
	SMN	smn:topic:publish smn:topic:list	作业通知
	KMS	kms:cmk:get kms:cmk:list kms:cmk:create kms:cmk:decrypt kms:cmk:encrypt kms:dek:create kms:dek:encrypt kms:dek:decrypt	创建数据连接时，使用KMS加解密

控制台功能	依赖服务	需配置角色/策略	具体功能
	MRS	mrs:cluster:get mrs:cluster:list mrs:job:submit mrs:job:delete mrs:job:stop mrs:sql:execute mrs:sql:cancel mrs:job:get mrs:job:list	MRS类型作业节点运行： MRS Presto SQL、MRS Spark、MRS Spark Python、MRS Flink Job、MRS MapReduce MRS Spark SQL、MRS Hive SQL
	DLI	dli:queue:submitJob dli:jobs:create dli:jobs:update dli:jobs:get dli:jobs:list dli:jobs:listAll	DLI类型作业节点运行： DLI SQL、DLI Spark
	OBS	obs:object:GetObject obs:object:PutObject obs:object:DeleteObject obs:bucket:GetBucketLocation obs:bucket>ListAllMyBuckets obs:bucket>ListBucket obs:bucket>ListBucketVersions obs:bucket>CreateBucket obs:bucket>DeleteBucket	OBS类型作业节点运行： Create OBS、Delete OBS、OBS Manager
	DWS	dws:cluster:list dws:cluster:getDetail dws:openAPICluster:getDetail	创建DWS数据连接

控制台功能	依赖服务	需配置角色/策略	具体功能
	CDM	cdm:cluster:get cdm:cluster:list cdm:job:operate	数据连接需要Agent的相关脚本、作业，以及CDM作业运行： RDS SQL、DWS SQL、Hive SQL、SPARK SQL、Shell、Python
	CES	ces:metricData:list	运维概览，查询DLI队列CPU
	GES	ges:graph:access ges:graph:operate ges:graph:list ges:graph:getDetail ges:metadata:create ges:metadata:operate ges:metadata:delete ges:metadata:list ges:metadata:getDetail ges:jobs:list ges:jobs:getDetail	Import GES作业节点运行
	ECS	ecs:servers:list ecs:servers:get ecs:servers:stop ecs:servers:start ecs:cloudServers:list	Open/Close Resource作业节点运行，创建主机连接

控制台功能	依赖服务	需配置角色/策略	具体功能
	DLI	dli:queue:submitJob dli:queue:cancelJob dli:group:useGroup dli:group:getGroup dli:group:updateGroup dli:group:deleteGroup dli:group:listAllGroup dli:database:createDatabase dli:database:dropDatabase dli:database:displayDatabase dli:database:displayAllDatabases dli:database:explain dli:database:createView dli:database:createTable dli:database:displayAllTables dli:database:createFunction dli:database:describeFunction dli:database:showFunctions dli:database:dropFunction dli:table:select dli:table:update dli:table:delete dli:table:dropTable dli:table:describeTable dli:table:showCreateTable	DLI类型作业/脚本运行

控制台功能	依赖服务	需配置角色/策略	具体功能
		dli:table:showPartitions dli:table:showSegments dli:table:showTableProperties dli:table:insertOverwriteTable dli:table:insertIntoTable dli:table:compaction dli:table:truncateTable dli:table:alterView dli:table:alterTableName dli:table:alterTableAddColumns dli:table:alterTableDropColumns dli:table:alterTableChangeColumn dli:table:alterTableSetLocation dli:table:alterTableAddPartition dli:table:alterTableNamePartition dli:table:alterTableSetProperties dli:table:alterTableRecoverPartition dli:table:alterTableDropPartition dli:column:select dli:jobs:create dli:jobs:delete dli:jobs:start dli:jobs:stop dli:jobs:update dli:jobs:export	

控制台功能	依赖服务	需配置角色/策略	具体功能
		dli:jobs:get dli:jobs:list dli:jobs:listAll dli:resource:useResource dli:resource:updateResource dli:resource:deleteResource dli:resource:getResource dli:resource:listAllResource dli:variable:update dli:variable:delete	
	IAM	iam:agencies:listAgencies	获取作业委托
	DIS	DIS Operator DIS User	DIS类型作业节点运行： DIS Stream、DIS Dump、 DIS Client
	SWR	SWR Admin	仅当在数据开发组件作业中使用 DLI Spark 节点选择自定义镜像时，需要容器镜像服务中的镜像读取权限。 推荐通过 镜像授权管理 ，添加所需镜像的读取权限。不推荐直接为用户授予SWR Admin系统角色，可能存在权限过大的风险。
数据目录	OBS	obs:object:GetObject obs:bucket:GetBucketStorage obs:bucket:GetBucketLocation obs:bucket>ListAllMyBuckets obs:bucket>ListBucket	OBS元数据采集
	DIS	dis:streams:list dis:transferTasks:list	DIS元数据采集

控制台功能	依赖服务	需配置角色/策略	具体功能
	CSS	css:cluster:list	CSS元数据采集
	GES	ges:graph:list ges:graph:getDetail ges:metadata:list ges:metadata:getDetail	GES元数据采集
	DLI	dli:database:displayDatabase dli:database:displayAllDatabases dli:table:select dli:table:describeTable dli:table:showPartitions dli:table:showTableProperties dli:jobs:create dli:jobs:get	DLI元数据采集&数据概要分析
	CDM	cdm:cluster:list	CSS元数据采集
数据质量	SMN	smn:topic:publish smn:topic:list	配置作业通知
	OBS	obs:object:GetObject obs:object:PutObject obs:bucket:GetBucketLocation obs:bucket:ListAllMyBuckets obs:bucket:ListBucket obs:bucket:CreateBucket	导出质量报告
	MRS	mrs:job:submit mrs:sql:execute mrs:sql:cancel mrs:job:get	MRS质量作业运行

控制台功能	依赖服务	需配置角色/策略	具体功能
	DLI	dli:queue:submitJob dli:jobs:get dli:jobs:listAll	DLI质量作业运行
数据安全	DLI	dli:queue:submitJob dli:queue:cancelJob dli:database:displayDatabase dli:database:displayAllDatabases dli:database:displayAllTables dli:table:describeTable dli:jobs:create dli:jobs:stop dli:jobs:get dli:resource:deleteResource dli:resource:getResource dli:resource:listAllResource	DLI权限管控
	DWS	dws:cluster:list dws:cluster:getDetail dws:openAPICluster:getDetail	DWS权限管控
	MRS	mrs:cluster:list mrs:job:submit mrs:job:stop	MRS权限管控
	KMS	kms:cmk:list kms:cmk:encrypt kms:cmk:decrypt	使用KMS加解密
	CDM	任意cdm权限，例如 cdm:cluster:get	DWS和MRS权限管控

表 10-3 开发者用户的依赖服务最小权限合集

权限类型	角色与策略权限-系统角色	角色与策略权限-自定义策略	角色与策略权限-自定义策略
是否必配	必配	必配	必配

<p>权限</p>	<ul style="list-style-type: none"> DIS Operator DIS User (可选, 不推荐) SWR Admin <p>说明</p> <p>仅当在数据开发组件作业中使用 DLI Spark 节点选择自定义镜像时, 需要容器镜像服务中的镜像读取权限。</p> <p>推荐通过镜像授权管理, 添加所需镜像的读取权限。不推荐直接为用户授予 SWR Admin 系统角色, 可能存在权限过大的风险。</p>	<p>依赖的全局级 (global级) 云服务的自定义策略</p> <p>DataArtsStudio_PermissionsOfDependentServices_global:</p> <pre>{ "Version": "1.1", "Statement": [{ "Effect": "Allow", "Action": ["obs:object:GetObject", "obs:object:PutObject", "obs:object:DeleteObject", "obs:bucket:GetBucketStorage", "obs:bucket:GetBucketLocation", "obs:bucket:ListAllMyBuckets", "obs:bucket:ListBucket", "obs:bucket:ListBucketVersions", "obs:bucket:CreateBucket", "obs:bucket:DeleteBucket", "rms:resources:list", "iam:agencies:listAgencies"] }] }</pre>	<p>依赖的项目级 (region级) 云服务的自定义策略</p> <p>DataArtsStudio_PermissionsOfDependentServices_region:</p> <pre>{ "Version": "1.1", "Statement": [{ "Effect": "Allow", "Action": ["cdm:cluster:get", "cdm:cluster:list", "cdm:cluster:create", "cdm:link:operate", "cdm:job:operate", "ces:*:get", "ces:*:list", "cloudtable:*:get", "cloudtable:*:list", "css:*:get", "css:*:list", "dis:streams:list", "dis:transferTasks:list", "dli:queue:submitJob", "dli:queue:cancelJob", "dli:table:insertOverwriteTable", "dli:table:insertIntoTable", "dli:table:alterView", "dli:table:alterTableRename", "dli:table:compaction", "dli:table:truncateTable", "dli:table:alterTableDropColumns", "dli:table:alterTableSetProperties", "dli:table:alterTableChangeColumn", "dli:table:showSegments", "dli:table:alterTableRecoverPartition", "dli:table:dropTable", "dli:table:update", "dli:table:alterTableDropPartition", "dli:table:alterTableAddPartition", "dli:table:alterTableAddColumns", "dli:table:alterTableRenamePartition", "dli:table:delete", "dli:table:alterTableSetLocation", "dli:table:describeTable", "dli:table:showPartitions", "dli:table:showCreateTable", "dli:table:showTableProperties", "dli:table:select", "dli:resource:updateResource", "dli:resource:useResource", "dli:resource:getResource",] }] }</pre>
------------------	---	---	--

			<p>"dli:resource:listAllResource",</p> <p>"dli:resource:deleteResource", "dli:database:explain",</p> <p>"dli:database:createDatabase",</p> <p>"dli:database:dropFunction",</p> <p>"dli:database:createFunction",</p> <p>"dli:database:displayAllDatabases", "dli:database:displayAllTables",</p> <p>"dli:database:displayDatabase",</p> <p>"dli:database:describeFunction", "dli:database:createView", "dli:database:createTable",</p> <p>"dli:database:showFunctions",</p> <p>"dli:database:dropDatabase", "dli:group:useGroup", "dli:group:updateGroup", "dli:group:listAllGroup", "dli:group:getGroup", "dli:group:deleteGroup", "dli:column:select", "dli:jobs:start", "dli:jobs:export", "dli:jobs:update", "dli:jobs:list", "dli:jobs:listAll", "dli:jobs:get", "dli:jobs:delete", "dli:jobs:create", "dli:jobs:stop", "dli:variable:update", "dli:variable:delete", "dws:cluster:list", "dws:cluster:getDetail",</p> <p>"dws:openAPICluster:getDetail", "ecs:servers:get", "ecs:servers:list", "ecs:servers:stop", "ecs:servers:start", "ecs:flavors:get",</p> <p>"ecs:cloudServerFlavors:get", "ecs:cloudServers:list", "ecs:availabilityZones:list", "ges:graph:access", "ges:metadata:create", "ges:jobs:list", "ges:graph:operate", "ges:jobs:getDetail", "ges:graph:getDetail", "ges:graph:list", "ges:metadata:list", "ges:metadata:getDetail", "ges:metadata:delete", "ges:metadata:operate", "kms:cmk:get", "kms:cmk:list", "kms:cmk:create", "kms:cmk:decrypt",</p>
--	--	--	---

			<pre>"kms:cmk:encrypt", "kms:dek:create", "kms:dek:encrypt", "kms:dek:decrypt", "mrs:cluster:get", "mrs:cluster:list", "mrs:job:get", "mrs:job:list", "mrs:job:submit", "mrs:job:stop", "mrs:job:delete", "mrs:sql:execute", "mrs:sql:cancel", "rds:*:get", "rds:*:list", "smn:topic:publish", "smn:topic:list", "vpc:publicIps:list", "vpc:publicIps:get", "vpc:vpcs:get", "vpc:vpcs:list", "vpc:port:get", "vpc:subnets:get", "vpc:securityGroups:get", "vpc:firewalls:list", "vpc:routeTables:list", "vpc:subNetworkInterfaces:list"] }]</pre>
--	--	--	--

相关链接

- [IAM产品介绍](#)
- [DataArts Studio细粒度授权实践](#)

11 权限列表

工作空间成员共有管理员、开发者、部署者、运维者和访客五种预置角色。

- **管理员**：工作空间管理员，拥有工作空间内所有的业务操作权限。建议将项目负责人、开发责任人、运维管理员设置为管理员角色。
- **开发者**：开发者拥有工作空间内创建、管理工作项的业务操作权限。建议将任务开发、任务处理的用户设置为开发者。
- **运维者**：运维者具备工作空间内运维调度等业务的操作权限，但无法更改工作项及配置。建议将运维管理、状态监控的用户设置为运维者。
- **访客**：访客可以查看工作空间内的数据，但无法操作业务。建议将只查看空间内容、不进行操作的用户设置为访客。
- **部署者**：企业模式独有，具备工作空间内任务包发布的相关操作权限。在企业模式中，开发者提交脚本或作业版本后，系统会对应产生发布任务。开发者确认发包后，需要部署者审批通过，才能将修改后的作业同步到生产环境。

本章节将为您介绍各预置角色的权限差异说明。

须知

后文权限点中的“操作”权限，指的是非增删改查的资源操作，例如导入、导出、任务执行、任务取消、任务启动、开始调度等操作。

管理中心

权限点	管理员	开发者	运维者	访客
查询集群列表 (MRS&DWS&CD M)	Y	Y	Y	Y
新建数据库	Y	Y	N	N
删除数据库	Y	Y	N	N
查询数据库	Y	Y	Y	Y
编辑数据库	Y	Y	N	N

权限点	管理员	开发者	运维者	访客
新建数据表	Y	Y	N	N
删除数据表	Y	Y	N	N
查询数据表	Y	Y	Y	Y
编辑数据表	Y	Y	N	N
新建资源迁移	Y	Y	N	N
操作资源迁移	Y	Y	Y	N
查询资源迁移	Y	Y	Y	Y
新建数据连接	Y	Y	N	N
删除数据连接	Y	Y	N	N
操作数据连接	Y	Y	Y	N
查询数据连接	Y	Y	Y	Y
编辑数据连接	Y	Y	N	N
删除RDS驱动包	Y	N	N	N
操作RDS驱动包	Y	N	N	N
查询RDS驱动包	Y	Y	Y	Y
新建DLI资源映射配置	N	N	N	N
删除DLI资源映射配置	N	N	N	N
查询DLI资源映射配置	N	N	N	N

数据架构

权限点	管理员	开发者	运维者	访客
新建原子指标	Y	Y	N	N
删除原子指标	Y	Y	N	N
查询原子指标	Y	Y	Y	Y
编辑原子指标	Y	Y	N	N
新建逻辑实体/物理表	Y	Y	N	N

权限点	管理员	开发者	运维者	访客
删除逻辑实体/物理表	Y	Y	N	N
查询逻辑实体/物理表	Y	Y	Y	Y
编辑逻辑实体/物理表	Y	Y	N	N
新建配置中心	Y	N	N	N
删除配置中心	Y	N	N	N
查询配置中心	Y	Y	Y	Y
编辑配置中心	Y	N	N	N
新建主题设计	Y	Y	N	N
删除主题设计	Y	Y	N	N
查询主题设计	Y	Y	Y	Y
编辑主题设计	Y	Y	N	N
新建业务指标	Y	Y	N	N
删除业务指标	Y	Y	N	N
查询业务指标	Y	Y	Y	Y
编辑业务指标	Y	Y	N	N
新建汇总表	Y	Y	N	N
删除汇总表	Y	Y	N	N
查询汇总表	Y	Y	Y	Y
编辑汇总表	Y	Y	N	N
新建通用配置	Y	Y	N	N
删除通用配置 (指已发布状态的逻辑实体或表, 对其在编辑并保存后产生的草稿的删除操作)	Y	Y	N	N
操作通用配置 (指逻辑实体或表的导入、导出、发布、下线、同步、逆向等操作)	Y	Y	Y	N

权限点	管理员	开发者	运维者	访客
查询通用配置 (指已发布状态的逻辑实体或表, 对其在编辑并保存后产生的草稿的查询操作)	Y	Y	Y	Y
编辑通用配置 (指已发布状态的逻辑实体或表, 对其在编辑并保存后产生的草稿的编辑操作)	Y	Y	N	N
删除维度表	Y	Y	N	N
查询维度表	Y	Y	Y	Y
新建流程设计	Y	Y	N	N
删除流程设计	Y	Y	N	N
查询流程设计	Y	Y	Y	Y
编辑流程设计	Y	Y	N	N
新建码表管理	Y	Y	N	N
删除码表管理	Y	Y	N	N
查询码表管理	Y	Y	Y	Y
编辑码表管理	Y	Y	N	N
新建模型	Y	Y	N	N
删除模型	Y	Y	N	N
查询模型	Y	Y	Y	Y
编辑模型	Y	Y	N	N
新建衍生/复合指标	Y	Y	N	N
删除衍生/复合指标	Y	Y	N	N
操作衍生/复合指标	Y	Y	N	N
查询衍生/复合指标	Y	Y	Y	Y
编辑衍生/复合指标	Y	Y	N	N

权限点	管理员	开发者	运维者	访客
新建关联质量规则	Y	Y	N	N
删除关联质量规则	Y	Y	N	N
查询关联质量规则	Y	Y	Y	Y
编辑关联质量规则	Y	Y	N	N
新建事实表	Y	Y	N	N
删除事实表	Y	Y	N	N
查询事实表	Y	Y	Y	Y
编辑事实表	Y	Y	N	N
新建目录	Y	Y	N	N
删除目录	Y	Y	N	N
查询目录	Y	Y	Y	Y
编辑目录	Y	Y	N	N
新建维度	Y	Y	N	N
删除维度	Y	Y	N	N
查询维度	Y	Y	Y	Y
编辑维度	Y	Y	N	N
新建时间限定	Y	Y	N	N
删除时间限定	Y	Y	N	N
查询时间限定	Y	Y	Y	Y
编辑时间限定	Y	Y	N	N
新建数据标准	Y	Y	N	N
删除数据标准	Y	Y	N	N
查询数据标准	Y	Y	Y	Y
编辑数据标准	Y	Y	N	N

数据集成

权限点	管理员	开发者	运维者	访客
新建集群	Y	Y	N	N
删除集群	Y	Y	N	N

权限点	管理员	开发者	运维者	访客
操作集群	Y	Y	Y	N
查询集群	Y	Y	Y	Y
编辑集群	Y	Y	N	N
操作连接	Y	Y	Y	N
查询连接	N	N	N	N
操作作业	Y	Y	Y	N
查询作业	N	N	N	N

数据开发

权限点	管理员	开发者	部署者	运维者	访客
新建Schema	Y	Y	N	N	N
删除Schema	Y	Y	N	N	N
查询Schema	Y	Y	N	Y	Y
编辑Schema	Y	Y	N	N	N
操作备份	Y	Y	N	Y	N
查询备份	Y	Y	N	Y	Y
新建补数据	Y	Y	N	N	N
操作补数据	Y	Y	N	Y	N
查询补数据	Y	Y	N	Y	Y
操作脏数据	Y	Y	N	Y	N
操作恢复备份	Y	N	N	Y	N
查询恢复备份	Y	Y	N	Y	Y
新建目录	Y	Y	N	N	N
删除目录	Y	Y	N	N	N
查询目录	Y	Y	N	Y	Y
编辑目录	Y	Y	N	N	N
新建通知	Y	Y	N	N	N
删除通知	Y	Y	N	N	N
查询通知	Y	Y	N	Y	Y

权限点	管理员	开发者	部署者	运维者	访客
编辑通知	Y	Y	N	N	N
新建数据库	Y	Y	N	N	N
删除数据库	Y	Y	N	N	N
查询数据库	Y	Y	N	Y	Y
编辑数据库	Y	Y	N	N	N
新建解决方案	Y	Y	N	N	N
删除解决方案	Y	Y	N	N	N
操作解决方案	Y	Y	N	Y	N
查询解决方案	Y	Y	N	Y	Y
编辑解决方案	Y	Y	N	N	N
查询IAM代理	Y	Y	Y	Y	Y
更新IAM代理	Y	N	N	N	N
操作环境变量	Y	Y	N	N	N
查询环境变量	Y	Y	N	Y	Y
编辑环境变量	Y	Y	N	N	N
操作作业节点	Y	Y	N	Y	N
查看发布包	Y	Y	Y	Y	Y
操作发布包	Y	N	Y	Y	N
新建数据连接	Y	Y	N	N	N
删除数据连接	Y	Y	N	N	N
操作数据连接	Y	Y	N	Y	N
查询数据连接	Y	Y	N	Y	Y
编辑数据连接	Y	Y	N	N	N
撤销发布	Y	Y	Y	Y	N
新建数据表	Y	Y	N	N	N
删除数据表	Y	Y	N	N	N
查询数据表	Y	Y	N	Y	Y
编辑数据表	Y	Y	N	N	N
操作作业实例	Y	Y	N	Y	N
查询作业实例	Y	Y	N	Y	Y

权限点	管理员	开发者	部署者	运维者	访客
新建资源	Y	Y	N	N	N
删除资源	Y	Y	N	N	N
操作资源	Y	Y	N	Y	N
查询资源	Y	Y	N	Y	Y
编辑资源	Y	Y	N	N	N
编辑环境变量映射	N	N	N	N	N
操作脚本编辑锁	Y	Y	N	Y	N
新建脚本	Y	Y	N	N	N
删除脚本	Y	Y	N	N	N
操作脚本	Y	Y	N	Y	N
查询脚本	Y	Y	N	Y	Y
编辑脚本	Y	Y	N	N	N
新建作业标签	Y	Y	N	Y	N
删除作业标签	Y	Y	N	Y	N
查询作业标签	Y	Y	N	Y	Y
新建作业	Y	Y	N	N	N
删除作业	Y	Y	N	N	N
操作作业	Y	Y	N	Y	N
查询作业	Y	Y	N	Y	Y
编辑作业	Y	Y	N	Y	N
查询作业编辑锁详情	Y	Y	N	Y	Y
操作作业编辑锁	Y	Y	N	Y	N
创建基线	Y	N	N	Y	N
查询基线	Y	Y	N	Y	N
删除基线	Y	N	N	Y	N
修改基线	Y	N	N	Y	N
查询基线实例	Y	Y	N	Y	N
获取保障作业概要信息列表	Y	Y	N	Y	N

权限点	管理员	开发者	部署者	运维者	访客
查询事件	Y	Y	N	Y	N
更新事件	Y	N	N	Y	N

数据质量

权限点	管理员	开发者	运维者	访客
数据质量监控				
查询总览	Y	Y	Y	Y
操作实例	Y	Y	Y	N
查询实例	Y	Y	Y	Y
新建规则模板	Y	N	N	N
删除规则模板	Y	N	N	N
操作规则模板	Y	N	N	N
查询规则模板	Y	Y	Y	Y
编辑规则模板	Y	N	N	N
查询运行结果	Y	Y	N	N
新建规则	Y	Y	N	N
删除规则	Y	Y	N	N
操作规则	Y	Y	Y	N
查询规则	Y	Y	Y	Y
编辑规则	Y	Y	N	N
编辑质量评分	Y	N	N	N
新建目录	Y	Y	N	N
删除目录	Y	Y	N	N
查询目录	Y	Y	Y	Y
编辑目录	Y	Y	N	N
业务指标监控				
查询总览	Y	Y	Y	Y
操作实例	Y	Y	Y	N
查询实例	Y	Y	Y	Y

权限点	管理员	开发者	运维者	访客
新建业务场景	Y	Y	N	N
删除业务场景	Y	Y	N	N
操作业务场景	Y	Y	Y	N
查询业务场景	Y	Y	Y	Y
编辑业务场景	Y	Y	N	N
新建指标	Y	Y	N	N
删除指标	Y	Y	N	N
查询指标	Y	Y	Y	Y
编辑指标	Y	Y	N	N
新建规则	Y	Y	N	N
删除规则	Y	Y	N	N
查询规则	Y	Y	Y	Y
编辑规则	Y	Y	N	N
新建目录	Y	Y	N	N
删除目录	Y	Y	N	N
查询目录	Y	Y	Y	Y
编辑目录	Y	Y	N	N

数据目录

权限点	管理员	开发者	运维者	访客
查询数据源数据	Y	Y	Y	N
操作任务实例	Y	Y	Y	N
查询任务实例	Y	Y	Y	Y
新建采集任务	Y	Y	N	N
删除采集任务	Y	Y	N	N
操作采集任务	Y	Y	Y	N
查询采集任务	Y	Y	Y	Y
编辑采集任务	Y	Y	N	N
编辑审批单	Y	Y	N	N

权限点	管理员	开发者	运维者	访客
编辑资产报告	Y	Y	N	N
新建标签	Y	Y	N	N
删除标签	Y	Y	N	N
查询标签	Y	Y	Y	Y
编辑标签	Y	Y	N	N
新建资产	Y	Y	N	N
删除资产	Y	Y	N	N
操作资产	Y	Y	Y	N
查询资产	Y	Y	Y	Y
编辑资产	Y	Y	N	N
新建目录	Y	Y	N	N
删除目录	Y	Y	N	N
查询目录	Y	Y	Y	Y
编辑目录	Y	Y	N	N
新建分类	Y	Y	N	N
删除分类	Y	Y	N	N
查询分类	Y	Y	Y	Y
编辑分类	Y	Y	N	N
新建数据权限规则	Y	N	N	N
删除数据权限规则	Y	N	N	N
查询数据权限规则	Y	Y	Y	Y
编辑数据权限规则	Y	N	N	N

数据服务

权限点	管理员	开发者	运维者	访客
新建流控策略	Y	Y	N	N
删除流控策略	Y	Y	N	N
操作流控策略	Y	Y	Y	N
查询流控策略	Y	Y	Y	Y

权限点	管理员	开发者	运维者	访客
编辑流控策略	Y	Y	N	N
新建应用	Y	Y	N	N
删除应用	Y	Y	N	N
操作应用	Y	Y	Y	N
查询应用	Y	Y	Y	Y
编辑应用	Y	Y	N	N
操作审核	Y	Y	Y	N
查询审核	Y	Y	Y	Y
新建API目录	Y	Y	Y	N
删除API目录	Y	Y	Y	N
查询API目录	Y	Y	Y	Y
编辑API目录	Y	Y	Y	N
操作集群	Y	Y	N	N
查询集群	Y	Y	Y	Y
新建审核人	Y	N	N	N
删除审核人	Y	N	N	N
操作审核人	Y	Y	Y	N
查询审核人	Y	Y	Y	N
新建API	Y	Y	N	N
删除API	Y	Y	N	N
操作API	Y	Y	Y	N
查询API	Y	Y	Y	Y
编辑API	Y	Y	N	N
查询数据源	Y	Y	N	N
查询总览	Y	Y	Y	Y

数据安全

权限点	管理员	开发者	运维者	访客
查询总览	Y	Y	Y	Y

权限点	管理员	开发者	运维者	访客
新建数据溯源任务	Y	Y	N	N
删除数据溯源任务	Y	Y	N	N
操作数据溯源任务	Y	Y	N	N
查询数据溯源任务	Y	Y	Y	Y
编辑数据溯源任务	Y	Y	N	N
操作安全任务调度	Y	Y	N	Y
新建敏感发现任务	Y	Y	Y	N
删除敏感发现任务	Y	Y	N	N
操作敏感发现任务	Y	Y	Y	N
查询敏感发现任务	Y	Y	Y	Y
编辑敏感发现任务	Y	Y	N	N
查询数据源	Y	Y	Y	Y
新建访问权限管理	Y	Y	N	N
删除访问权限管理	Y	Y	N	N
查询访问权限管理	Y	Y	Y	Y
编辑访问权限管理	Y	Y	N	N
查询资源权限配置	Y	Y	N	N
新建脱敏策略	Y	Y	N	N
删除脱敏策略	Y	Y	N	N
操作脱敏策略	Y	Y	Y	Y
查询脱敏策略	Y	Y	Y	Y
编辑脱敏策略	Y	Y	N	N
新建密级	Y	Y	N	N
删除密级	Y	Y	N	N
查询密级	Y	Y	Y	Y
编辑密级	Y	Y	N	N
新建规则分组	Y	Y	Y	N
删除规则分组	Y	Y	N	N
操作规则分组	Y	Y	Y	N
查询规则分组	Y	Y	Y	Y

权限点	管理员	开发者	运维者	访客
编辑规则分组	Y	Y	Y	N
新建数据脱敏任务	Y	Y	N	N
删除数据脱敏任务	Y	Y	N	N
操作数据脱敏任务	Y	Y	N	N
查询数据脱敏任务	Y	Y	Y	Y
编辑数据脱敏任务	Y	Y	N	N
新建数据水印嵌入	Y	Y	N	N
删除数据水印嵌入	Y	Y	N	N
操作数据水印嵌入	Y	Y	N	N
查询数据水印嵌入	Y	Y	Y	Y
编辑数据水印嵌入	Y	Y	N	N

12 约束与限制

浏览器限制

您需要使用支持的浏览器版本登录DataArts Studio。

表 12-1 浏览器兼容性

浏览器版本	建议版本	建议操作系统	备注
Google Chrome	115, 114, 113	Windows 10	分辨率最佳可视范围为最小1366*768px，最大为1920*1080px。其中1920*1080px为最佳显示分辨率，界面自适应为最优显示。

使用限制

使用DataArts Studio前，您需要认真阅读并了解以下使用限制。

表 12-2 DataArts Studio 使用限制一览表

组件	约束限制
公共	<ol style="list-style-type: none">1. DataArts Studio必须基于华为云底座部署。资源隔离场景下，支持以全栈专属云模式部署，另外也支持以华为云Stack和HCS Online混合云模式部署。 关于全栈专属云、华为云Stack和HCS Online的适用场景和差异等更多信息，欢迎通过咨询了解。2. DataArts Studio基于数据湖底座提供数据一站式集成、开发、治理等能力，本身不具备存储和计算的能力，需要配合数据湖底座使用。3. 每个企业项目下最多绑定一个DataArts Studio实例。当企业项目下已绑定实例时，再次购买实例会失败。4. DataArts Studio各组件对不同数据源的支持程度不一，您需要按照您的业务需求来选择数据湖底座。DataArts Studio平台当前支持的数据湖产品请参见DataArts Studio支持的数据源。

组件	约束限制
管理中心	<ol style="list-style-type: none"> 1. 由于管理中心的限制，数据治理各组件（如数据架构、数据质量、数据目录等）暂不支持包含中文和“.”字符的库表名。 2. DataArts Studio实例赠送的CDM集群，由于本身规格有限，推荐仅作为DataArts Studio管理中心数据连接的Agent代理使用。 3. 建议为管理中心数据连接的Agent和CDM迁移作业规划相互独立的CDM集群，避免双方使用同一集群，导致业务高峰期时资源抢占引起业务不可用。 4. CDM集群作为管理中心数据连接Agent时，单集群无法连接多个MRS安全集群。建议您按照业务情况规划多个Agent与MRS安全集群一一映射。 5. CDM集群作为管理中心数据连接Agent时，单集群的并发活动线程最大为200。即当多个数据连接共用同一Agent时，通过这些数据连接提交SQL脚本、Shell脚本、Python脚本等任务的同时运行上限为200，超出的任务将排队等待。建议您按照业务量情况规划多个Agent分担压力。 6. 单工作空间允许创建的数据连接个数最多200个。 7. 管理中心相关开放API并发限制为100qps。
数据集成	<ol style="list-style-type: none"> 1. CDM作业支持自动备份和恢复，将备份数据存储到OBS中，该功能需要您手动开启。详情请参见CDM作业自动备份/恢复章节。 2. 数据集成CDM集群为单集群部署，集群故障可能会导致业务、数据损失。建议您使用数据开发作业CDM Job节点调用CDM作业，并选择两个CDM集群以提升可靠性。详情请参见CDM Job节点章节。 <p>关于数据集成中的更多约束限制，请参考CDM约束与限制。</p>
数据开发	<ol style="list-style-type: none"> 1. 数据开发脚本、作业等资产支持备份管理，将备份数据存储到OBS中，该功能需要您手动开启。详情请参见备份管理章节。 2. 单工作空间允许创建的作业个数最多1万个。 3. RDS SQL、DWS SQL、HIVE SQL、SPARK SQL、DLI SQL脚本执行结果页面展示最多1千条，且数据量少于3MB。超过1千条数据可以使用转储功能，转储最多支持1万条。

组件	约束限制
数据架构	<ol style="list-style-type: none"> 1. 数据架构当前支持关系建模和维度建模（仅支持星形模型）这两种建模方式。 2. 数据架构支持最大导入文件大小为4Mb；支持最大导入指标个数为3000个；支持一次最大导出500张表。 3. 单工作空间中创建各类对象的配额如下： <ul style="list-style-type: none"> ● 主题5000个。 ● 数据标准目录500条，个数20000个。 ● 原子指标、衍生指标、复合指标各5000条。 4. 配置中心中各类对象的自定义项配额如下： <ul style="list-style-type: none"> ● 主题自定义项10条。 ● 表自定义项30条。 ● 属性自定义项10条。 ● 业务指标自定义项50条。
数据质量	<ol style="list-style-type: none"> 1. 数据质量作业执行时长依赖数据引擎，如果底层数据引擎资源不足，可能会导致运行速度变慢。 2. 单个数据质量作业最多可以配置50条规则，如有需要可拆分为多个质量作业。
数据目录	<ol style="list-style-type: none"> 1. 元数据采集任务通过执行引擎相关的DDL SQL获取，不建议单个任务采集超过1000张表。如有需要可拆分为多个采集任务，另外调度时间和频次也需要根据业务需要进行合理设置，避免对引擎造成较大的访问和连接压力，设置建议如下： <ul style="list-style-type: none"> ● 若业务对元数据时效性要求为1天，则设置调度周期=max(1天，单次采集周期时间)，其他情况同理。 ● 若业务压力集中在白天，则设置调度时间在夜间，其他情况同理，选择数据源压力最小的时间段。 2. 数据血缘的产生依赖于数据开发中调度运行的作业，测试运行的作业不会产生血缘。
数据服务	<ol style="list-style-type: none"> 1. 数据服务共享版仅供开发测试使用，专享版性能优于共享版，推荐使用数据服务专享版。 2. 数据服务集群与工作空间绑定，创建后暂不支持修改规格或升级版本。

13 与其他云服务的关系

统一身份认证服务

DataArts Studio使用统一身份认证服务（Identity and Access Management，简称IAM）实现认证和鉴权功能。

云审计服务

DataArts Studio使用云审计服务（Cloud Trace Service，简称CTS）审计用户在管理控制台页面的操作，可用于检视是否存在非法或越权操作，完善服务安全管理。

弹性云服务器服务

DataArts Studio使用弹性云服务器（Elastic Cloud Server，简称ECS）进行CDM集群和数据服务集群的创建，另外DataArts Studio可以通过主机连接在ECS上执行Shell或Python脚本。

虚拟私有云服务

DataArts Studio使用虚拟私有云服务（Virtual Private Cloud，简称VPC）的创建隔离的网络环境。

弹性公网 IP 服务

DataArts Studio使用弹性公网IP服务（Elastic IP，简称EIP）打通与公网间的网络通信。

对象存储服务

DataArts Studio使用对象存储服务（Object Storage Service，简称OBS）的桶存储日志信息。

消息通知服务

DataArts Studio使用消息通知服务（Simple Message Notification，简称SMN）依据用户的订阅需求主动推送通知消息，使用户可以在触发告警（如质量监控）时能立即接收到通知。

云专线服务

DataArts Studio使用云专线服务（Direct Connect，简称DC）打通与第三方数据中心的网络通信。

API 网关服务

DataArts Studio通过API网关服务（API Gateway，简称APIG）对外开放各组件的API接口。

数据湖探索服务

DataArts Studio支持将数据湖探索服务（Data Lake Insight，简称DLI）作为数据湖底座，进行数据集成、开发、治理与开放。

MapReduce 服务

DataArts Studio支持将MapReduce服务（简称MRS）作为数据湖底座，进行数据集成、开发与治理。

云数据仓库服务

DataArts Studio支持将云数据仓库服务（GaussDB(DWS)，简称DWS）作为数据湖底座，进行数据集成、开发、治理与开放。

云数据库服务

DataArts Studio支持将云数据库服务（Relational Database Service，简称RDS）作为数据源，进行数据集成、开发与开放。

云数据迁移服务

DataArts Studio需要通过Agent来访问租户资源或服务，需要依赖云数据迁移服务（Cloud Data Migration，简称CDM）。其中云数据迁移与数据集成成为同一服务的两种形态，云数据迁移为独立服务形态，数据集成为DataArts Studio的组件形态。