

智能数据湖 AI Datalake

用户指南

文档版本 01
发布日期 2026-04-14



版权所有 © 华为技术有限公司 2026。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。

安全声明

漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

目录

1 准备工作	1
1.1 注册华为账号并开通华为云	1
1.2 开通 AI Datalake 服务并授权	2
1.3 通过 IAM 授予使用 AI Datalake 的权限	10
1.3.1 通过 IAM 角色或策略授予使用 AI Datalake 的权限	11
1.3.2 通过 IAM 身份策略授予使用 AI Datalake 的权限	13
2 创建 AI Datalake 工作空间	16
2.1 工作空间概述	16
2.2 创建工作空间	18
2.3 管理工作空间	19
3 创建 AI Datalake 计算资源池	21
3.1 计算资源池概述	21
3.2 购买预留资源池	25
3.3 配置计算资源池网络	29
3.3.1 为什么要配置计算资源池网络	29
3.3.2 创建网络	30
3.4 调整计算资源池容量	32
3.5 管理计算资源池	34
4 创建 AI Datalake Aura 引擎端点	36
4.1 Aura 引擎端点概述	36
4.2 创建 Aura 引擎端点	40
4.3 管理 Aura 引擎端点	45
5 使用 DataArts Studio 开发 AI Datalake 作业	50
6 管理数据与元数据	53
6.1 了解数据目录、数据库和表	53
6.2 创建数据目录、数据库和表	55
7 监控 AI Datalake 计算资源池	57
8 使用 LTS 管理 AI Datalake 作业日志	59
9 使用 CTS 审计 AI Datalake 服务	63

1 准备工作

1.1 注册华为账号并开通华为云

使用AI DataLake服务之前，请先注册华为账号并开通华为云，完成实名认证并进行服务授权。

📖 说明

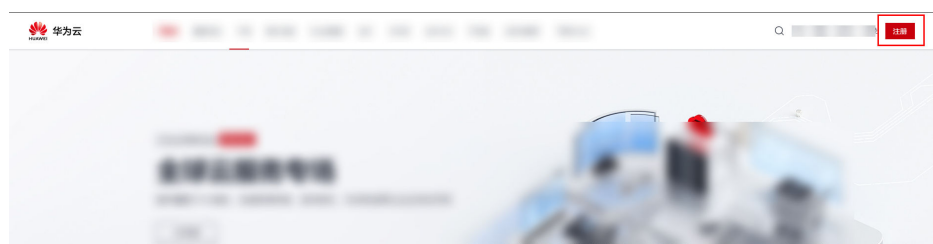
注册华为账号并开通华为云后，如需对DLI资源进行精细管理，请使用IAM服务创建IAM用户及用户组，并授权，以使得IAM用户获得具体的操作权限。

注册华为账号并开通华为云

如果您已完成华为账号注册，可跳过该步骤。

步骤1 打开注册[华为云](#)官方网站。

步骤2 单击华为云官网右上角“注册”进入注册页面。



步骤3 在注册页面，根据提示信息完成注册。具体操作可参见[账号注册](#)。

步骤4 根据页面引导勾选服务条款，单击“开通”，开通华为云。

----结束

实名认证

根据国家法律规定，所有用户必须完成实名认证后才能使用云服务。

如果您已完成实名认证，可跳过该步骤。具体操作可参见[实名认证](#)。

1.2 开通 AI DataLake 服务并授权

为保证正常使用AI DataLake服务，您需要开通AI DataLake服务，并授权允许AI DataLake服务代表用户访问其他云服务，才可以在AI DataLake服务中执行创建工作空间，开发作业等操作。

本节操作介绍如何开通AI DataLake，并设置云资源访问授权。

前提条件

您已经注册了华为账号开通华为云。具体操作可参见[账号注册](#)。实名认证指引请参考[实名认证](#)。

约束与限制

授权需要主账号或者用户组Admin中的子账号进行操作。

操作步骤

1. 配置AI DataLake云资源访问授权

- a. 登录AI DataLake服务管理控制台，进入AI DataLake服务页面。
当前页面会展示完整AI DataLake服务依赖的其他云服务资源。
为保证正常使用AI DataLake服务，请您授权允许AI DataLake服务代表用户访问其他云服务。详细的委托策略请参考[表1-2](#)。
- b. 根据需要勾选对应委托权限。

2. 阅读并同意AI DataLake服务声明

仔细阅读AI DataLake服务条款，并勾选同意该服务声明。

3. 单击“立即授权”。

同意授权后，AI DataLake将在统一身份认证服务IAM为您创建名为ai_datalake_admin_trust的委托，授权成功后，可以进入IAM服务委托列表查看该未委托的基本信息和授权记录。



注意

委托ai_datalake_admin_trust创建成功后，请勿删除，一旦删除将变更授权范围，影响相关业务权限。

AI DataLake 云资源授权列表

表 1-1 AI DataLake 云资源授权列表

策略类型	策略类型	是否必选	说明	不勾选对业务的影响
基础通用服务	AI_DATA LAKE_COMMON_POLICY	必选	委托必须的基础权限策略。	基础权限验证失败： <ul style="list-style-type: none"> 不能获取IAM用户信息。 不能读取OBS桶及数据。
访问自身服务	AI_DATA LAKE_SELF_POLICY	必选	访问AI DataLake服务所需的权限。	核心数据分析功能完全不可用 <ul style="list-style-type: none"> 不能创建/管理工作空间。 不能执行SQL查询和分析。 不能创建/查看/取消作业。
日志功能	AI_DATA LAKE_LOGS_POLICY	必选	LTS日志服务功能。	问题排查和审计追踪困难： <ul style="list-style-type: none"> 不能使用LTS日志服务。 作业运行日志无法记录和查看。
LakeFormation	AI_DATA LAKE_LAKEFORMATION_POLICY	必选	LakeFormation 数据湖功能。	数据目录功能失效： <ul style="list-style-type: none"> 不能管理数据访问权限。 不能进行数据治理和权限管控。
消息通知	AI_DATA LAKE_SMN_POLICY	可选	SMN通知消息。	作业运行完成/失败等事件无法通过SMN推送通知。
AOM 监控	AI_DATA LAKE_AOM_POLICY	可选	AOM监控告警功能。	运维监控能力缺失： <ul style="list-style-type: none"> 业运行指标无法上报至AOM。 不能使用AOM进行性能监控和告警。
镜像包注册	AI_DATA LAKE_SWR_POLICY	可选	SWR镜像包功能。	依赖特定镜像的高级功能受限： <ul style="list-style-type: none"> 不能使用SWR中的自定义镜像包。 不能注册和管理自定义运行环境。

策略类型	策略类型	是否必选	说明	不勾选对业务的影响
VPCEP	AI_DATA LAKE_V PCEP_POLICY	可选	VPCEP网络打通功能。	跨服务网络打通受限： <ul style="list-style-type: none">• 不能通过VPCEP打通私有网络。• 不能访问其他私有网络服务。
DNS 功能	AI_DATA LAKE_D NS_POLICY	可选	DNS服务功能。	跨域网络访问配置困难： <ul style="list-style-type: none">• 不能使用DNS服务进行域名解析配置。• 自定义域名访问受限。
OBS 存储	AI_DATA LAKE_O BS_POLICY	可选	OBS对象存储功能。	数据存储管理功能受限： <ul style="list-style-type: none">• 不能通过AI DataLake直接创建和管理OBS桶。• 不能上传/下载数据对象。• 不能配置桶的加密和生命周期策略。

AI DataLake将在统一身份认证服务IAM为您创建名为ai_datalake_admin_trust的委托，该委托包含的策略和Action请参考[表1-2](#)。

表 1-2 AI DataLake 云资源授权列表

策略类型	策略类型	Action	是否必选	说明
基础通用服务	AI_DATA LAKE_COMMON_POLICY	iam:agencies:listAgencies 列举委托 iam:roles:getRole 查询角色 iam:roles:listRoles 列举角色 iam:permissions:listRolesForAgency 查询委托的角色权限 iam:users:getUser 查询用户 obs:bucket:ListAllMyBuckets 列举所有桶 obs:bucket:ListBucket 列举桶 obs:object:GetObjectVersion 获取对象版本 obs:object:GetObject 获取对象	必选	委托必须的基础权限策略

策略类型	策略类型	Action	是否必选	说明
访问自身服务	AI_DATALAKE_SELF_POLICY	DataArtsFabric:workspace:list列举工作空间 DataArtsFabric:endpoint:show查看端点 DataArtsFabric:workspace:listRoute查询工作空间路由 DataArtsFabric:endpoint:list列举端点 DataArtsFabric:job:dropJobInstance删除作业实例 DataArtsFabric:job:listJobInstance列举作业实例 DataArtsFabric:workspace:showSession查看会话 DataArtsFabric:workspace:listMessagePolicy列举消息策略 DataArtsFabric:workspace:executeStatement执行语句 DataArtsFabric:workspace:showStatementResult查看语句结果 DataArtsFabric:workspace:deleteSqlSession删除SQL会话 DataArtsFabric:workspace:cancelStatement取消语句 DataArtsFabric:workspace:createSqlSession创建SQL会话 DataArtsFabric:workspace:listSessionStatementRecords列举会话语句记录 DataArtsFabric:workspace:listSqlSessions列举SQL会话 DataArtsFabric:workspace:listStatementRecords列举语句记录 DataArtsFabric:workspace:showSqlSession查看SQL会话	必选	访问AI Datalake服务所需的权限

策略类型	策略类型	Action	是否必选	说明
日志功能	AI_DATALAKE_LTS_POLICY	lts:groups:get 查询日志组 lts:groups:list 列举日志组 lts:groups:create 创建日志组 lts:topics:get 查询日志流 lts:topics:list 列举日志流 lts:topics:create 创建日志流	必选	LTS日志服务功能

策略类型	策略类型	Action	是否必选	说明
LakeFormation	AI_DATALAKE_LAKEFORMATION_POLICY	lakeformation:accessTenant:grant授予租户访问权限 lakeformation:access:delete删除访问权限 lakeformation:access:create创建访问权限 lakeformation:access:describe查询访问权限 lakeformation:table:describe查询表详情 lakeformation:credential:describe查询凭证信息 lakeformation:policy:describe查询策略信息 lakeformation:table:drop删除表 lakeformation:table:create创建表 lakeformation:policy:create创建策略 lakeformation:database:describe查询数据库详情 lakeformation:table:alter修改表属性 lakeformation:policy:describe查询策略信息 lakeformation:credential:describe查询凭证信息 lakeformation:function:describe查询函数详情 lakeformation:catalog:describe查询Catalog详情 lakeformation:database:create创建数据库 lakeformation:database:drop删除数据库 lakeformation:function:create创建函数 lakeformation:function:alter修改函数属性 lakeformation:function:drop删除函数 lakeformation:policy:drop删除策略	必选	LakeFormation 数据湖功能

策略类型	策略类型	Action	是否必选	说明
		lakeformation:role:describe 查询角色详情 lakeformation:user:describe 查询用户详情 lakeformation:transaction:operate 操作事务		
消息通知	AI_DATALAKE_SMN_POLICY	smn:topic:publish 发布主题消息	可选	SMN通知消息
AOM 监控	AI_DATALAKE_AOM_POLICY	aom:metric:put 上报指标 aom:alarm:put 上报告警	可选	AOM监控告警功能
镜像包注册	AI_DATALAKE_SWR_POLICY	swr:repo:listRepoDomains 列举镜像仓库组织 swr:repo:listRepoTags 列举镜像标签 swr:repo:createRepoDomain 创建镜像仓库 swr:repo:getRepo 获取镜像仓库	可选	SWR镜像包功能
VPCEP	AI_DATALAKE_VPCEP_POLICY	vpcep:epservices:get 查询终端节点服务 vpcep:connections:update 更新连接 vpcep:permissions:update 更新权限 vpcep:permissions:list 列举权限 vpcep:epserviceDesc:get 查询终端节点服务描述 vpcep:endpoints:get 查询终端节点 vpcep:endpoints:delete 删除终端节点 vpcep:endpoints:create 创建终端节点 vpcep:connections:list 列举连接 vpcep:endpoints:list 列举终端节点	可选	VPCEP网络打通功能

策略类型	策略类型	Action	是否必选	说明
DNS 功能	AI_DATALAKE_DNS_POLICY	dns:zone:get 查询Zone dns:zone:create 创建Zone dns:zone:list 列举Zone dns:zone:delete 删除Zone dns:recordset:update 更新记录集 dns:recordset:create 创建记录集 dns:recordset:delete 删除记录集 dns:recordset:list 列举记录集	可选	DNS服务功能
OBS 存储	AI_DATALAKE_OBS_POLICY	obs:bucket:CreateBucket创建桶 obs:bucket:DeleteBucket删除桶 obs:bucket:PutLifecycleConfiguration设置生命周期配置 obs:bucket:ListBucketMultipartUploads列举分段上传 obs:bucket:HeadBucket判断桶是否存在 obs:bucket:ListAllMyBuckets列举所有桶 obs:bucket:ListBucket列举桶 obs:bucket:PutEncryptionConfiguration设置加密配置 obs:bucket:PutBucketObjectLockConfiguration设置对象锁定配置 obs:object:PutObject上传对象 obs:object:GetObject获取对象 obs:bucket:GetLifecycleConfiguration获取生命周期配置	可选	OBS对象存储功能

1.3 通过 IAM 授予使用 AI DataLake 的权限

1.3.1 通过 IAM 角色或策略授予使用 AI DataLake 的权限

如果您需要对您所拥有的AI DataLake进行角色与策略的权限管理，您可以使用[统一身份认证服务](#)（Identity and Access Management，简称IAM），通过IAM，您可以：

- 根据企业的业务组织，在您的华为账号中，给企业中不同职能部门的员工创建IAM用户，让员工拥有唯一安全凭证，并使用AI DataLake资源。
- 根据企业用户的职能，设置不同的访问权限，以达到用户之间的权限隔离。
- 将AI DataLake资源委托给更专业、高效的其他华为账号或者云服务，这些账号或者云服务可以根据权限进行代运维。

如果华为账号已经能满足您的要求，不需要创建独立的IAM用户，您可以跳过本章节，不影响您使用AI DataLake服务的其他功能。

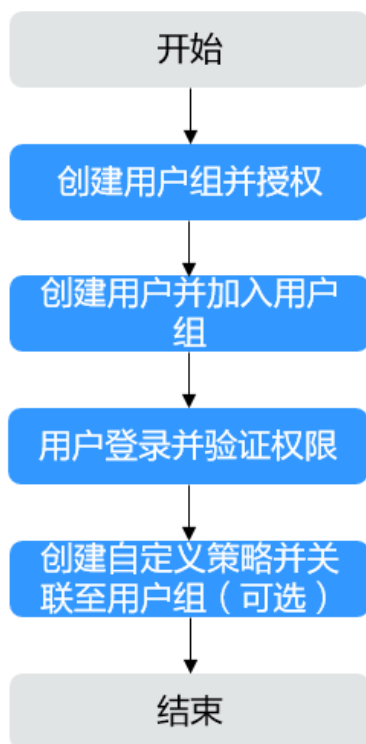
本章节为您介绍基于角色与策略的授权方法，操作流程如[图1-1](#)所示。

前提条件

给用户组授权之前，请您了解用户组可以添加的AI DataLake权限，并结合实际需求进行选择，AI DataLake支持的系统权限，请参见[角色与策略权限管理](#)。若您需要对除AI DataLake之外的其它服务授权，IAM支持服务的所有权限请参见[权限策略](#)。

示例流程

图 1-1 给用户授予 AI DataLake 权限流程



1. 创建用户组并授权

在IAM控制台创建用户组，并授予AI DataLake只读权限“AI DataLakeReadOnlyPolicy”。

2. 创建用户并加入用户组

在IAM控制台创建用户，并将其加入1中创建的用户组。

3. 用户登录并验证权限。

新创建的用户登录控制台，切换至授权区域，验证权限：

- 在“服务列表”中选择AI DataLake服务，进入AI DataLake主界面，单击左上角工作空间列表下拉框，选择“创建工作空间”，尝试创建AI DataLake工作空间，如果无法创建AI DataLake工作空间（假设当前权限仅包含AI DataLake ReadOnlyPolicy），表示“AI DataLake ReadOnlyPolicy”已生效。
- 在“服务列表”中选择除AI DataLake服务外（假设当前策略仅包含AI DataLake ReadOnlyPolicy）的任一服务，如果提示权限不足，表示“AI DataLake ReadOnlyPolicy”已生效。

AI DataLake 自定义策略样例

如果系统预置的AI DataLake权限，不满足您的授权要求，可以创建自定义策略。自定义策略中可以添加的授权项（Action）请参考[身份策略授权参考](#)。

目前华为云云服务平台支持以下两种方式创建自定义策略：

- 可视化视图创建自定义策略：无需了解策略语法，按可视化视图导航栏选择云服务、操作、资源、条件等策略内容，可自动生成策略。
- JSON视图创建自定义策略：可以在选择策略模板后，根据具体需求编辑策略内容；也可以直接在编辑框内编写JSON格式的策略内容。

具体创建步骤请参见：[创建自定义策略](#)。下面为您介绍常用的AI DataLake自定义策略样例。

- 示例1：授权用户创建、查看、修改工作空间权限

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "DataArtsFabric:workspace:list",
        "DataArtsFabric:workspace:alter",
        "DataArtsFabric:workspace:create"
      ]
    }
  ]
}
```

- 示例2：拒绝用户删除集群

拒绝策略需要同时配合其他策略使用，否则没有实际作用。用户被授予的策略中，一个授权项的作用如果同时存在Allow和Deny，则遵循**Deny优先原则**。

如果您给用户授予AI DataLake Minimal FullAccess的系统策略，但不希望用户拥有AI DataLake admin中定义的删除云服务器权限，您可以创建一条拒绝删除云服务的自定义策略，然后同时将AI DataLake Admin和拒绝策略授予用户，根据Deny优先原则，则用户可以对AI DataLake执行除了删除集群外的所有操作。拒绝策略示例如下：

```
{
  "Version": "1.1",
  "Statement": [
```

```

    {
      "Effect": "Deny",
      "Action": [
        "DataArtsFabric:workspace:drop"
      ]
    }
  ]
}

```

• 示例3：多个授权项策略

一个自定义策略中可以包含多个授权项，且除了可以包含本服务的授权项外，还可以包含其他服务的授权项，可以包含的其他服务必须跟本服务同属性，即都是项目级服务或都是全局级服务。多个授权语句策略描述如下：

```

{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "DataArtsFabric:workspace:list",
        "DataArtsFabric:workspace:alter",
        "DataArtsFabric:workspace:create"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "ecs:servers:get"
      ]
    }
  ]
}

```

AI DataLake 资源

资源是服务中存在的对象。在AI DataLake中，资源是指集群资源，您可以在创建自定义策略时，通过指定资源路径来选择特定资源。

表 1-3 AI DataLake 的指定资源与对应路径

指定资源	资源路径
workspace	DataArtsFabric:<region>:<account-id>:workspace:<workspace-id>

1.3.2 通过 IAM 身份策略授予使用 AI DataLake 的权限

如果您需要对您所拥有的AI DataLake进行身份策略的权限管理，您可以使用[统一身份认证服务](#)（Identity and Access Management，简称IAM），通过IAM，您可以：

- 根据企业的业务组织，在您的华为账号中，给企业中不同职能部门的员工创建用户或用户组，让员工拥有唯一安全凭证，并使用AI DataLake资源。
- 根据企业用户的职能，设置不同的访问权限，以达到用户之间的权限隔离。
- 将AI DataLake资源委托给更专业、高效的其他华为账号或者云服务，这些账号或者云服务可以根据权限进行代运维。

如果华为账号已经能满足您的要求，您可以跳过本章节，不影响您使用AI DataLake服务的其他功能。

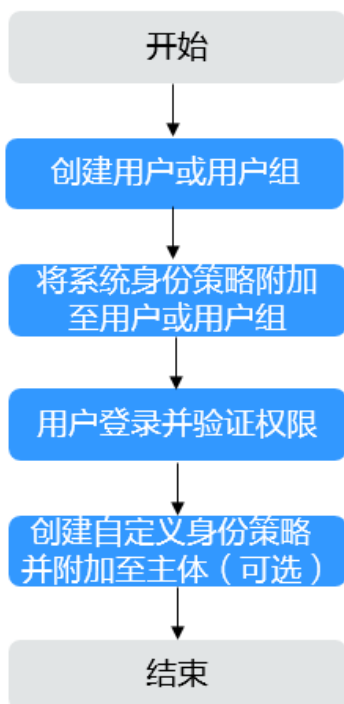
本章节为您介绍基于策略授权的授权方法，操作流程如图1-2所示。

前提条件

在授权操作前，请您了解可以添加的AI DataLake权限，并结合实际需求进行选择。AI DataLake支持的系统策略，请参见[身份策略权限管理](#)。若您需要对除AI DataLake之外的其他服务授权，IAM支持服务的所有权限请参见[权限策略](#)。

示例流程

图 1-2 给用户授予 AI DataLake 权限流程



1. **创建用户**或**创建用户组**
在IAM控制台创建用户或用户组。
2. **将系统策略附加至用户或用户组**
为用户所在的用户组授予AI DataLake只读权限的系统策略“AI DataLakeReadOnlyPolicy”，或将策略附加至用户所在的用户组。
3. **用户登录**并验证权限
使用已授权的用户登录控制台，验证权限：
 - 在“服务列表”中选择AI DataLake服务，进入AI DataLake主界面，单击左上角工作空间列表下拉框，选择“创建工作空间”，尝试创建AI DataLake工作空间，如果无法创建AI DataLake工作空间（假设当前权限仅包含AI DataLake ReadOnlyPolicy），表示“AI DataLake ReadOnlyPolicy”已生效。
 - 在“服务列表”中选择除AI DataLake服务外（假设当前策略仅包含AI DataLake ReadOnlyPolicy）的任一服务，若提示权限不足，表示“AI DataLake ReadOnlyPolicy”已生效。

AI DataLake 自定义策略样例

如果系统预置的AI DataLake系统策略，不满足您的授权要求，可以创建自定义身份策略。

目前华为云云服务平台支持以下两种方式创建自定义身份策略：

- 可视化视图创建自定义身份策略：无需了解策略语法，按可视化视图导航栏选择云服务、操作、资源、条件等策略内容，可自动生成策略。
- JSON视图创建自定义身份策略：可以在选择策略模板后，根据具体需求编辑策略内容；也可以直接在编辑框内编写JSON格式的策略内容。

具体创建步骤请参见：[创建自定义身份策略并附加至主体](#)。

下面为您介绍常用的AI DataLake自定义身份策略样例。

- 示例1：授权用户提交、查看、停止作业权限

```
{
  "Version": "5.0",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "DataArtsFabric:workspace:list",
        "DataArtsFabric:workspace:alter",
        "DataArtsFabric:workspace:create"
      ],
      "Resource": [
        "DataArtsFabric:*:*:workspace:*"
      ]
    }
  ]
}
```

- 示例2：多个授权项策略

一个自定义策略中可以包含多个授权项，且除了可以包含本服务的授权项外，还可以包含其他服务的授权项。多个授权语句策略描述如下：

```
{
  "Version": "5.0",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "DataArtsFabric:workspace:list",
        "DataArtsFabric:workspace:alter",
        "DataArtsFabric:workspace:create"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "vpc:vpcs:get"
      ]
    }
  ]
}
```

2 创建 AI DataLake 工作空间

2.1 工作空间概述

工作空间概述

在使用AI DataLake进行数据分析的场景中，管理员首先需要考虑的就是项目隔离、权限隔离与管理、资源分配问题。AI DataLake的工作空间正是为解决这些问题而设计，通过环境隔离、资源绑定的基本逻辑，为您提供一个独立的开发环境，并基于此环境创建计算资源、选择引擎、配置端点和开发作业，从而实现高效有序且安全隔离的数据开发。

工作空间的核心功能

工作空间作为独立的数据开发和管理单元，提供了以下核心功能：

- 每个工作空间拥有独立的计算资源，不同工作空间之间资源完全隔离，互不影响。所有数据分析任务均在工作空间内完成。
- 一个工作空间绑定一个LakeFormation实例，使用LakeFormation管理元数据和数据访问权限。
- 一个工作空间即可满足不同的数据处理需求，在工作空间下创建计算资源，并根据业务需求选择引擎创建端点，开发作业，每个空间可以新建不同的端点，用于连接计算引擎和计算资源，且支持专属预留资源和按需弹性资源。
- 支持通过基于IAM的细粒度权限控制不同的IAM用户对空间的访问权限，即开放所有空间的权限，不区分空间授权。
- 不同工作空间之间的数据、权限、作业相互隔离，保障项目数据安全。

工作空间约束与限制

表 2-1 工作空间约束限制

类型	说明
配额限制	每个账号在每个区域最多可创建5个工作空间。如需扩大配额参考 配额限制 。

类型	说明
命名规则	工作空间的名称长度为 4-32 个字符，只能包含字母、数字、下划线（_）和连字符（-），且不能以连字符（-）开头或结尾。
区域限制	工作空间创建后不能更换区域，也不支持迁移至其他区域。
生命周期	工作空间删除后无法恢复，请谨慎操作。
绑定Lakeformation实例规则	<ul style="list-style-type: none"> ● 一个工作空间只能绑定一个LakeFormation实例。 ● 不同工作空间可以绑定同一个LakeFormation实例。 ● 工作空间创建后，不支持切换绑定的Lakeformation实例。
权限要求	开通AI DataLake服务时需要授予AI_DATALAKE_SELF_POLICY权限，否则无法创建/管理工作空间。

工作空间与 LakeFormation 实例的关系

在AI DataLake服务架构中工作空间是业务与权限的隔离单元，LakeFormation实例是元数据与权限的数据服务载体。

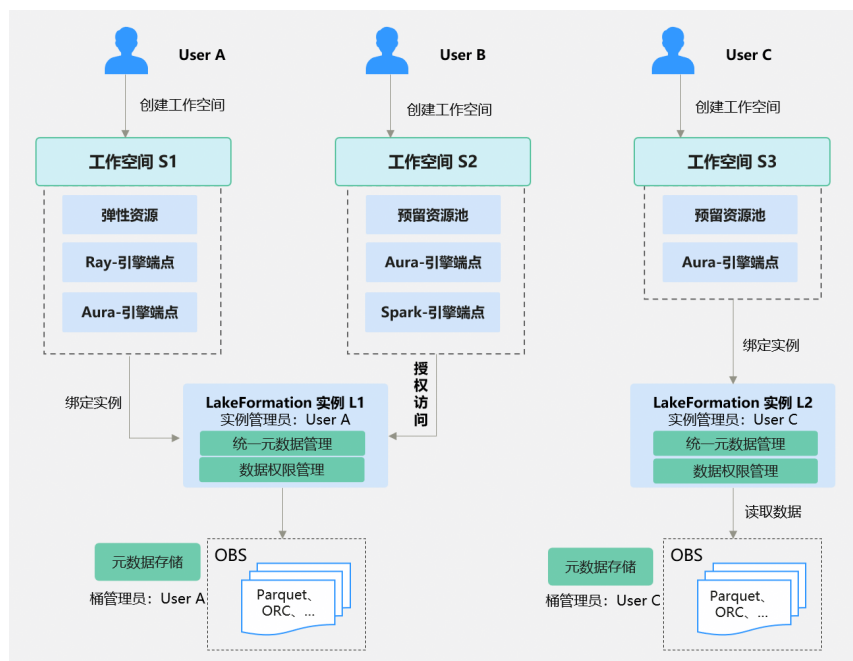
- LakeFormation实例提供元数据服务和权限管理功能。
- 工作空间是面向业务的逻辑空间，在绑定一个LakeFormation实例后才能使用元数据能力。
- 一个LakeFormation实例可被多个工作空间绑定，实现元数据共享 + 权限隔离。

在实际工作时：

- 工作空间本身不存储元数据，所有元数据操作（建库、建表、授权）都转发给绑定的LakeFormation实例执行。
- LakeFormation实例不感知业务隔离，仅提供统一的元数据读写与权限校验。而LakeFormation管理的数据都存在OBS中，因此在创建LakeFormation实例时需要绑定OBS桶。

如**图2-1**所示，用户User A创建工作空间S1，并绑定LakeFormation实例L1。用户User B创建工作空间S2后，通过LakeFormation的权限授权配置，实现S2访问L1中创建的元数据与对应OBS数据，从而达成多工作空间共享同一元数据实例、按需权限隔离与互访的效果。

图 2-1 工作空间与 LakeFormation 实例的关系



2.2 创建工作空间

操作场景

工作空间是AI DataLake为您提供的独立的数据分析工作环境，用于隔离不同业务、部门或项目的资源与权限。

本节操作介绍创建工作空间的操作步骤。

约束与限制

- 工作空间创建后，区域无法修改，请先选择区域，再创建工作空间。
- 每个账号在单个区域内可创建的工作空间数量有限，具体配额请参见[配额限制](#)。
- 删除工作空间前，请先确保空间下已无正在使用的端点或作业等数据。

前提条件

- 已注册华为云账号，并完成实名认证。
- 已开通AI DataLake服务。
- 已规划并创建LakeFormation实例、数据目录、数据库。
- 当前用户具备创建/查看/修改工作空间的权限，具体操作请参考[权限管理](#)。

操作步骤

- 步骤1** 登录AI DataLake管理控制台。
- 步骤2** 在左侧导航栏中，单击工作空间区域。
- 步骤3** 在下拉列表中选择“创建工作空间”，或者选择“工作空间管理 > 创建”。

步骤4 配置工作空间的基本信息：

表 2-2 工作空间的基本信息

类型	参数	是否必填	说明
基础信息	工作空间名称	是	工作空间的具体名称，同一个账号下的工作空间不可重名。 <ul style="list-style-type: none"> 名称只能包含字母、中文、数字、中划线、下划线。 名称字符长度为4~32位。 工作空间名称不区分大小写，系统会自动转换为小写。
	描述	否	对工作空间的简要描述。
	企业项目	否	如果所建工作空间属于企业项目，可选择对应的企业项目。 企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理。 关于如何设置企业项目请参考《 企业管理用户指南 》。 说明 只有开通了企业管理服务的用户才显示该参数。
多模数据管理	LakeFormation实例	是	选择当前工作空间关联的LakeFormation实例，每个工作空间需关联1个LakeFormation实例，空间创建后已关联的实例不支持修改。 <ul style="list-style-type: none"> 选择已创建的LakeFormation实例，用于统一管理元数据和数据访问控制。 如无可实例，请先创建LakeFormation实例，具体操作请参考创建LakeFormation实例。

步骤5 确认所有配置信息无误，单击“立即创建”按钮，完成工作空间创建。

工作空间创建成功后会返回工作空间管理界面，您可以在工作空间管理列表中查看新创建的工作空间。

----结束

后续操作

创建资源池：作业需要计算资源才能运行，因此在创建工作空间后，需要创建计算资源池来分配计算能力，具体操作请参考[创建计算资源池](#)。

2.3 管理工作空间

操作场景

工作空间创建成功后，可在工作空间列表查看详细信息，也可以修改或删除工作空间。

查看工作空间详细信息

- 步骤1** 登录AI DataLake管理控制台。
 - 步骤2** 在左侧导航栏中，单击工作空间区域，在下拉列表中选择“工作空间管理”。
 - 步骤3** 在工作空间列表页面可查看已创建的所有工作空间。
 - 步骤4** 在搜索框输入关键字搜索需要的工作空间，单击空间名称，进入空间详情页面。
 - 步骤5** 在工作空间详情页面可查看空间名称、创建人、描述、企业项目、空间ID、创建时间和所关联的LakeFormation实例。
- 结束

删除工作空间

- 步骤1** 登录AI DataLake管理控制台。
 - 步骤2** 在左侧导航栏中，单击工作空间区域，在下拉列表中选择“工作空间管理”。
 - 步骤3** 在工作空间列表页面的搜索框输入关键字搜索待删除的工作空间，单击“操作”列的“删除”。
 - 步骤4** 在“删除工作空间”弹框中输入“DELETE”，单击“确定”删除工作空间。
- 结束

3 创建 AI DataLake 计算资源池

3.1 计算资源池概述

AI DataLake 计算资源池类型

AI DataLake提供的计算资源分为预留资源池和弹性资源两种，满足不同业务场景的算力需求。

- **预留资源池：**用户提前购买实例资源，资源类型覆盖CPU、NPU、GPU三种类型的算力，您可根据业务负载特点选择合适的计算资源组合。

预留资源池为核心业务提供独占资源保障，性能稳定可预测，适用于长期稳定运行的批处理任务、AI计算场景。

- **弹性资源：**无需提前购买，按实际使用量计费，自动响应业务需求。

如果在配置端点时您选择的资源模式是“混合模式”，那么系统将为您开通弹性资源，在预留资源不足时自动调度资源进行补充。

混合资源调配既保障了业务平稳运行，又兼顾了资源的成本优化。

图 3-1 AI DataLake 计算资源类型与混合资源调配

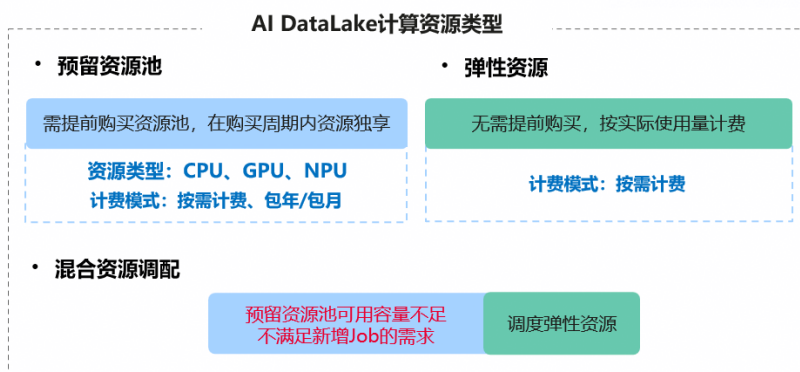


表 3-1 AI DataLake 计算资源类型与混合资源调配

资源类型	支持的计费模式	说明	优势	适用场景
预留资源	<ul style="list-style-type: none"> 按需计费 包年/包月 	您可通过按需/预付费用锁定一定周期内的资源使用量，期间资源可随时使用，到期后自动失效，需续费或重新购买，适合业务稳定、负载可预测的场景。	固定包年包月费用，比按需更加优惠。 预留资源，确保高峰期业务算力充足。	适用于生产项目：在此项目下，您的作业相对稳定（经过开发调试再上线），建议您使用包年/包月计费模式，可以更优惠。
弹性资源	按需计费	弹性资源无需购买，可以先使用资源，按任务实际使用量后付费，无需为业务峰值预留多余资源，有效控制成本，适合业务波动性较大或短期、偶发的场景。	按实际用量付费，更加灵活	适用于开发项目：开发调试时使用，作业随机性大，数据量小。按需计费能够帮您有效控制成本，将资源消耗控制在一定范围内。
混合资源调配	混合模式（预留资源池+弹性资源池）	使用的是预留资源池和弹性资源池。系统会优先计算预留资源池的可用容量是否满足新增Job的需求	与预留资源可以确保业务稳定运行，结合弹性资源池满足业务弹性需求	长期规划的业务需求，预估部分使用量的场景，在超过额度的使用量后按需计费。

AI DataLake 计算资源池使用模式

表 3-2 AI DataLake 计算资源池使用模式

资源使用模式	使用的资源类型	计费模式	资源特点	适用的引擎	适用场景
弹性模式	按需资源池	按需计费	无需提前购买，按实际使用量计费。	Aura	适用于资源使用不稳定，偶发性测试验证场景。 突发性业务负载、开发测试环境、业务探索期。

资源使用模式	使用的资源类型	计费模式	资源特点	适用的引擎	适用场景
预留模式	预留资源池	按需计费模式	需提前购买，按实际使用量计费，在购买周期内资源独享。 资源稳定，周期短。相比包年包月，无需承诺使用时长，灵活度高。	Aura	适用于资源使用稳定，周期短（几个小时）的业务。 例如：短期项目、临时性预留需求、负载波动大
	预留资源池	包年/包月模式	提前购买，固定周期付费，在购买周期内资源独享。 对比按需购买的方式，一次性购买较长周期的资源成本更优。	Aura	适用于资源使用稳定，波动小，周期较长（年/月） 例如：AI 模型训练、大规模数据批处理、核心业务分析
混合模式	预留资源池+弹性资源池	混合计费模式	混合模式下 ，系统会优先计算预留资源的可用容量是否满足新增Job的需求：Job所需资源大于预留资源池当前可用容量，系统自动触发调度弹性资源，作业将使用弹性资源池来执行，此时您需要按需支付按需弹性资源池的使用费用。	Aura	适用于资源使用稳定，波动大，特殊波动洪峰使用弹性资源池保障。 例如：日常有稳定负载，偶发性业务高峰场景。

预留资源池

- **资源类型：**
预留资源池支持三种异构计算资源，用户可根据业务负载特点灵活选择：

资源类型	说明	典型应用场景
CPU	通用计算处理器，适合数据处理、ETL、批处理等任务。	<ul style="list-style-type: none"> ● 通用计算增强型：适合于CPU密集型作业，例如实时流作业处理等。 ● 通用计算标准型：适合于大多数作业类型。 ● 通用计算内存增强型：内存密集型作业，例如离线批处理作业等。
GPU	GPU实例适用于图形渲染和大规模并行计算场景，适合深度学习训练和科学计算。GPU 拥有成百上千个计算核心，可以同时处理大量简单计算任务。	深度学习，例如神经网络训练、模型调优场景；图形处理场景，例如图像识别、目标检测；视频处理场景，例如视频分析、转码、视频渲染场景，等其他AI科学计算场景。
NPU	NPU适用于AI计算场景，NPU资源采用架构优化和指令集，专门加速AI推理任务。	AI计算场景、图像识别场景，推荐系统等AI计算设计场景。是当前人工智能训练与推理的重要基础设施。

- **计费模式：**

预留资源池支持按需和包年/包月两种计费模式：

- 按需计费：按实际使用量计费，在购买周期内资源独享。资源稳定，周期短。相比包年包月，无需承诺使用时长，灵活度高。
- 包年/包月：固定周期付费，在购买周期内资源独享。对比按需购买的方式，一次性购买较长周期的资源成本更优。

 **说明**

AI DataLake公测期间不支持使用包年/包月计费模式购买资源。

- **适用场景**

预留资源池适用于对稳定性要求高、负载可预测的业务场景，这些场景通常需要资源使用过程不被中断，具备充足的资源保证任务按时完成，因此可以通过预留资源获取更优的单位算力和成本。

弹性资源

- **资源特点：**

弹性资源无需购买，按需使用，按实际使用的资源量和运行时长计费，采用后付费模式。

弹性资源适用小规模或临时的数据处理需求。对于重要的或需要保证资源的作业，建议购买预留资源池来执行作业。

- **适用场景：**

弹性资源池适用于临时算力需求场景，例如在开发测试阶段不确定的资源需求，新业务上线初期，负载规模尚不明确。

还可以用于“混合模式”与预留资源搭配使用，应对业务负载有明显波峰波谷的业务场景。

混合资源调配

混合资源调配是将预留资源池与弹性资源池结合，实现成本与性能的平衡。

混合资源调配基本原理如图3-2所示：优先使用预留资源，预留资源不足时自动切换至弹性资源补充，

系统按以下优先级调度计算资源：

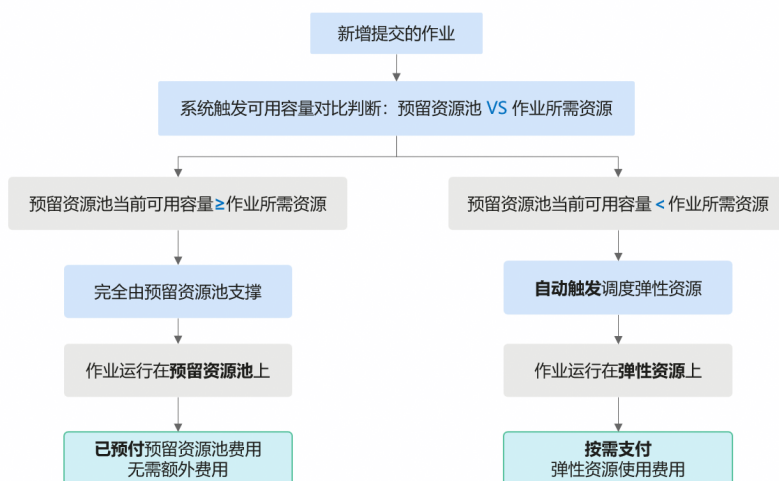
- 第一优先级：预留资源池中的空闲资源
- 第二优先级：弹性资源池中按需创建的资源

混合资源调配模式下，系统会优先计算预留资源的可用容量是否满足新增Job的需求：

- Job所需资源小于等于预留资源池当前可用容量，完全由预留资源池支撑，不会触发调度弹性资源池，也无需支付额外费用。
- Job所需资源大于预留资源池当前可用容量，系统自动触发调度弹性资源，作业将使用弹性资源池来执行，此时您需要按需支付按需弹性资源池的使用费用。

这种混合资源调配模式兼具预留资源的稳定性与按需资源的灵活性，有效应对业务变化。

图 3-2 混合资源调配原理图



3.2 购买预留资源池

操作场景

AI DataLake预留资源池即购买独立的计算资源，不管使用“包年/包月”或是“按需计费”购买的预留资源，在购买周期内资源都是独享，空闲（无作业运行）时不会释放。

预留资源池适用于日常有稳定的计算需求，作业运行规律、可预测的稳定负载业务场景。持续可用的计算资源保障了资源池不被其他用户抢占，性能更加稳定。

本节操作介绍预留资源池的适用场景、约束限制和购买方式。

预留资源池计费模式

AI DataLake预留资源池即购买独立的云服务计算资源，不管使用“包年/包月”或是“按需计费”购买的预留资源，在购买周期内资源都是独享，空闲（无作业运行）时不会释放。其中“包年/包月”的方式价格比“按需计费”模式更优惠。

- **包年/包月：**预付费模式，按订单的购买周期计费。在购买周期内资源独享，空闲（无作业运行）时不会释放，使用体验更佳，价格比按需计费模式更优惠。

适用于可预估资源使用周期的场景，例如已完成开发进入生产阶段的项目，推荐使用包年包月计费模式预留业务所需资源。

包年包月的预留资源池在使用过程中规格内按包周期计费，超出规格部分按需计费。

- **按需计费：**后付费模式，按实际使用量计费，在购买周期内资源独享，空闲时资源不被释放。

自创建起根据选购的云服务的规格类型按自然小时收费，秒级计费，按小时结算。计算费用=单价*小时数。

按需计费模式适用于测试项目，资源消耗不高，按需计费成本更低。

- **混合资源调配：**

混合资源调配是指预留资源池和弹性资源池组合使用。

当预留资源池容量不足以支撑业务使用时，系统自动调用弹性资源补充，确保业务平稳运行。自动调用的弹性资源池按需计费，将会产生预留资源池外的按需费用。

混合资源调配优先消耗预留资源，预留资源耗尽后按弹性资源计费，兼具预留资源的稳定性与按需资源的灵活性，有效应对业务变化。

如需降低该场景下的整体成本，则可以通过**变更预留资源池规格**的操作，扩大预留资源池的规格，且“扩大后的预留资源规格”**大于等于**“原预留资源池规格与调配的弹性资源池的和”。扩大预留资源池规格后，相比变更前（部分弹性资源池按需计费）更优惠。

约束与限制

- 预留资源池名称长度为4-32个字符，只能包含字母、数字、下划线（_）和连字符（-），且不能以连字符（-）开头或结尾。
- 资源池必须绑定网络后使用，可参考[创建网络](#)。

购买预留资源池

步骤1 登录AI DataLake管理控制台。

步骤2 在左侧导航栏单击“资源管理”。

步骤3 在资源管理页面单击“计算资源池”。

步骤4 单击界面右上角的“购买”。

步骤5 在“购买计算资源池”界面，填写具体参数，参数填写参考如下。

表 3-3 参数说明

参数名称	描述
区域	<p>区域指预留资源池的物理数据中心所在的位置，请就近选择靠近您业务的区域，可减少网络时延，提高访问速度。</p> <p>预留资源池创建后不能更换区域，也不支持迁移至其他区域，请谨慎选择。</p>
计费模式	<p>选择预留资源池的计费模式：</p> <p>按需计费：后付费模式，按实际使用量计费，在购买周期内资源独享，空闲时资源不被释放。</p> <p>自创建起根据选购的云服务的规格类型按自然小时收费，秒级计费，按小时结算。</p> <p>计算费用=单价*小时数</p>
资源池名称	<p>预留资源池的具体名称。</p> <p>预留资源池名称长度为1-63个字符，只能包含字母、数字、下划线（_）和连字符（-），且不能以连字符（-）开头或结尾。</p> <p>说明 预留资源池名称不区分大小写，系统会自动转换为小写。</p>
资源类型	<p>在创建预留资源池时，您需要选择合适的计算资源类型。AI Datalake支持三种计算资源类型：CPU（通用计算型）、GPU（异构计算型）和 NPU（AI计算型）。不同的资源类型适用于不同的业务场景，您可以根据实际业务需求进行选择。</p>
CPU 资源	<p>设置当前资源池中的CPU资源数量。</p> <p>CPU是通用型计算资源，适用于各种类型的计算任务。</p> <ul style="list-style-type: none"> ● CPU资源的特点：通用性强，适合各种类型的计算任务。相比于GPU和NPU的成本更低。擅长处理顺序执行的任务。 ● CPU资源的适用场景： <ul style="list-style-type: none"> - 通用计算增强型：适合于CPU密集型作业，例如实时流作业处理等。 - 通用计算标准型：适合于大多数作业类型。 - 通用计算内存增强型：内存密集型作业，例如离线批处理作业等。 ● CPU资源的具体规格请参考产品规格

参数名称	描述
GPU 资源	<p>设置当前资源池的GPU规格及资源数量，单击“添加规格”，可为当前资源池选择多个GPU规格。</p> <p>GPU资源适用于图形渲染和大规模并行计算场景，适合深度学习训练和科学计算。GPU 拥有成百上千个计算核心，可以同时处理大量简单计算任务。</p> <ul style="list-style-type: none"> GPU资源的特点：支持并行计算，适用于批处理作业场景，数千个核心同时计算，处理效率高。天然适合深度学习、神经网络、AI计算场景。 GPU资源的适用场景：深度学习，例如神经网络训练、模型调优场景；图形处理场景，例如图像识别、目标检测；视频处理场景，例如视频分析、转码、视频渲染场景，等其他AI科学计算场景。 GPU资源的具体规格请参考产品规格
NPU 资源	<p>设置当前资源池的NPU规格及资源数量，单击“添加规格”，可为当前资源池选择多个NPU规格。</p> <p>NPU资源适用于AI计算场景，NPU资源采用架构优化和指令集，专门加速AI推理任务。</p> <ul style="list-style-type: none"> NPU资源的特点：AI计算场景专用，具备高性能、低延迟、推理成本更优的特点。 NPU资源的适用场景：AI计算场景、图像识别场景，推荐系统等AI计算设计场景。是当前人工智能训练与推理的重要基础设施。 <p>在实际选型中，综合考虑模型规模、数据量、并发量、成本预算来选择最匹配的NPU资源方案，从而实现性能、成本和资源利用率的平衡。</p> <ul style="list-style-type: none"> NPU资源的具体规格请参考产品规格
AI DataLake 网络	<p>规划预留资源池所属的网段，弹性资源池网段与数据源网段不能重合。</p> <p>一个AI DataLake工作空间内仅支持一个网络，如果没有可用网络，单击“创建网络”新增，更多关于AI DataLake网络的介绍，可参考配置计算资源池网络。</p>

步骤6 参数填写完成后，单击“立即购买”。

步骤7 等待弹性资源池状态变成“可用”表示当前弹性资源池创建成功。

----结束

相关文档

- 资源池购买成功后，可将其关联至对应的引擎端点，完成资源分配，引擎即可调用算力执行任务，请参考[创建AI DataLake Aura引擎端点](#)。
- 计算资源池创建后只是具备计算能力的“孤岛”，只有通过网络连接才能做到与外部数据和服务的交互，请参考[配置计算资源池网络](#)。

3.3 配置计算资源池网络

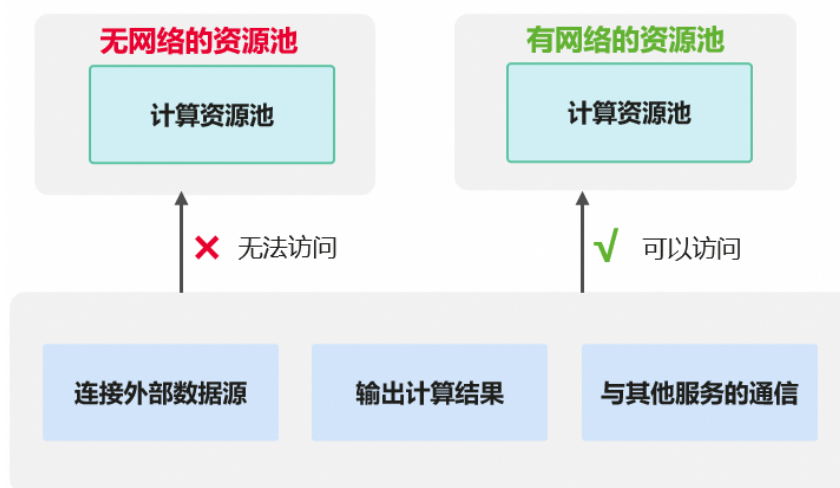
3.3.1 为什么要配置计算资源池网络

计算资源池创建后只是具备计算能力的“孤岛”，只有通过网络连接才能做到与外部数据和服务的交互。因此给计算资源池配置网络是使用计算资源分析数据的基础要求。

如图3-3所示，网络是计算资源池与外部环境进行数据交换和通信的唯一通道。没有网络配置，计算资源池将无法发挥任何作用。

本节内容以几个典型的计算资源池与外部交互的场景为例，介绍计算资源池配置网络的必要性。

图 3-3 AI DataLake 计算资源池类型

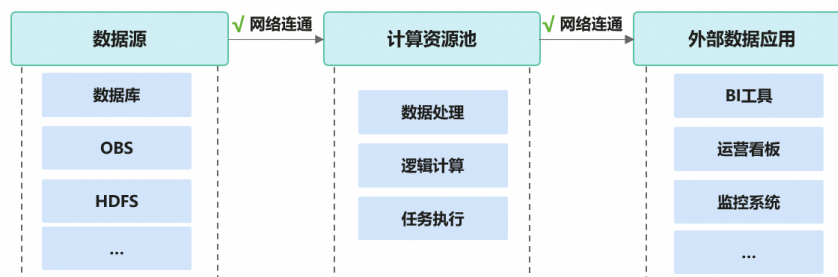


保障 AI DataLake 数据的流入与流出

计算资源池的核心职责是处理数据，而数据的输入和输出都依赖网络。

- 数据采集：外部的数据源（如业务数据库、日志服务器）需要将数据实时或批量地推送到计算资源池中进行处理。网络带宽决定了数据的写入速度。
- 分析结果输出：分析处理后的结果（如报表、模型预测结果）需要发送给外部应用（如 BI 工具、Web 服务）进行展示或进一步利用。

图 3-4 数据的流入与流出

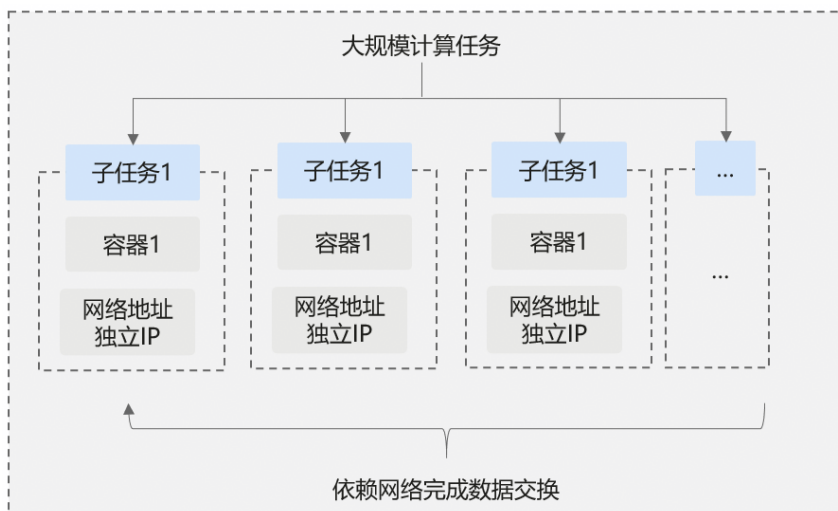


满足 AI DataLake 计算节间的通信需求

AI DataLake计算资源底层使用了容器资源，容器之间需要通过网络进行数据交换和任务协调，以完成大规模计算任务。

这个过程中会将计算任务需要分解为多个子任务，子任务之间需要通过网络进行数据交换。没有网络连接，容器将成为孤立个体，无法被访问。

图 3-5 AI DataLake 计算节间的通信需求



3.3.2 创建网络

计算资源池创建后需要通过网络连接才能做到与外部数据和服务的交互，同时为避免资源池内部网络与其他数据源网络规划冲突，在工作空间中需要创建一个网络，基于虚拟私有云（VPC）进行封装，对用户提供的CIDR网段的选择项。

基于VPC对等连接，可实现跨VPC的资源互通，帮助用户高效共享资源、降低数据传输延迟，并提升业务系统的稳定性与连续性。

创建网络有两种方式，用户在在工作空间的资源管理页面创建，也可在创建计算资源池时统一创建。

图 3-6 AI DataLake 网络包含的资源示例

网络示例

- **VPC:**
 - VPC ID: vpc-xxxx
 - 网段: 192.168.0.0/16
- **子网:**
 - 子网ID: subnet-xxxxx
 - 子网网段: 192.168.1.0/24
- **安全组:**
 - 安全组ID: sg-xxxxx
 - 规则: SSH (22) 、内网互通 (全部)

约束与限制

表 3-4 网络约束限制

类型	说明
配额限制	一个AI DataLake工作空间内只支持创建一个网络。
网段的规划建议	<p>购买计算资源时配置的网段，主要是为了给每个计算节点分配独立的私网IP地址，即计算资源池的每个Pod同样需要独立IP。这些IP用于节点间的数据交换、Shuffle通信、访问数据源以及服务发现等。</p> <ul style="list-style-type: none">• 网段越大，可容纳节点/容器数量越多，扩容更灵活。• 子网掩码越大可用IP数量越少。• 网段不支持更改或删除，请合理规划IP范围，避免IP资源不足导致无法扩容。

类型	说明
网段的配置要求	<ul style="list-style-type: none"> ● 必须配置私网IP地址段 ● 可用网段范围： <ul style="list-style-type: none"> - 10.0.0.0/8 ~ 10.255.0.0/22 - 172.16.0.0/12 ~ 172.31.0.0/22 - 192.168.0.0/16 ~ 192.168.0.0/22 ● 不能与数据源网段重合。

创建网络

- 步骤1** 登录AI DataLake管理控制台。
- 步骤2** 在左侧导航栏单击“资源管理”。
- 步骤3** 在资源管理页面单击“网络”。
- 步骤4** 在网络配置页面右上角单击“创建网络”，输入网络对应的网段地址后单击“确定”。

图 3-7 创建网络



- 步骤5** 网络创建成功后，在网络列表中可查看对应的网络ID信息等内容。

----结束

3.4 调整计算资源池容量

随着业务负载的不断增长，现有的计算资源池容量可能无法满足日益增长的算力需求。为了保障业务的稳定运行，管理员需要通过弹性扩容的方式动态调整计算资源池的规模，以适应业务峰谷变化。

约束与限制

- 计算资源池包含多种资源类型（CPU、GPU、NPU）时，同一扩缩容任务只支持调整一种资源类型。如需同时调整多种资源，需分多次执行扩缩容任务。
- 仅在资源池状态为“运行中”时可操作扩容与缩容。
- 扩容与缩容的生效时间与规格、实例数量相关，不可预估。

- 扩容前请先确认作业是否已经完成，系统优先确保作业运行完成。若30分钟内作业未结束，系统将执行强制扩容，请注意长作业的失败风险。

扩缩容计算资源池

步骤1 登录AI DataLake管理控制台。

步骤2 在左侧导航栏单击“资源管理”。

步骤3 在计算资源池页面中待操作的计算资源池的“操作”列下，单击“更多 > 扩缩容”。

步骤4 选择待扩容的资源类型、资源规格以及资源目标数量，单击“下一步”。

图 3-8 资源池扩缩容



1 扩缩容 ————— 2 变更确认

资源池扩缩容

资源池名称

计费模式
按需计费

资源类型

CPU GPU NPU

规格 ?
pi5.8xlarge.4

当前数量
1

目标数量 ?
- | 1 | +

需预留 1 个以保障端点运行，最大可扩容至 10240 个。[查看端点](#)

步骤5 确认资源规格及变更数量无误后，单击“确定”。

任务提交后，资源池将进入“扩容中”或者“缩容中”状态，等待资源池扩容完成。

扩容前请先确认作业是否已经完成，系统优先确保作业运行完成。若30分钟内作业未结束，系统将执行强制扩容，请注意长作业的失败风险。

待资源扩容完成后，资源池状态将自动变更为“可用”，表示扩容成功完成，新的计算资源已就绪可供使用。

----结束

3.5 管理计算资源池

管理员需要定期监控和了解当前工作空间内计算资源池的运行状态、资源使用情况以及关联的端点信息，以便进行资源规划和运维管理。

当某些计算资源池不再使用时，为了释放系统资源并避免资源浪费，管理员需要将其从系统中删除。删除操作是不可逆的，请确保已确认不再需要该资源池。

查看计算资源池信息

- 步骤1** 登录AI DataLake管理控制台。
- 步骤2** 在左侧导航栏单击“资源管理”。
- 步骤3** 在计算资源池页面可查看当前工作空间内已创建的资源池列表信息。
- 步骤4** 单击资源池名称，可进入资源池概览信息页面，查看基本信息、资源信息。


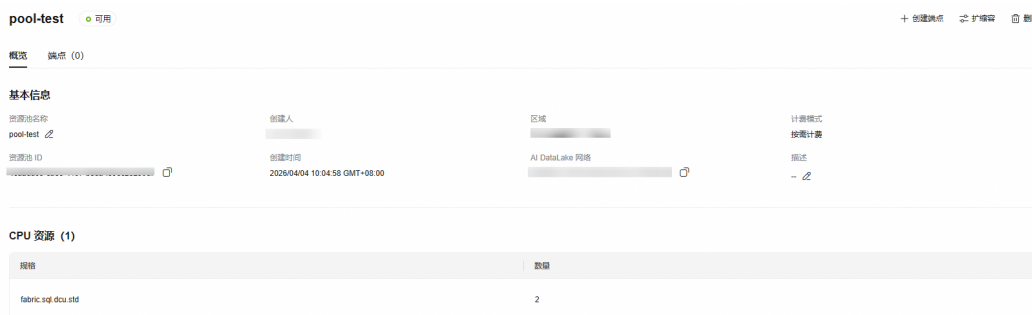
单击  ，可修改资源池名称及描述信息。

图 3-9 查看资源池概览信息



计算资源池在生命周期内会处于不同的状态，了解各种状态的含义有助于管理员更好地进行资源管理和问题排查。

表 3-5 资源池状态说明

状态	说明
创建中	表示系统正在根据用户配置创建新的计算资源池。在该阶段，系统正在分配计算资源、初始化配置、建立网络连接等操作。
创建失败	表示计算资源池在创建过程中未能成功完成，资源池未能正常创建成功。可根据报错信息进行问题定位，例如检查当前账号下可用资源（CPU/GPU/NPU）是否充足、当前用户权限是否正常。

状态	说明
删除中	表示系统正在执行资源池的删除操作，正在释放计算资源、清理关联配置等。
冻结	表示计算资源池被系统锁定，无法正常使用。冻结状态下的资源池保留配置信息，但计算资源已暂停分配，无法承载业务负载。
扩容中	系统正在对计算资源池进行扩容操作，正在增加 CPU、内存、GPU、NPU 等计算资源的配置。
缩容中	系统正在对计算资源池进行缩容操作，正在减少 CPU、内存、GPU、NPU 等计算资源的配置。
删除中	系统正在删除计算资源池。

步骤5 在“端点”页签，可查看当前关联至本资源池的端点信息。

----结束

删除计算资源池

步骤1 登录AI DataLake管理控制台。

步骤2 在左侧导航栏单击“资源管理”。

步骤3 在计算资源池页面中待操作的计算资源池的“操作”列下，单击“更多 > 删除”。

步骤4 在弹出的窗口中按照界面提示输入“DELETE”，单击“确定”。

当前资源池下存在已关联的端点时，不支持直接删除资源池，请先删除相关端点后再重试。

步骤5 等待资源池删除成功。

----结束

4 创建 AI DataLake Aura 引擎端点

4.1 Aura 引擎端点概述

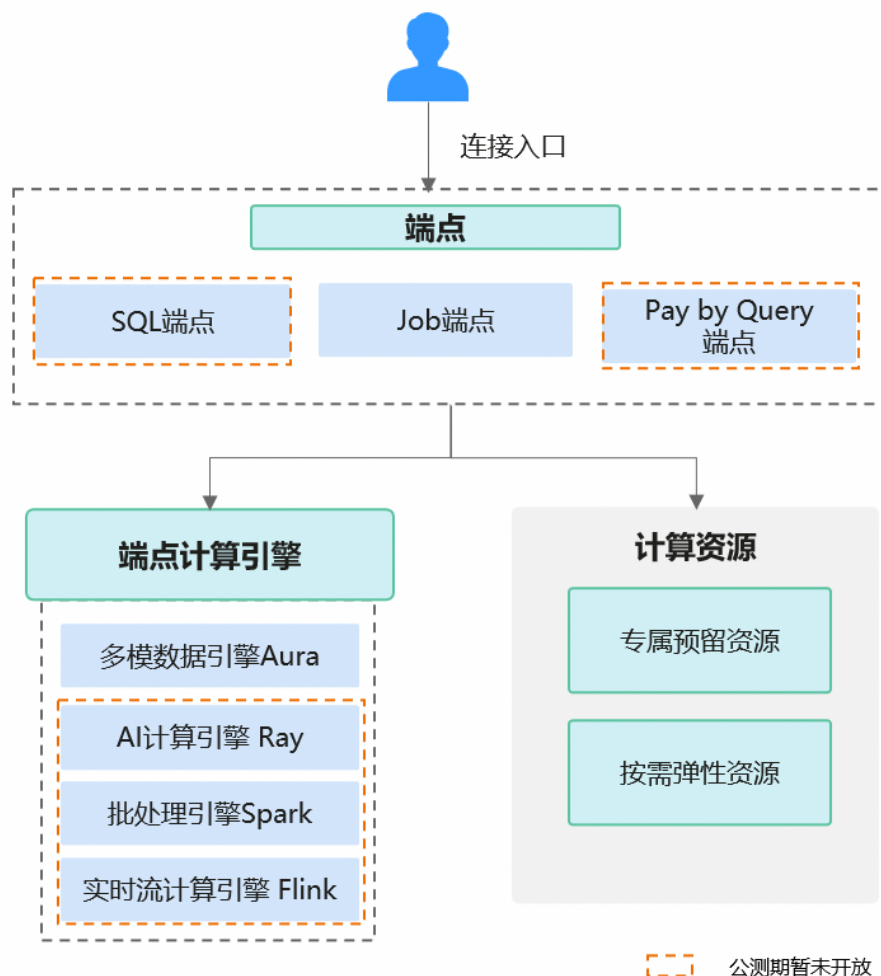
端点类型

在工作空间中提交作业时，是通过端点来连接引擎与计算资源。创建端点时，AI DataLake自动将引擎和计算资源池绑定，在后续提交作业时无需再手动配置作业与引擎和资源的映射关系。

端点类型分为端点引擎类型和端点类型，两个维护互相结合，满足不同业务场景的需求。

- 端点引擎类型是指端点绑定的计算引擎，决定了数据处理的计算方式和能力。
- 端点类型是指端点的使用模式，决定了连接方式、计费模式和适用场景。

图 4-1 端点类型介绍



Aura 引擎支持配置的端点类型

Aura引擎支持配置三种类型的端点，分别适用于不同的使用场景和业务需求。您可以根据工作负载特点选择合适的端点类型。

- **Job端点：**

Job端点用于提交和管理批量作业，适用于定时任务和大规模数据处理场景。

支持批量提交作业、作业调度与编排（在DataArts Studio空间完成）、任务状态监控，适合大规模数据处理。常应用于ETL数据抽取转换加载、定时批处理任务、大规模数据清洗等定时任务和大规模数据处理场景。

Aura 引擎支持的资源使用模式

Aura引擎支持两种资源使用模式：预留模式和混合模式。用户可以根据业务负载特点选择合适的模式，优化成本、平衡资源性能。

表 4-1 Aura 引擎支持的资源使用模式

类型	说明	详细说明
预留模式	独享性能，成本最优。预留资源，单价最低，确保业务基线稳定。适用于负载稳定业务场景。	资源使用模式：预留模式
混合模式	基线保障，自动扩容。优先消耗预留资源，高峰期自动触发弹性补位。适用于有规律波动的业务。	资源使用模式：混合模式

资源使用模式：预留模式

- 模式介绍

预留模式通过预留专属计算资源，确保业务所需的计算资源的稳定性，同时获得最优的单价成本。

在预留模式下，用户预先购买一定量的计算资源，这些资源始终为用户保留独享。无论是否使用这些资源，都需要按预留的资源数量支付费用。这种模式适合负载稳定、持续运行的业务，能够提供稳定的性能保障和最优的成本效率。

图 4-2 预留资源模式示意图

预留模式

独享性能，成本最优。预留资源，单价最低，确保业务基线稳定。适用于负载稳定业务场景。



- 核心特性

- 资源独享，性能稳定：预留资源完全独享，不存在资源争抢问题，性能更加稳定。
- 成本最优：预留资源单价在所有模式成本最优。

- 适用场景

- 负载长期稳定运行的业务：资源使用量稳定，例如持续运行的服务，日均处理量波动较小的业务。
- 性能保障要求高的业务：对查询延迟有严格要求，需要稳定的资源保障响应性能的应用。独享资源可以确保性能不会受到其他用户的影响。
- 成本优化场景的业务需求：长期运行的生产环境的业务，在资源利用率不变的情况下可以通过使用预留资源池降低资源成本。

资源使用模式：混合模式

- 模式介绍

混合模式结合了预留资源和弹性资源，优先消耗预留资源，高峰期预留资源不足时调度弹性资源组合使用。预留资源保证了业务基线的稳定性，弹性资源的自动调度又具备应对突发负载的弹性能力。

图 4-3 混合资源模式示意图

混合模式

基线保障，自动扩容。优先消耗预留资源，高峰期自动触发弹性补位。适用于有规律波动的业务。



- 核心特性:

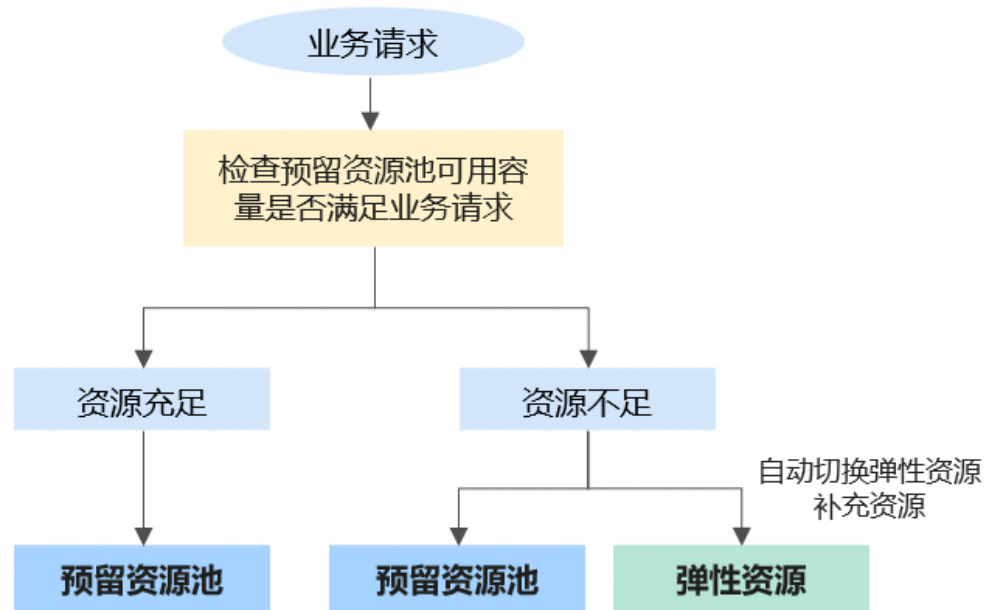
- 预先购买的预留资源作为业务运行的基线，无论何时都可以使用，不受弹性资源状态影响。
- 而当业务负载超过预留资源的容量时，系统会自动调度弹性资源，无需人工干预。调度的弹性资源按实际使用量计费，避免资源浪费。

- 资源消耗顺序

混合资源调配顺序如图4-4所示，系统按以下优先级调度计算资源：

- 第一优先级：预留资源中的空闲资源
- 第二优先级：按需创建的弹性资源

图 4-4 混合资源调配原理图



混合模式下，系统会优先计算预留资源的可用容量是否满足新增作业的需求：

- 作业所需资源小于等于预留资源池当前可用容量，完全由预留资源池支撑，不会触发调度弹性资源。
- 作业所需资源大于预留资源池当前可用容量，系统自动触发调度弹性资源，作业将使用弹性资源来执行

• **适用场景**

混合模式适用于以下业务场景：

- **有规律波动的业务场景：**
业务量在一天中呈现明显的周期性变化，业务负载呈现规律性的高低起伏。例如白天需要更多资源，夜间资源需求较少。
- **兼顾性能与成本的业务场景：**
通过日常使用预留资源确保业务稳定运行，提供可预期的基础性能。同时保障应对高峰或突发场景的能力，通过弹性资源自动扩展资源，无需人工干预。
- **一些不可预测的突发负载场景：**
可能存在突发的业务场景需要快速扩展资源，应对数据分析需求，这种需求通常是临时的短期的业务需求。

4.2 创建 Aura 引擎端点

操作场景

创建Aura端点是用户与AI DataLake引擎服务建立连接的核心操作，作为调用引擎服务的关键步骤，为用户提供了一个安全、稳定的访问入口。

- **一键创建：** 仅需简单配置端点类型和资源模式，系统自动完成底层资源分配。

- **专属入口**：每个端点对应唯一的访问地址，用户可通过该入口直接调用引擎服务。
- **多模态支持**：统一入口支持结构化数据、半结构化数据、非结构化数据等多种数据类型的处理与分析。

约束与限制

表 4-2 端点约束限制

类型	说明
配额限制	<ul style="list-style-type: none">● 每个账号最多创建1000个端点。● 一个工作空间最多创建50个端点。
命名规则	工作空间的名称长度为1-64个字符，只能包含小写字母、数字、中划线（-），以字母开头，字母或数字结尾。

前提条件

- 已创建工作空间，详细操作请参见[创建工作空间](#)。
- 已创建计算资源，详细操作请参见[购买预留资源池](#)。

创建 Aura 引擎端点

- 步骤1** 登录AI DataLake管理控制台。
- 步骤2** 在页面左上角切换至对应的工作空间。
- 步骤3** 在左侧导航栏选择“引擎管理 > 多模数据引擎Aura”进入Aura引擎端点列表页面。
- 步骤4** 单击页面右上角的“创建端点”，配置以下参数并单击“立即创建”。

表 4-3 创建 Aura 引擎端点

参数	参数说明
端点类型	选择端点类型。 <ul style="list-style-type: none">● Job 端点：用于执行定时调度任务，保障大规模数据稳定处理。
端点名称	输入端点名称。 <ul style="list-style-type: none">● 名称只能包含小写字母、数字、中划线，且只能以字母开头，以字母或数字结尾。● 输入长度不能超过63个字符。
添加描述	可单击“添加描述”添加端点描述信息。

参数	参数说明
资源使用模式	<p>选择资源使用模式。</p> <ul style="list-style-type: none">• 预留模式：预留模式通过预留专属计算资源，确保业务所需的计算资源的稳定性，同时获得最优的单价成本。详细介绍请参见资源使用模式：预留模式。• 混合模式：混合模式结合了预留资源和弹性资源，优先消耗预留资源，高峰期预留资源不足时调度弹性资源组合使用。预留资源保证了业务基线的稳定性，弹性资源的自动调度又具备应对突发负载的弹性能力。详细介绍请参见资源使用模式：混合模式。
选择资源池	<p>在下拉框中选择已创建的资源池。如果下拉框中没有可选的资源池，可以单击“购买资源池”进行购买。</p>
CPU资源	<p>配置对应CPU的保障配额及最大配额。CPU为通用计算处理器，适合数据处理、ETL、批处理等任务。</p> <ul style="list-style-type: none">• 静态余量：资源池中此规格资源尚未被“保障分配”的剩余容量，即资源池可用容量。• 保障配额 (Min)：资源池分配给当前端点的最低可用资源，确保核心业务有资源可用。<ul style="list-style-type: none">- 预留模式： 该数值需与最大配额值相同，且需大于0。- 混合模式 该数值需小于等于最大配额值，且需大于0。• 最大配额 (Max)：资源池分配给当前端点的资源上限，防止资源池资源被完全占用。<ul style="list-style-type: none">- 预留模式：参数取值范围为0~<i>静态余量值</i>，且需与保障配额值相同。- 混合模式：最大配额即为分配的弹性资源的数量，参数取值可根据实际业务进行配置，即预估业务需使用的最大CPU配额。参数值需大于0。

参数	参数说明
GPU资源	<p>配置对应GPU的保障配额及最大配额。GPU为图形处理器，具有强大的并行计算能力。</p> <ul style="list-style-type: none"> ● 静态余量：资源池中此规格资源尚未被“保障分配”的计算卡数量。 ● 保障配额 (Min)：资源池分配给当前端点的最低可用资源，确保核心业务有资源可用。 <ul style="list-style-type: none"> - 预留模式： 该数值需与最大配额值相同，且需大于0。 - 混合模式 该数值需小于等于最大配额值，且需大于0。 ● 最大配额 (Max)：资源池分配给当前端点的资源上限，防止资源池资源被完全占用。 <ul style="list-style-type: none"> - 预留模式：参数取值范围为0~静态余量值，且需与保障配额值相同。 - 混合模式：最大配额即为分配的弹性资源的数量，参数取值可根据实际业务进行配置，即预估业务需使用的最大CPU配额。参数值需大于0。
NPU资源	<p>配置对应NPU的保障配额及最大配额。NPU为神经网络处理器，擅长AI推理任务。</p> <ul style="list-style-type: none"> ● 静态余量：资源池中此规格资源尚未被“保障分配”的计算卡数量。 ● 保障配额 (Min)：资源池分配给当前端点的最低可用资源，确保核心业务有资源可用。 <ul style="list-style-type: none"> - 预留模式： 该数值需与最大配额值相同，且需大于0。 - 混合模式 该数值需小于等于最大配额值，且需大于0。 ● 最大配额 (Max)：资源池分配给当前端点的资源上限，防止资源池资源被完全占用。 <ul style="list-style-type: none"> - 预留模式：参数取值范围为0~静态余量值，且需与保障配额值相同。 - 混合模式：最大配额即为分配的弹性资源的数量，参数取值可根据实际业务进行配置，即预估业务需使用的最大CPU配额。参数值需大于0。
日志对接LTS	<p>是否对接LTS，如果不对接LTS，则无法查看对应作业的日志，了解更多对接详情请参见使用LTS管理AI DataLake作业日志。</p> <p>对接LTS后日志会投递到云日志服务（Log Tank Service，简称LTS）进行管理。可以使用LTS对云服务日志进行关键词搜索、运营数据统计分析、运行状况监控告警等多种操作。</p> <p>勾选该参数后，还需配置“日志组”和“日志流”参数。</p>

参数	参数说明
日志组	<p>在下拉框中选择日志组，如果下拉框中没有可选的日志组，可以单击“创建日志组”，在右侧弹框中参考表4-4配置相关参数后，单击“确定”创建日志组。</p> <p>日志组（LogGroup）是云日志服务进行日志管理的基本单位，用于对日志流进行分类，一个日志组下面可以创建多个日志流。日志组本身不存储任何日志数据，仅方便管理日志流，每个账号下可以创建100个日志组。</p>
日志流	<p>在下拉框中选择日志流，如果下拉框中没有可选的日志流，可以单击“创建日志流”，在右侧弹框中参考表4-5配置相关参数后，单击“确定”创建日志流。</p> <p>云日志服务是以日志流（LogStream）作为日志管理维度。日志采集后，以日志流为单位，将不同类型的日志分类存储在不同的日志流上，方便对日志进一步分类管理。</p>

表 4-4 创建日志组

参数	参数说明
日志组名称	输入日志组名称，只能包含英文字母、数字、中文、中划线、下划线及小数点，且不能以小数点、下划线开头或以小数点结尾，不能与其他日志组名称或原始名称相同。
日志流名称	输入日志流名称，只能包含英文字母、数字、中文、中划线、下划线及小数点，且不能以小数点、下划线开头或以小数点结尾。
日志存储时间（天）	<p>设置日志存储时间。</p> <p>日志数据默认存储30天，取值范围为1~30天，超出存储时间的日志将会被自动化删除，可以按需将日志数据转储至OBS桶中长期保存。创建日志组免费，使用阶段按照日志量收费。</p>
企业项目	选择企业项目。
标签管理	单击“添加标签”添加标签，并设置标签键值。

表 4-5 创建日志流

参数	参数说明
日志组名称	日志组名称，不支持修改。
日志流名称	输入日志流名称，只能包含英文字母、数字、中文、中划线、下划线及小数点，且不能以小数点、下划线开头或以小数点结尾。
日志存储时间（天）	<p>设置日志存储时间。</p> <p>日志数据默认存储30天，取值范围为1~30天，超出存储时间的日志将会被自动化删除，可以按需将日志数据转储至OBS桶中长期保存。创建日志组免费，使用阶段按照日志量收费。</p>

参数	参数说明
企业项目	选择企业项目。
标签管理	单击“添加标签”添加标签，并设置标签键值。

步骤5 端点创建后，可在列表中查看相关信息，端点状态变为“已就绪”后即可提交作业到该端点中运行。

----结束

后续操作

Aura端点创建完成后，即可在DataArts Studio作业开发页面提交一个单任务Shell或Pipeline类型的AI DataLake作业至端点运行，详细操作请参见[作业开发](#)。

4.3 管理 Aura 引擎端点

操作场景

Aura端点创建成功后，可在端点列表查看详细信息，也可以修改或删除端点。

查看 Aura 引擎端点详情

步骤1 登录AI DataLake管理控制台。

步骤2 在页面左上角切换至对应的工作空间。

步骤3 在左侧导航栏选择“引擎管理 > 多模数据引擎Aura”进入Aura引擎端点列表页面，即可查看已创建的端点信息。

表 4-6 查看端点信息

参数	参数说明	示例
名称	已创建的端点名称。	test
类型	端点类型。	Job

参数	参数说明	示例
状态	端点状态。 <ul style="list-style-type: none"> 已就绪：端点创建成功，且正在运行中。仅该类状态的端点支持提交作业。 未就绪：端点创建失败、端点修改失败或端点删除失败。 冻结：端点不可用。 删除中：正在删除端点。 准备中：正在创建端点。 	已就绪
资源使用模式	端点的资源使用模式。	预留模式
资源池	端点所属的资源池。	-
创建时间	创建端点的时间。	2026/03/26 14:25:15 GMT+08:00
更新时间	更新端点的时间。	2026/03/26 14:25:18 GMT+08:00
创建者	创建端点的用户。	-
操作	已创建的端点支持的操作。 <ul style="list-style-type: none"> 修改：可单击“修改”，修改端点名称、CPU资源、GPU资源、NPU资源、日志对接LTS等配置，并单击“确定” 删除：如果端点不再使用，可单击端点所在行的“删除”删除端点。 	-

步骤4 单击端点名称可进入端点概览页面，可查看表4-7中的相关信息。

表 4-7 端点概览信息

类别	参数	参数说明	示例
基本信息	端点名称	已创建的端点名称，可单击修改按钮进行修改。	test
	端点类型	Aura端点类型。	Job

类别	参数	参数说明	示例
	创建人	创建端点的用户。	-
	创建时间	创建端点的时间。	2026/03/26 14:25:15 GMT+08:00
	端点 ID	端点的ID。	-
	描述	端点的描述信息。	-
	资源使用模式	端点的资源使用模式。	预留模式
	资源池	端点所属的资源池名称。	-
日志配置 (仅创建 Aura端点时 开启“日志 对接LTS”功 能后才显示)	日志转存储服务	端点运行的作业的日志转存储服务，仅支持对接LTS服务。	云日志服务 LTS
	日志组	对接LTS后，存储日志的日志组ID。可单击日志组ID跳转至LTS服务的日志管理页面，执行相关管理操作。	-
	日志流	对接LTS后，存储日志的日志流组ID。可单击日志流ID跳转至LTS服务的日志流详细信息页面。	-

步骤5 选择“资源配置”页签，可查看端点的资源配置信息，包括规格、最大值和最小值。

步骤6 选择“作业运行历史”，可在列表中查看端点运行的作业信息。

表 4-8 作业运行历史信息

参数	参数说明	示例
运行名称/运行 ID	运行的作业名称和作业ID。	-
状态	作业的运行状态。 <ul style="list-style-type: none"> ● 成功：作业运行成功。 ● 失败：作业运行失败。 ● 停止中：正在停止作业。 ● 运行中：正在运行作业。 ● 已停止：作业停止成功。 	成功

参数	参数说明	示例
创建时间	创建作业的时间。	2026/04/07 11:36:38 GMT+08:00
执行时间	作业开始执行的时间。	2026/04/07 11:36:38 GMT+08:00
完成时间	作业运行的完成时间。	2026/04/07 11:41:31 GMT+08:00
提交人	提交作业的用户名称。	-
操作	单击“查看详情”，可查看作业运行详情，详细介绍请参见表4-9。	-

表 4-9 作业运行详情

参数	参数说明	示例
作业运行名称	运行的作业名称。	-
运行状态	作业的运行状态。 <ul style="list-style-type: none"> • 成功：作业运行成功。 • 失败：作业运行失败。 • 停止中：正在停止作业。 • 运行中：正在运行作业。 • 已停止：作业停止成功。 	成功
运行端点名称	运行作业的Aura端点名称。	test
作业运行 ID	运行的作业ID。	-
创建时间	创建作业的时间。	2026/04/07 11:36:38 GMT+08:00
运行时长（秒）	作业的运行时长。	4664
Job 版本	Job版本。	-
创建人	创建作业的用户名称。	-
日志组 ID	对接LTS后，存储日志的日志组ID。	-
日志流 ID	对接LTS后，存储日志的日志流组ID。	-

参数	参数说明	示例
作业参数	作业配置的相关参数。	-

----结束

修改/删除 Aura 引擎端点

步骤1 登录AI DataLake管理控制台。

步骤2 在页面左上角切换至对应的工作空间。

步骤3 在左侧导航栏选择“引擎管理 > 多模数据引擎Aura”进入Aura引擎端点列表页面，即可查看已创建的端点。

- 如果需要修改端点信息，可单击端点所在行的“修改”，在右侧的修改引擎端点弹框中可修改端点名称、CPU资源、GPU资源、NPU资源、日志对接LTS等配置，单击“确定”。
- 如果端点不再使用，可单击端点所在行的“删除”删除端点。



删除端点时请确保作业已运行完成。

----结束

5 使用 DataArts Studio 开发 AI DataLake 作业

在AI DataLake完成引擎端点的配置后，您就已经具备了提交作业的必要条件：计算资源和引擎的配置。

接下来将要依托**DataArts Studio**的作业开发模块完成数据查询、数据处理等任务。本节操作介绍使用**DataArts Studio**完成作业开发的关键操作指导。

在AI DataLake的快速入门中，以Aura Job为例介绍了完整的作业开发流程，了解更多请参见[快速上手使用AI DataLake查询数据](#)。

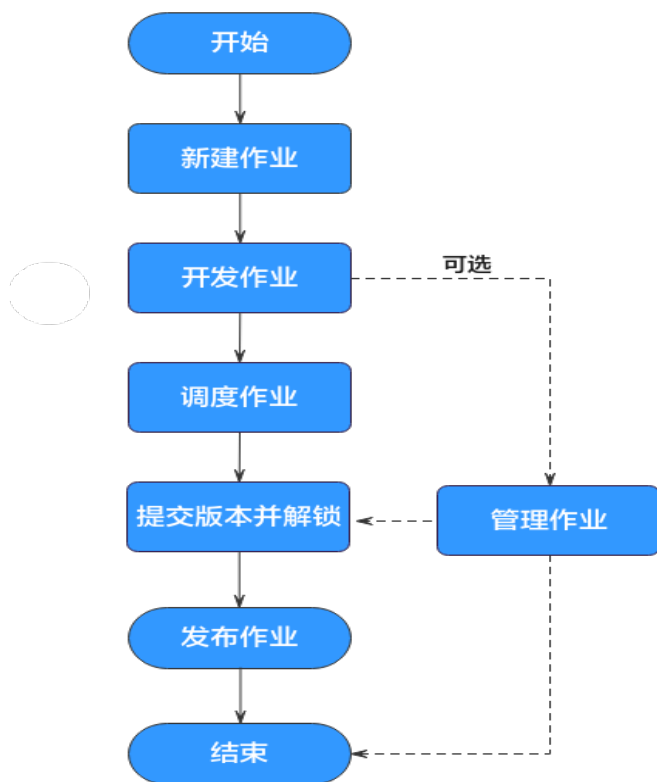
作业开发前的准备工作

- 开通DataArts Studio。
- 已在DataArts Studio创建工作空间。

创建并开发作业

开发作业前，您可以通过[图5-1](#)了解DataArts Studio数据开发模块作业开发的基本流程。

图 5-1 作业开发流程



1. 新建作业：当前提供两种作业类型：批处理和实时处理，分别应用于批量数据处理和实时连接性数据处理，其中批处理作业支持Pipeline和单节点作业两种模式，具体请参见[新建作业](#)。
2. 开发作业：基于新建的作业，进行作业开发，您可以进行编排、配置节点。具体请参见[开发Pipeline作业](#)。
3. 调度作业：配置作业调度任务。具体请参见[调度作业](#)。
 - 如果您的作业是批处理作业，您可以配置作业级别的调度任务，即以作业为一个整体进行调度，支持单次调度、周期调度、事件驱动调度三种调度方式。具体请参见[配置作业调度任务（批处理作业）](#)。
 - 如果您的作业是实时处理作业，您可以配置节点级别的调度任务，即每一个节点可以独立调度，支持单次调度、周期调度、事件驱动调度三种调度方式。具体请参见[配置节点调度任务（实时作业）](#)。
4. 提交版本并解锁：作业调度配置完成后，您需要提交版本并解锁，提交版本并解锁后才能用于调度运行，便于其他开发者修改。具体请参见[提交版本](#)。
5. （可选）管理作业：作业开发完成后，您可以根据需要，进行作业管理。具体请参见[（可选）管理作业](#)。
6. 发布作业。企业模式下需要发布作业，具体请参见[发布作业任务](#)。

在 AI DataLake Aura 端点页面查看作业运行信息

步骤1 登录AI DataLake管理控制台。

步骤2 在页面左上角切换至对应的工作空间。

步骤3 在左侧导航栏选择“引擎管理 > 多模数据引擎Aura”进入Aura引擎端点列表页面，单击运行作业的端点名称进入概览页面。

步骤4 选择“作业运行历史”，在列表中即可查看作业的运行状态，“成功”即表示作业执行成功。

步骤5 也可单击“查看详情”查看作业详细信息。

表 5-1 作业运行详情

参数	参数说明	示例
作业运行名称	运行的作业名称。	-
运行状态	作业的运行状态。 <ul style="list-style-type: none">成功：作业运行成功。失败：作业运行失败。停止中：正在停止作业。运行中：正在运行作业。已停止：作业停止成功。	成功
运行端点名称	运行作业的Aura端点名称。	test
作业运行 ID	运行的作业ID。	-
创建时间	创建作业的时间。	2026/04/07 11:36:38 GMT+08:00
运行时长（秒）	作业的运行时长。	4664
Job 版本	Job版本。	-
创建人	创建作业的用户名称。	-
日志组 ID	对接LTS后，存储日志的日志组ID。	-
日志流 ID	对接LTS后，存储日志的日志流组ID。	-
作业参数	作业配置的相关参数。	-

----结束

6 管理数据与元数据

6.1 了解数据目录、数据库和表

数据库和表是AI DataLake作业开发的基础，在执行作业前您需要根据业务场景规划数据库和表。

AI DataLake使用LakeFormation提供元数据管理、数据权限管理、数据接入管理等操作。

本节操作介绍数据目录、数据库和表的基本概念。

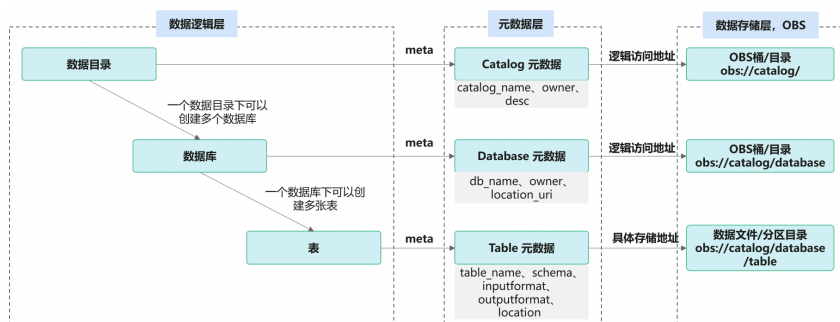
数据目录与数据库和表的关系

如图6-1所示，

数据目录负责管理元数据，数据库负责分类，表负责描述具体数据集，OBS负责存储数据，而Lake Formation负责数据统一权限控制和治理。

- 数据目录记录数据存放在哪里、表结构是什么、数据格式如何以及谁有权限访问这些数据。
- 数据库是数据目录中的逻辑分类单位，用来把一组相关的表组织在一起。
- 表则是对具体数据集的定义，描述了底层数据在OBS桶中的位置、字段、分区和格式等信息。

图 6-1 数据目录与数据库和表的关系



数据目录

数据目录就是LakeFormation中保存元数据的管理体系，可以统一管理所有数据库，记录数据库的位置、权限、元数据（描述数据的数据），即

数据目录并不是直接存储数据内容，而是保存“关于数据的数据”，即云数据。

数据库

数据库是数据目录的子单元，它是按照数据结构来组织、存储和管理数据的，通常用于存储、检索和管理结构化数据，由多个数据表组成，这些数据表通过键和索引相互关联。

表

表是数据库最重要的组成部分，是结构化存储数据的二维表格，它由行和列组成。每一行代表一个数据项，每一列代表数据的一个属性或特征。表用于组织和存储特定类型的数据，使得数据可以被有效地查询和分析。

数据库是一个框架，表是其实质内容。一个数据库包含一个或者多个表。

说明

创建数据库和表时，有权限控制，需要对其他用户授权，其他用户才可查看该用户新建的数据库和表。

表的元数据

元数据（Metadata）是用来定义数据类型的数据。主要是描述数据自身信息，包含源、大小、格式或其它数据特征。数据库字段中，元数据用于诠释数据仓库的内容。

创建表时，会定义元数据，由列名、类型、列描述三列组成。

数据库和表的约束与限制

表 6-1 DLI 资源相关约束限制

限制项	说明
数据库	<ul style="list-style-type: none">“default”为内置数据库，不能创建名为“default”的数据库。DLI支持创建的数据库的最大数量为50个。
数据表	<ul style="list-style-type: none">DLI支持创建的表的最大数量为5000个。DLI支持创建表类型：<ul style="list-style-type: none">Managed：数据存储位置为DLI的表。External：数据存储位置为OBS的表。View：视图，视图只能通过SQL语句创建。跨源表：表类型同样为External。创建DLI表时不支持指定存储路径。

限制项	说明
数据导入	<ul style="list-style-type: none"> 仅支持将OBS上的数据导入DLI或OBS中。 支持将OBS中CSV, Parquet, ORC, JSON和Avro格式的数据导入到在DLI中创建的表。 将CSV格式数据导入分区表, 需在数据源中将分区列放在最后一列。 导入数据的编码格式仅支持UTF-8。
数据导出	<ul style="list-style-type: none"> 只支持将DLI表(表类型为“Managed”)中的数据导出到OBS桶中, 且导出的路径必须指定到文件夹级别。 导出文件格式为json格式, 且文本格式仅支持UTF-8。 支持跨账号导出数据, 即B账户对A账户授权后, A账户拥有B账户OBS桶的元数据信息和权限信息的读取权限, 以及路径的读写权限, 则A账户可将数据导出至B账户的OBS路径中。

表管理页面

在“数据管理”页面中, 单击对应数据库名称或“操作”列中的“表管理”, 可进入其表管理页面。

表管理页面显示用户在当前数据库中创建所有的表, 您可以查看表类型, 数据存储位置等信息。表列表默认按创建时间排列, 创建时间最近的表显示在最前端。

6.2 创建数据目录、数据库和表

在使用AI DataLake提交作业前, 您需要先在LakeFormation创建一个实例, 后续的操作, 如管理元数据、设置元数据权限等, 都是基于您创建的实例进行的。

本节分别给出在LakeFormation创建实例、管理元数据、配置数据权限、将数据迁移至当前LakeFormation实例的操作说明, 详细的操作指导可以参考LakeFormation服务的相关帮助文档。

创建 LakeFormation 实例

- 在使用LakeFormation之前, 首先需要创建一个实例, 后续的操作, 如管理元数据、设置元数据权限等, 都是基于您创建的实例进行的。
详情请参考[创建LakeFormation实例](#)。
- LakeFormation提供实例的配置、扩容、删除等基本功能, 帮助您便捷地进行实例管理, 加速实现数据湖承载的业务规划和部署。
详情请参考[管理LakeFormation实例](#)。

管理 LakeFormation 元数据

LakeFormation提供Catalog、数据库、数据表等的创建、修改、查看、删除等功能, 并支持配置元数据生命周期。帮助您便捷地进行数据湖初始化构建以及持续运营, 集中式地统一管理LakeFormation实例下所有的元数据, 加速实现数据湖承载的业务规划和部署。

详情请参考[创建LakeFormation元数据](#)和[管理LakeFormation元数据](#)章节。

管理数据权限

LakeFormation提供针对Catalog、数据库、数据表等数据资源的授权、取消、查看等功能。帮助您对数据湖实现便捷的统一的数据权限管理。

详情请参考[配置LakeFormation元数据权限](#)和[管理LakeFormation数据权限](#)章节。

将数据迁移至 LakeFormation 实例

- LakeFormation支持将外部服务的元数据及其权限迁移至当前LakeFormation实例中，对元数据及权限进行统一管理。支持元数据以及权限比对，用来确认迁移的一致性。

详情请参考[迁移元数据及权限至LakeFormation](#)。

- LakeFormation提供统一的接入管理能力，用户可以通过创建接入客户端的方式为指定的客户端环境建立网络连接通道，同时可以在客户端详情中查看接入IP等信息，用于其他云服务接入LakeFormation实例。

详情请参考[创建LakeFormation接入客户端](#)。

7 监控 AI DataLake 计算资源池

在日常运维工作中，管理员需要实时掌握计算资源池的运行状态，了解CPU、内存、NPU、GPU 等资源的实际使用情况。

通过可视化监控，管理员可以：

- 直观了解资源池中各类资源的实时使用情况。
- 判断是否需要资源池的扩缩容操作。
- 优化资源管理，避免资源浪费或不足。
- 保障业务稳定，及时发现资源瓶颈。

查看资源池监控

步骤1 登录AI DataLake管理控制台。

步骤2 在左侧导航栏单击“运维监控 > 资源池监控”。

步骤3 在资源池监控页面中，可查看当前工作空间下，资源池的资源使用信息。

图 7-1 查看资源池监控

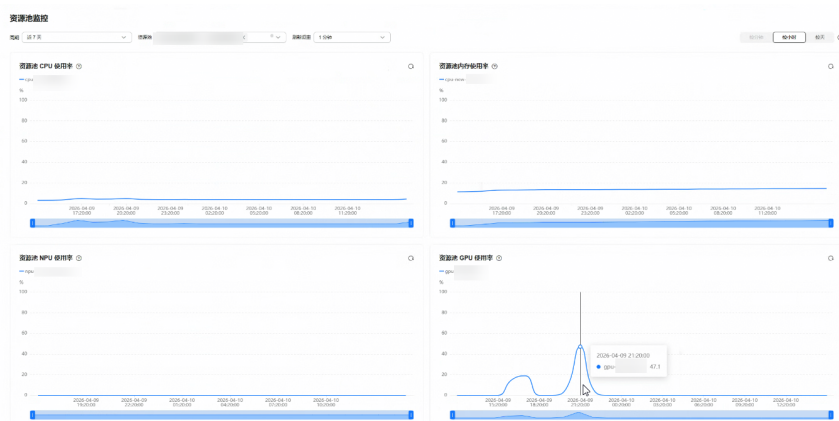


表 7-1 资源池监控指标

指标名称	单位	描述
资源池CPU使用率	百分比	当前计算资源池中所有计算节点的CPU资源被实际占用的比例，反映了资源池处理通用计算任务的负载情况。
资源池内存使用率	百分比	当前计算资源池中所有计算节点的内存资源被实际占用的比例。内存是影响数据处理任务性能的关键因素，尤其对于大数据分析、内存计算等场景至关重要。
资源池NPU使用率	百分比	当前计算资源池中NPU加速器的使用比例，NPU专门用于加速深度学习推理和训练任务。
资源池GPU使用率	百分比	当前计算资源池中GPU加速器的使用比例。GPU凭借其强大的并行计算能力，广泛应用于深度学习、科学计算、图形渲染等领域。

步骤4 运维人员可以通过配置监控参数来查看计算资源池的实时运行状态，系统提供了灵活的监控配置选项，支持自定义监控数据统计周期、选择特定资源池以及调整页面刷新频率，满足不同场景下的监控需求。设置监控数据统计周期后，在界面右上角可选择“按分钟”、“按小时”或者“按天”的维度查看图表，图表数据会按照所选聚合粒度计算平均值呈现。

- 监控数据统计周期为近1小时/6小时，监控图表内的数据周期最小为1分钟。
- 监控数据统计周期为近1天/7天，监控图表内的数据周期最小为1小时，图表不支持按分钟呈现。
- 监控数据统计周期为近30天，监控图表内的数据周期最小为1天，图表不支持按分钟或者按小时呈现。
- 自定义监控数据统计周期最大选择30天，超过1天时监控图表内的数据周期最小为1小时，超过7天时监控图表内的数据周期最小为1天。

----结束

8 使用 LTS 管理 AI DataLake 作业日志

在大数据作业开发、调试和线上运维过程中，用户对日志的实时查看、检索和分析能力提出了更高的要求。为了提高日志管理的效率和用户体验，AI DataLake提供了将Job标准输出日志自动投递到云日志服务（Log Tank Service，简称LTS）的功能。只需在AI DataLake控制台配置端点绑定LTS，系统将自动完成日志的实时采集、查询和分析，从而简化日志管理流程，提高问题定位和故障排查的效率，同时确保日志数据的安全隔离，满足多租户场景下的安全与合规要求。

约束与限制

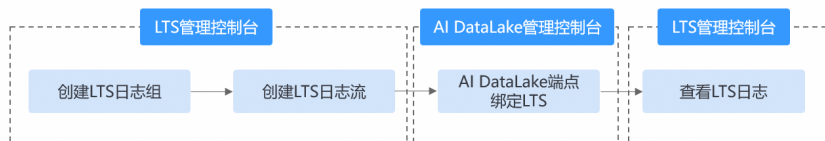
- 仅Aura Job支持配置AI DataLake对接LTS日志流。
- 每个端点仅支持对接一个LTS日志流。
- 一个日志流支持对接多个Aura端点。
- 端点绑定LTS需要具备以下权限，请参考[表8-1](#)。

表 8-1 端点绑定 LTS 需要具备的权限

权限类型	API	角色与策略授权	身份与策略授权
创建端点权限	POST /v2/workspaces/{workspace_id}/endpoints	DataArtsFabric:endpoint:create	DataArtsFabric:endpoint:create
读取日志组权限	GET /v2/{project_id}/groups	lts:groups:get	lts:logGroup:listLogGroup
读取日志流权限	GET /v2/{project_id}/groups/{log_group_id}/streams	lts:topics:get	logStream:listLogStream

操作流程

图 8-1 创建程序包



使用LTS管理AI DataLake作业日志分为以下操作步骤

1. **步骤1: 创建LTS日志组:** 日志组 (LogGroup) 是云日志服务进行日志管理的基本单位, 用于对日志流进行分组。
2. **步骤2: 创建LTS日志流:** 日志采集后, 以日志流为单位, 将不同类型的日志分类存储在不同的日志流上, 方便对日志进一步分类管理。
3. **步骤3: 创建端点绑定LTS:** 所有使用端点提交的Aura Job作业的日志将通过绑定的LTS存储在对应的日志组和日志流中。
4. **步骤4: 查看LTS日志:** 绑定LTS后即可根据LTS支持存储的日志类型搜索和查看日志。

基本概念

本节介绍LTS相关的基本概念, 便于您通过云日志服务快速高效地进行设备运维管理、用户业务趋势分析、安全监控审计等操作。

日志组

日志组 (LogGroup) 是云日志服务进行日志管理的基本单位, 用于对日志流进行分类, 一个日志组下面可以创建多个日志流。日志组本身不存储任何日志数据, 仅方便用户管理日志流。

日志流

日志流 (LogStream) 是日志读写的基本单位。日志采集后, 以日志流为单位, 将不同类型的日志分类存储在不同的日志流上, 方便对日志进一步分类管理。如果日志较多, 需要分门别类, 建议您创建多个日志流, 并给日志流做好命名, 方便后续快速查找日志。

步骤 1: 创建 LTS 日志组

日志组 (LogGroup) 是云日志服务进行日志管理的基本单位, 用于对日志流进行分组, 一个日志组下面可以创建多个日志流。日志组本身不存储任何日志数据, 仅方便您管理日志流。

一个日志组通常对应公司内的某一个项目/业务, 建议将某个项目/业务下的多个应用/服务的日志流归属到同一个日志组下。当公司有多个项目时, 具体的项目人员只需要查看所属项目对应的日志组下的日志流即可, 其它项目的日志流不会对其产生干扰。


1. 登录[云日志服务控制台](#)。
2. 进入“日志管理”页面, 单击“创建日志组”。
3. 在“创建日志组”页面中配置LTS日志的相关参数。

- 本例设置创建日志组名称：lts-group-aura
 - 日志存储时间：30天
- 更多参数说明请参考[创建日志组](#)。
4. 单击“确定”，日志组创建成功，即可在日志组列表下方生成一条日志组信息。

步骤 2：创建 LTS 日志流

云日志服务以日志流（LogStream）作为日志管理维度。每个日志流只能属于一个日志组，每个日志组中可创建多个日志流。

日志采集后，以日志流为单位，将不同类型的日志分类存储在不同的日志流上，方便对日志进一步分类管理。如果日志较多，需要分门别类，建议您创建多个日志流，并给日志流做好命名，方便后续快速查找日志。例如，您可以将操作日志、访问日志等接入不同的日志流，查询日志时可以进入对应的日志流快速查看日志。支持按照业务需求对不同的日志流添加对应的标签，方便运维人员管理业务。

1. 登录[云日志服务控制台](#)。
2. 单击日志组名称对应的 。
3. 单击“创建日志流”，在创建日志流页面，填写日志流相关信息。
本例创建日志流参数配置如下：
 - 日志流名称：lts-topic-spark001
 - 日志存储：开启
 - 日志存储时间：使用日志组的存储时间更多参数说明请参考[创建日志流](#)。
4. 单击“确定”，完成日志流的创建，在日志组下方生成一条日志流信息。

步骤 3：创建端点绑定 LTS

1. 登录AI DataLake控制台。
2. 选择“多模数据引擎 Aura”。
3. 单击“创建端点”并配置参数“日志对接LTS”。
本例创建日志流参数配置如下：
 - 日志组：选择[步骤1：创建LTS日志组](#)中创建的日志组。
 - 日志流：选择[步骤2：创建LTS日志流](#)中创建的日志流。更多参数说明请参考[创建日志流](#)。
4. 单击“确定”，完成端点和LTS日志组与日志流的绑定。
完成绑定后，所有使用该端点提交的Aura Job作业的日志将通过绑定的LTS存储在对应的日志组和日志流中：
 - 运行中的作业：从绑定开始后产生的日志将存储在LTS中。
 - 新提交的作业：所有日志都将存储在LTS中。

步骤 4：查看 LTS 日志

请选择使用该端点上运行任一Aura job作业。

1. 登录AI DataLake控制台。
2. 单击“多模数据引擎 Aura > 端点名称 > 作业历史”，查看作业执行历史。

- 记录对应的作业名称。
- 登录LTS日志流页面，在搜索对应的作业名称，即可查看对应的作业日志。

图 8-2 在 LTS 中查询日志



相关信息：LTS 支持存储的日志类型

表 8-2 LTS 支持存储的日志类型

logType	含义	对应的节点路径
Aura Job	Aura job标准输出日志	采集容器标准输出（stdout.log和stderr.log）

9 使用 CTS 审计 AI DataLake 服务

通过云审计服务，您可以记录与AI DataLake服务相关的操作事件，便于日后的查询、审计和回溯。

表 9-1 云审计服务支持的 AI DataLake 操作列表

操作名称	事件名称	资源类型
注册租户协议	createV2Agreement	agreement
删除用户注册协议	deleteV2Agreement	V2Agreement
创建V2工作空间	createV2Workspace	workspace
更新V2工作空间	updateV2Workspace	workspace
删除V2工作空间	deleteV2Workspace	workspace
创建服务委托	createAgency	Agency
删除服务委托	deleteAgency	Agency
创建V2计算资源	createV2Compute	compute
更新V2计算资源	updateV2Compute	compute
删除V2计算资源	deleteV2Compute	compute
创建V2端点	createV2Endpoint	endpoint
更新V2端点	updateV2Endpoint	endpoint
删除V2端点	deleteV2Endpoint	endpoint
创建v2网络配置	CreateV2Network	V2Network
创建网络连接	createV2NetworkConnecti on	connection
删除网络连接	deleteV2NetworkConnecti on	connection
创建工作空间	createWorkspace	workspace

操作名称	事件名称	资源类型
更新工作空间	updateWorkspace	workspace
删除工作空间	deleteWorkspace	workspace
更新AOM监控采集配置	updateMetricsConfig	endpoint
删除指定作业的指定版本	deleteJobVersion	job
创建作业	createJob	job
更新作业	updateJob	job
删除指定作业	deleteJob	Job
创建服务	createService	service
删除服务的某个指定版本	deleteServiceVersion	service
更新服务	updateService	service
删除未使用的服务	deleteService	service
注册镜像包 用户想要使用自定义镜像去创建ray集群或者ray服务时，需要先创建镜像包。此接口为异步接口，配套使用接口ListCaps查看镜像包具体信息，无特殊场景。	createCap	cap
更新镜像包 当需要修改镜像包信息，如描述信息时，需要使用此接口。此接口为同步接口，无配套使用接口和特殊场景。	updateCap	cap
删除指定镜像包 当需要删除镜像包的时候，需要使用此接口。此接口为同步接口，无配套使用接口和特殊场景。	deleteCap	cap
删除某个镜像包的某一版本	deleteCapVersion	cap
添加计算资源	createCompute	compute
删除计算资源	deleteCompute	compute
变更计算资源	updateCompute	compute
注册租户协议	createAgreement	agreement
删除用户注册协议	deleteAgreement	Agreement
创建服务委托	createAgency	Agency

操作名称	事件名称	资源类型
删除服务委托	deleteAgency	Agency
创建端点	createEndpoint	endpoint
修改端点	updateEndpoint	endpoint
删除端点	deleteEndpoint	endpoint
批量添加资源标签	BatchCreateWorkspaceTags	tags
批量删除资源标签	BatchDeleteWorkspaceTags	tags
创建消息通知策略	createMessageNotificationPolicy	messageNotificationPolicy
删除消息通知策略	deleteMessageNotificationPolicy	messageNotificationPolicy
订阅端点	subscribeEndpoint	endpoint
创建网络连接	createNetwork	network
删除网络连接	deleteNetwork	network
创建缓存实例	createCache	cache
查询缓存实例列表	listCache	cache
查询缓存实例详情	showCache	cache
删除缓存实例 需要删除某个缓存时调用，传入具体需要删除的缓存ID。此接口为异步接口，配套使用接口ShowCache接口查询单个缓存实例详情信息，无特殊场景。	deleteCache	cache
注册自定义镜像 用户想要使用自定义镜像去创建ray集群或者ray服务时，需要先注册自定义镜像。此接口为异步接口，配套使用接口ListImages查看镜像具体信息，无特殊场景。	createImage	image
删除指定镜像包。	deleteImage	image
更新自定义镜像信息 当需要修改自定义镜像信息，如描述信息时，需要使用此接口。此接口为同步接口，无配套使用接口和特殊场景。	updateImage	image

操作名称	事件名称	资源类型
创建appinstance副本	createAppInstanceReplicas	appInstance
创建路由信息	createRoute	routes
删除appinstance副本	batchDeleteAppInstanceReplicas	appInstance
删除作业	deleteJobRun	jobRun
删除服务实例	deleteServiceInstance	serviceInstance
删除端点下所有的作业	clearJobRun	jobRun
删除部署的App实例	deleteAppInstance	appInstance
发起调用请求	InvokeModelService	ModelService
取消作业运行	cancelJobRun	jobRun
执行SQL	closeSession	sql
批量删除appinstance关联的引擎实例	batchDeleteAppInstanceRelationEngines	appInstance
更新App实例	updateAppInstance	AppInstance
更新已部署的服务实例	updateServiceInstance	serviceInstance
校验路由的session id是否合法，由承载租户调用	checkSessionValid	session
根据路由id删除指定路由信息	deleteRoute	routes
获取Session信息	showSession	session
获取Sql Session列表	createSession	sql
获取Statement结果、执行状态	executeStatement	sql
运行作业	createJobRun	jobRun
部署服务	createAppInstance	appInstance
部署服务	createServiceInstance	serviceInstance

关于如何开通云审计服务以及如何查看追踪事件，请参考《[云审计服务快速入门](#)》中的相关章节。

关于云审计服务事件结构的关键字段详解，请参见《[云审计服务用户指南](#)》中的[事件结构](#)和[事件样例](#)。