



应用托管平台 用户操作手册

天翼云科技有限公司

一、产品介绍.....	4
1.产品定义	4
2.产品优势	4
3.功能特性	5
4.应用场景	5
4.1serverless 应用部署	5
4.2 专属资源应用部署	6
5.资源池限制	6
二、快速入门.....	7
1.准备工作	7
2.购买应用	8
2.1 通过应用市场/MCP 市场部署应用	8
2.2 自定义部署应用	9
三、操作指导.....	11
1.应用市场	12
2.MCP 市场	13
3.我的应用	15
4.应用实例	16
4.1 创建公共资源应用实例	16
4.2 创建独享资源应用实例	19
4.3 查看应用实例	21
4.4 访问实例	26

4.5 启动服务	26
4.6 停止服务	27
4.7 更新服务	28
4.8 扩缩容	31
4.9 删除服务	32
5.访问控制	33
5.1 创建访问控制策略	33
5.2 访问控制策略管理	35
6.资源组管理	39
6.1 查看资源组	39
6.2 管理资源组	40
7.项目空间	40
7.1 创建项目空间	40
7.2 项目空间管理	42

一、产品介绍

1.产品定义

一站式应用托管平台，免运维底层资源，提供丰富的应用资源，支持跨域算力分配和分布式服务部署，实现应用构建、部署与运维的全生命周期管理能力。



2.产品优势

丰富的应用资源

汇聚不同类型的应用，覆盖多种用途和行业需求。可方便地浏览、搜索选择适合的应用。

开箱即用部署

提供应用一键部署的功能，极大简化应用的安装和配置过程，轻松上手。

便捷访问

为部署的应用提供公网访问能力，可通过公网便捷访问已部署的应用。

灵活运维

提供实例级别的日志、监控，满足不同维度的运维需求。

3.功能特性

应用无差异部署

基于云原生能力和部署工具链的集成，提升部署效率和降低运维成本。

跨域算力分配

基于跨域算力动态分配算力源到最优位置，支持应用分布式部署，实现故障分钟级切换。

应用全生命周期管理

提供全面的应用管理能力，包括服务调用、启动、停止、更新、扩缩容等关键操作，实现高效的运维管理。

项目权限隔离

支持完善的项目空间管理能力，实现多项目、多团队的高效协作与资源隔离，灵活配置人员、角色与权限体系，实现对成员访问范围与操作权限的精细化控制。

4.应用场景

4.1serverless 应用部署

适用场景

开发、测试环境搭建，中小型应用或微服务

产品优势

按需资源： 以 Serverless SaaS 的方式按需使用算力，无须关注底层 IaaS/PaaS，平台全托管免运维

快速部署： 应用快速部署和启动

灵活管理： 随时启停应用，停止应用后不计费

4.2 专属资源应用部署

适用场景

大型 Web 应用与关键业务系统搭建，多组件部署

产品优势

安全性强： 用户独占的专属集群，数据安全性高，性能稳定

自主性强： 可自主分配资源，可将多个应用部署到同个专属资源组内，资源利用率高

使用限制

5.资源池限制

当前开放**云南-昆明、贵州-安顺**资源池，其他资源池请持续关注。

配额限制

使用应用托管平台时，需注意配额限制，具体如下表所示。如果资源池或者配额不能满足业务使用，可通过联系客户经理或工单申请调整资源池和配额。

公共资源算力配额

分类	限制数量	单位
CPU	500	核
内存	1000	GB
存储	5000	GB

资源售罄

在购买算力资源时，所选的产品规格可能已售罄，如果没有需要的产品规格，请联系客户经理或通过工单反馈，我们会根据资源库存情况，补充对应的产品规格。

二、快速入门

1.准备工作

1.1 注册天翼云账号

如果您已经拥有一个天翼云实名账号可直接跳过本步骤，如果您还没有天翼云账号，具体可参见[注册天翼云账号](#)。

1.2 实名认证

具体可参见[实名认证](#)模块对天翼云账号进行实名认证。

1.3 账户充值

您需要确保天翼云实名账户中的余额充足，具体天翼云账户充值步骤请参见[账户充值](#)。

2.购买应用

应用托管提供两种部署应用的方式。

2.1 通过应用市场/MCP 市场部署应用

应用市场/MCP 市场提供优质的应用资源，支持通过关键词搜索、分类筛选等方式快速定位所需应用，一键部署。

操作步骤：

1. 进入天翼云应用托管平台，点击左侧【应用市场】或【MCP 市场】。



2.选择应用点击对应的【立即部署】按钮，进入应用详情页。

MCP市场 / 应用部署



Wikipedia MCP 服务器

热门

wiki

这是一个 Model Context Protocol (MCP) 服务器，可以从维基百科检索信息，为大型语言模型 (LLMs) 提供上下文。这个工具帮助 AI 助手从维基百科获取事实信息，以可靠来源为基础进行响应。

总计

预估 ¥ 0.27 元/小时

资源类型

1c1g

访问策略

app-default-sg

刷新

如需新的应用访问策略, 可点击 [去创建](#)

公网访问

启用

☐ 我已阅读, 理解并接受 《算力互联网平台产品服务协议》 《中国电信天翼云用户协议》

立即部署

免责声明: 平台提供标准服务, 不对您使用本服务的合法性负责。请您遵守国家法律法规, 合法合规使用本服务。

应用详情

Wikipedia MCP 服务器

这是一个 Model Context Protocol (MCP) 服务器，可以从维基百科检索信息，为大型语言模型 (LLMs) 提供上下文。这个工具帮助 AI 助手从维基百科获取事实信息，以可靠来源为基础进行响应。

概述

Wikipedia MCP 服务器通过标准化的 Model Context Protocol 接口提供对维基百科信息的实时访问。这使得 LLMs 能够直接从维基百科检索准确且最新的信息，以增强其响应。

特性

- 搜索维基百科：查找与特定查询匹配的文章
- 检索文章内容：获取包含所有信息的完整文章文本
- 文章摘要：获取文章的简明摘要
- 部分提取：从文章中检索特定部分
- 链接发现：在文章中找到相关主题的链接
- 相关主题：发现与特定文章相关的主题
- 多语言支持：访问不同语言的维基百科

可用的 MCP 工具

Wikipedia MCP 服务器提供了以下工具，供 LLMs 与维基百科交互：

3.勾选协议，点击【立即订购】按钮，进入订单支付页面。

4.应用购买成功后，可在应用托管控制台-我的应用页面，查看已购买的应用。点击【开始部署】，可以再次部署该应用。点击【查看实例】按钮，可查看当前应用已部署的应用实例。

我的应用

Wikipedia MCP 服务器

应用工具

管理

AI

一个帮助AI助手从维基百科检索信息的模型上下文信息服务器。

开始部署

查看实例

mcp-计算器

MCP

A Model Context Protocol server for calculating. This server enables LLMs to use calculator for precise numerical calculations.

开始部署

查看实例

mcp-github-trending

MCP

一个通过GitHub API接口获取GitHub热门开源项目开发数据(MCP)服务器。

开始部署

查看实例

mcp-server-tmdb

MC

MCP

该MCP服务器与The Movie Database (TMDB) API集成，提供电影信息、搜索功能和推荐。

开始部署

查看实例

共4项

1 2 3 4

1

应用实例管理

新增应用实例

应用名称

mcp-github-trending

搜索

实例名称ID	状态	实例ID	应用信息	创建时间	更新时间	操作
mcp-github-trending-291112 ID: ef6b72ab73a0a0b0e0	已停止	8-1	mcp-github-trending 02 app-ef6b72ab73a0a0b0e0	2025-08-29 17:38:18	2025-08-29 18:32:18	启动 重启 删除

共1项

1 2 3 4

1

2.2 自定义部署应用

通过自定义部署参数进行应用的部署，适用于管理与运维自己的应用服务。



操作步骤：

1.登录应用托管控制台，在左侧导航栏单击【应用实例】， 点击【创建应用实例】按钮。



2.应用配置： 选择自定义应用， 输入实例名称。



3.资源配置： 资源类型选择【公共资源】， 付费类型当前支持按量付费，选择算力规格，包
括资源池、是否支持加速卡、地域、规格和实例数量。



4.网络配置：支持配置公网访问，开启后将为服务分配公网访问 ip。访问策略为服务的公网
访问提供安全防护，可根据需要调整访问策略内容。一个服务只能绑定一个访问策略。



5.参数配置： 配置服务的镜像，支持填写镜像地址（私有仓库需提供授权）。



6.高级配置：

参数名称

参数说明



参数名称	参数说明
启动命令	配置服务的启动命令
环境变量	配置服务的环境变量
网络变量	配置调用端口（默认 8000）、本地监听端口（默认 8000）以及协议信息（http）。使用 80、8080、443、8443 敏感调用端口时，需先备案，再联系运营人员启用端口
滚动升级	配置服务最大不可用百分比，该参数仅影响服务升级。采用逐步用新版本实例替换旧版本实例的升级方式，请求会均衡分布到新旧版本实例上

高级配置

启动命令

运行命令

请输入运行命令

运行参数

单个填写

批量填写(换行进行分隔)

请输入参数(param)

+

添加

运行命令

python /var/ft_mnist/mnist_with_summaries.py

运行参数

-log_dir/train

-learning_rate=0.01

环境变量

变量键	变量值	操作
	暂无数据	添加

网络变量

调用端口	本地监听端口	协议	操作
			添加

添加80、8080、443、8443敏感端口时，需先备案，再由运营后端放敏感端口。

滚动升级

最大不可用比例

-

40

+

%

滚动更新中不可用实例数占预期实例数的百分比

7.完成以上配置，确认部署应用的资源配置费用，并勾选协议点击【部署应用】后，完成实例的创建。

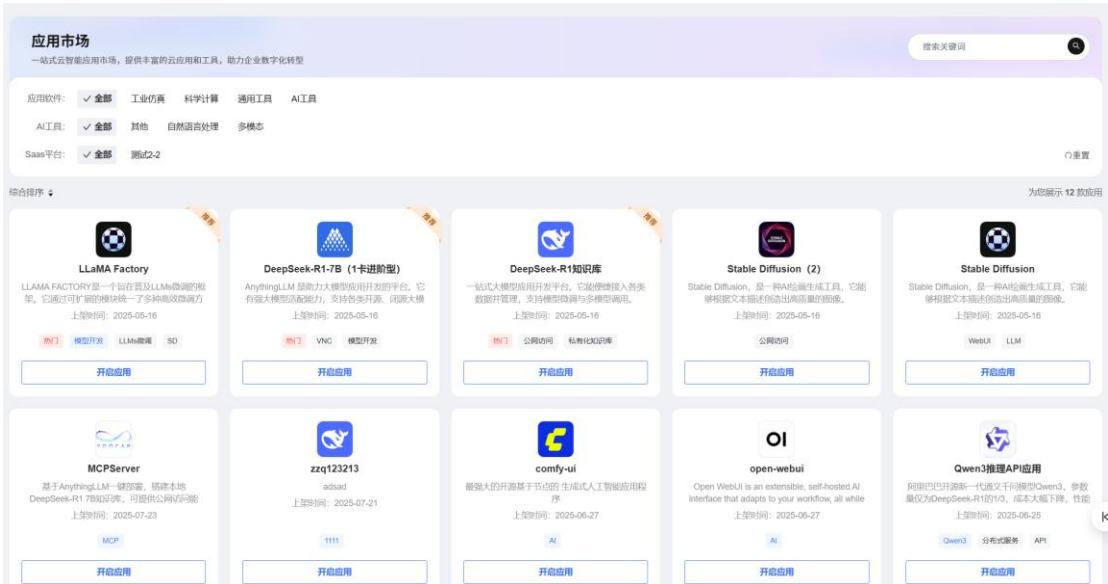
三、操作指导

1.应用市场

应用是指软件应用程序，例如开源软件工具、工业仿真软件、游戏软件等。应用托管平台平台提供全栈式应用托管方案，涵盖应用构建与部署、模型算法、SaaS 应用、业务调度、立体化监控等应用生命周期管理，基于异构算力统一调度引擎，实现对算力芯片，算力架构无差异化的应用部署，开箱即用。

操作步骤：

1.进入【应用托管控制台】， 点击【应用市场】。



2.选择应用点击对应的【开启应用】按钮，进入应用详情页。



3.选择访问策略，填入需要的参数，选择是否公网访问，勾选协议，点击【开启应用】按钮，进入订单支付页面。

4.应用购买成功后，可在【应用】【应用实例】，查看已购买的应用实例

应用实例管理						
+ 新建应用实例						
实例名称ID	状态	实例数	应用信息	创建时间	更新时间	操作
 [redacted]	创建失败	0 / 1	LLaMA Factory  app-df9ba0a4f93c390a0g	2025-07-30 09:57:31	2025-07-30 11:57:24	访问 启动 停止 更多
 [redacted]	运行中	1 / 1	时间 MCP 服务器  app-df4e6b5d3c3c330a0	2025-07-24 16:22:25	2025-07-30 10:15:40	访问 启动 停止 更多
 [redacted]	创建失败	0 / 1	LLaMA Factory  app-df9ba0a4f93c390a0g	2025-07-29 16:47:34	2025-07-29 17:17:50	访问 启动 停止 更多
 [redacted]	创建失败	0 / 1	enhance prompt server  app-df4e6b5d3c3c330a0g	2025-07-29 09:48:19	2025-07-29 10:23:43	访问 启动 停止 更多

5.待应用处于运行中状态时，可点击【访问】操作，点击访问地址进行应用使用。

应用实例管理						
+ 新建应用实例						
实例名称ID	状态	实例数	应用信息	创建时间	更新时间	操作
 [redacted]	创建失败	0 / 1	LLaMA Factory  app-df9ba0a4f93c390a0g	2025-07-30 09:57:31	2025-07-30 11:57:24	访问 启动 停止 更多
 [redacted]	运行中	1 / 1	时间 MCP 服务器  app-df4e6b5d3c3c330a0	2025-07-24 16:22:25	2025-07-30 10:15:40	访问 启动 停止 更多
 [redacted]	创建失败	0 / 1	LLaMA Factory  app-df9ba0a4f93c390a0g	2025-07-29 16:47:34	2025-07-29 17:17:50	访问 启动 停止 更多
 [redacted]	创建失败	0 / 1	enhance prompt server  app-df4e6b5d3c3c330a0g	2025-07-29 09:48:19	2025-07-29 10:23:43	访问 启动 停止 更多
mcp-cal ID: df4e6b5d3c3c330a0	已停止	0 / 1	mcp-计算图  app-df4e6b5d3c3c330a0g	2025-06-25 17:46:52	2025-07-23 17:58:01	访问 启动 停止 更多
oeshi ID: df4e6b5d3c3c330a0	已停止	0 / 1	数据示例  app-df4e6b5d3c3c330a0g	2025-07-09 15:17:35	2025-07-10 09:09:07	访问 启动 停止 更多

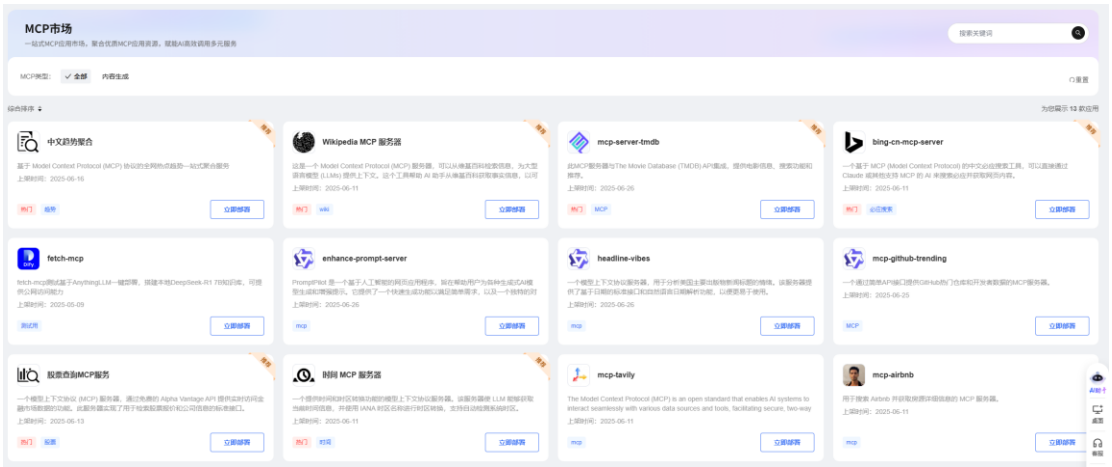
2.MCP 市场

MCP 是指 Model Calling Platform 相关的服务资源，涵盖内容生成、网页搜索、效率工具等多种类型，全面支撑企业各类智能应用场景。

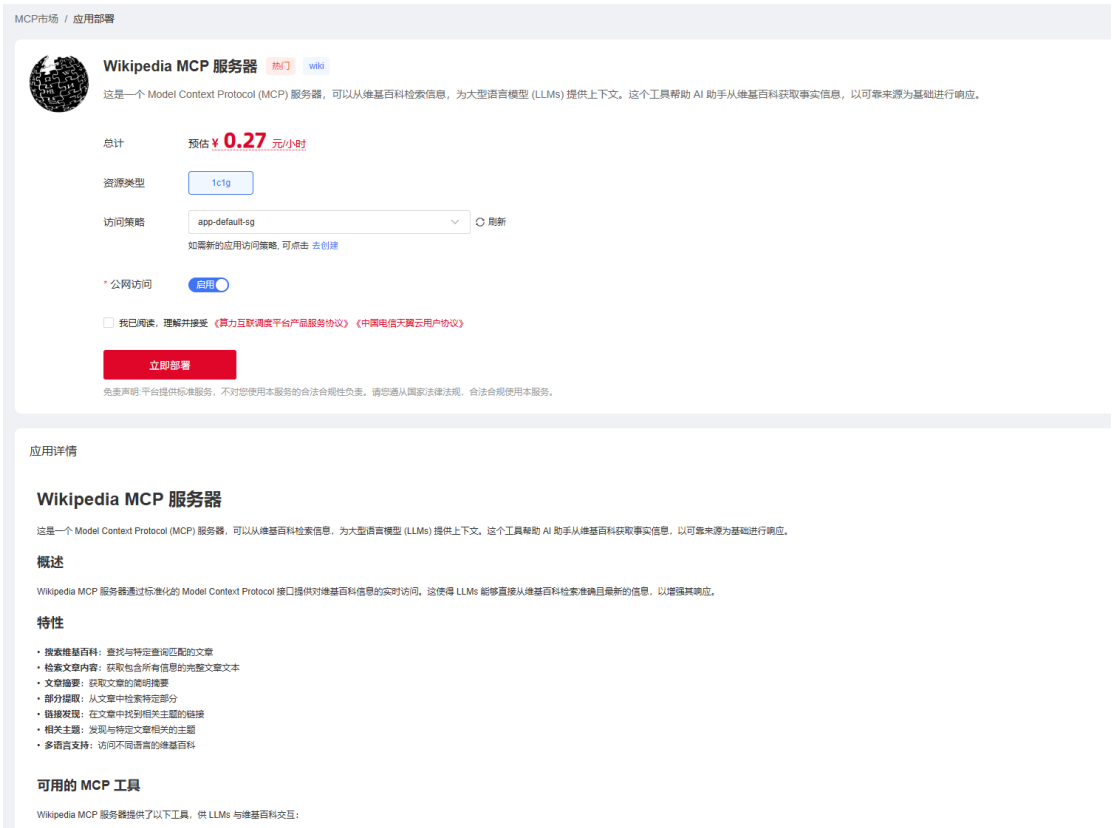
操作步骤：



1.进入天翼云应用托管平台，点击【MCP 市场】。



2.选择应用点击对应的【立即部署】按钮，进入应用详情页。



3.勾选协议，点击【立即订购】按钮，进入订单支付页面。

4.应用购买成功后，可在应用托管控制台-我的应用页面，查看已购买的应用。点击【开始部署】，可以再次部署该应用。点击【查看实例】按钮，可查看当前应用已部署的应用实例。

我的应用



mcp-mcp-1.0

[应用工具](#) [管理](#) [AI](#)

一个用于管理MCP应用的工具。您可以在这里查看和管理您的MCP应用。

立即部署

查看实例



mcp-计算

[MCP](#)

A Model Control Protocol server for calculating. This server enables LLMO to use calculator for precise numerical calculations.

立即部署

查看实例



mcp-github-trending

[MCP](#)

一个通过GitHub API接口提供GitHub热门仓库和开发者的MCP服务器。

立即部署

查看实例



mcp-server-mdb

[MCP](#)

此MCP服务器与The Movie Database (TMDB) API集成，提供电影信息、搜索功能和推荐。

立即部署

查看实例

共4条

10

1

 默认 1 页

应用实例管理

应用名称

mcp-github-trending

实例名称	状态	实例ID	应用信息	创建时间	更新时间	操作
mcp-github-trending-251112 ID: d1b2c725b5c0a0b0e0	已停止	8/1	mcp-github-trending 02 app-d1b2c725b5c0a0b0e0	2025-06-25 17:38:18	2025-06-26 15:32:15	设置 删除 更多

共1条

10

1

 默认 1 页

3.我的应用

操作步骤

1.登录应用托管控制台，在左侧导航栏单击【我的应用】，查看已部署的应用信息。

我的应用



mcp-mcp-1.0

[应用工具](#) [管理](#) [AI](#)

一个用于管理MCP应用的工具。您可以在这里查看和管理您的MCP应用。

立即部署

查看实例



mcp-计算

[MCP](#)

A Model Control Protocol server for calculating. This server enables LLMO to use calculator for precise numerical calculations.

立即部署

查看实例



mcp-github-trending

[MCP](#)

一个通过GitHub API接口提供GitHub热门仓库和开发者的MCP服务器。

立即部署

查看实例



mcp-server-mdb

[MCP](#)

此MCP服务器与The Movie Database (TMDB) API集成，提供电影信息、搜索功能和推荐。

立即部署

查看实例

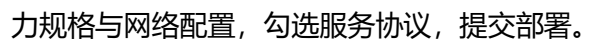
共4条

10

1

 默认 1 页

2.点击【立即部署】，可再次部署该应用，填写应用实例名称、应用基础配置信息，选择算



3. 点击【查看实例】，可查看该应用已部署的应用实例。

4.应用实例

4.1 创建公共资源应用实例

4.1.1 创建应用/MCP 市场应用实例

操作步骤

1.登录应用托管控制台，在左侧导航栏单击【应用实例】，单击【创建应用实例】按钮。

应用实例管理						实例名称	搜索关键字	🔍
实例名称ID	状态	实例类型	应用名称	创建时间	更新时间	操作		
 myapp-001 id: p1a8a3d43cm0a0ee0	 已停止	0.1	myapp-计算节点 app-01a3cm3d3qg3q3q	2025-07-25 17:40:52	2025-07-22 15:30:55	启动 刷新 删除		



2.应用配置：选择应用市场/MCP 市场的应用，输入实例名称。

应用配置

应用类型

应用市场

AI工具

SaaS平台

MCP

自定义应用

DeepSeek-R1-7B (1卡...)

open-webui

Qwen3推理API应用

dify-cpu

AnythingLLM 是助力大模型应用开发的平台。它有强大模型适配能力，支持各类开源、闭源大模型，可按需灵活调用，提供可视化构建界面，以低代码甚至无代码方式，让用户快速搭建聊天机器人、智能助手等应用。

更多应用详情

实例名称

请输入内容

长度为4-53字符，只能包含小写字母、数字及分隔符，分隔符不能连续出现，且需以小写字母开头，数字或小写字母结尾

3.资源配置：付费类型当前支持按量付费，选择算力规格，包括资源池、是否支持加速卡、地域、规格和实例数量。

资源配置

付费类型

按量付费

先使用后付费，按量计费

其他选项

网络策略

支持加速卡

地域

规格

实例数量

内网10

加速策略

--

华北-内网出-按量计费

CPU 1核 | 内存 1GB

1

+

已选择1个算力规格

4.网络配置：支持配置公网访问，开启后将为服务分配公网访问 ip。访问策略为服务的公网访问提供安全防护，可根据需要调整访问策略内容。一个服务只能绑定一个访问策略。

网络配置

公网访问

开启

访问策略

app-ef19a1f1c1g

删除

如需删除访问策略，请前往 访问策略

5.完成以上配置，并勾选协议点击【部署应用】后，完成实例的创建。

4.1.2 创建自定义应用实例

操作步骤

1.登录应用托管控制台，在左侧导航栏单击【应用实例】， 点击【创建应用实例】按钮。

应用实例管理

创建应用实例

实例名称

搜索实例

实例名称	状态	实例数	应用市场	创建时间	更新时间	操作
mcp-cpu ef19a1f1c1g	已停止	1/1	mcp-计算类 app-ef19a1f1c1g	2025-06-25 17:46:52	2025-07-22 15:36:58	详情 刷新 停止 删除



2.应用配置：选择自定义应用，输入实例名称。

应用配置

* 应用类型

应用软件

AI工具

Saas平台

MCP

自定义应用

* 实例名称

请输入内容

长度为4-53字符，只能包含小写字母、数字及分隔符，分隔符不能连续出现，且需以小写字母开头，数字或小写字母结尾

3.资源配置：资源类型选择【公共资源】，付费类型当前支持按量付费，选择算力规格，包括资源池、是否支持加速卡、地域、规格和实例数量。

资源配置

* 资源类型

公共资源

私有资源

* 付费类型

按量付费

先使用后付费，按量计费

* 算力规格

资源池

无加速卡

地域

规格

实例数量

资源池	无加速卡	地域	规格	实例数量
华东1A	无加速卡	华北1A	CPU 1核 / 内存 1GB	1

已选择1个算力规格

4.网络配置：支持配置公网访问，开启后将为服务分配公网访问 ip。访问策略为服务的公网访问提供安全防护，可根据需要调整访问策略内容。一个服务只能绑定一个访问策略。

网络配置

* 公网访问

开启

* 访问策略

默认策略

无策略

如策略未设置，则默认策略为默认策略，可点击 查看策略

5. 参数配置：配置服务的镜像，支持填写镜像地址（私有仓库需提供授权）。

参数配置

* 镜像类型

镜像地址

请输入镜像仓库地址/镜像名/版本

☐ 私有化镜像仓库授权

6.高级配置：

参数名称	参数说明
启动命令	配置服务的启动命令
环境变量	配置服务的环境变量
网络变量	配置调用端口（默认 8000）、本地监听端口（默认 8000）

参数名称	参数说明
	以及协议信息（http）。使用 80、8080、443、8443 敏感调用端口时，需先备案，再联系运营人员启用端口
滚动升级	配置服务最大不可用百分比，该参数仅影响服务升级。采用逐步用新版本实例替换旧版本实例的升级方式，请求会均衡分布到新旧版本实例上

高级配置

启动命令

运行命令

请输入运行命令

运行参数

☒ 单个填写 ☐ 批量填写(换行进行分隔)

请输入参数[param]

+

添加

运行命令

python /var/ft_mnist/mnist_with_summaries.py

运行参数

-log_dir=train

-learning_rate=0.01

环境变量

变量键	变量值	操作
	暂无数据	
	添加	

网络变量

调用端口	本地监听端口	协议	操作
		暂无数据	
		添加	

添加80、8080、443、8443敏感端口时，需先备案，再由运营后开放敏感端口。

* 滚动升级

最大不可用比例

-

40

+

%

滚动更新中不可用实例数占前期实例数的百分比

7.完成以上配置，确认部署应用的资源配置费用，并勾选协议点击【部署应用】后，完成实例的创建。

4.2 创建独享资源应用实例

4.2.1 创建自定义应用实例

操作步骤



1.登录应用托管控制台，在左侧导航栏单击【应用实例】，点击【创建应用实例】按钮。



2.应用配置：选择自定义应用，输入实例名称。



3.资源配置：资源类型选择【独享资源】，选择资源组进行应用部署。



4.参数配置：

参数名称	参数说明
Helm chart 地址	指定 Helm Chart 的存储位置。如果您的 Helm Chart 存放在私有仓库中，需要进行授权设置
命名空间	制定应用部署的命名空间
chart values	自定义 Helm Chart 的配置参数, 根据实际需求调整各个参数, 如镜像版本、副本数量、服务类型等
超时等待时间	定义 Helm 部署过程中拉取 Chart 的最大等待时间



参数配置

1 Helm chart地址

oci://registry.aliyuncs.com/acs-chart-repository/chart-name--tag

获取chart values

2 是否安装默认

☐

3 命名空间

默认命名空间

4 chart values

```
1
```

5 定时部署时间

0

300

0

Helm chart说明

简介

1 helm chart是kubernetes应用包的一个标准格式，包含所有必需的配置文件，如manifests、templates、charts等。

2 helm chart是kubernetes应用包的一个标准格式，包含所有必需的配置文件，如manifests、templates、charts等。

3 helm chart是kubernetes应用包的一个标准格式，包含所有必需的配置文件，如manifests、templates、charts等。

4 helm chart是kubernetes应用包的一个标准格式，包含所有必需的配置文件，如manifests、templates、charts等。

5 helm chart是kubernetes应用包的一个标准格式，包含所有必需的配置文件，如manifests、templates、charts等。

5.完成以上配置，并勾选协议点击【部署应用】后，完成实例的创建。

4.3 查看应用实例

4.3.1 查看实例概览

1.登录应用托管控制台，在左侧导航栏单击【应用实例】， 点击应用实例名称进入详情页。

应用实例管理

新增应用实例

实例名称

状态

实例ID

应用名称

创建时间

更新时间

操作

	Dify-V1.0.0-259765	运行中	1/1	Dify-V1.0.0-259765	2025-07-30 15:24:31	2025-07-30 15:28:23	访问 日志 更多
--	--------------------	-----	-----	--------------------	---------------------	---------------------	--

2.支持查看部署的应用信息、访问信息和资源配置信息。

Dify-V1.0.0-259765

运行中

查看日志 停止 部署 删除

部署方式

部署时间

创建时间

概览 部署 日志 编辑 变更记录

应用信息

应用名称

应用描述

访问信息

访问地址

公网地址

资源配置

资源规格

资源用量

实例数

4.3.2 查看配置

- 1.登录应用托管控制台，在左侧导航栏单击【应用实例】， 点击应用实例名称进入详情页。
- 2.点击【查看配置】， 可查看该应用实例的部署配置信息。



- 3.通过应用市场或 MCP 市场部署的应用， 展示部署的环境变量参数。



- 4.通过应用实例自定义部署的应用， 展示部署时完整的部署参数。通过公共资源部署的应用展示镜像地址、启动命令、环境变量、网络配置等参数。通过独享资源部署的应用展示 helm 地址、命名空间、chart values 等参数。

镜像地址harbor.ctyuncdn.cn/esx-dev/mcp/mcp-server-time

启动命令运行命令: -

启动参数: --pass-environment --port=8001 --sse-host 0.0.0.0 -- python -m mcp_server_time

环境变量	变量键	变量值
暂无数据		

网络配置	调用端口	本地监听端口	协议
	8000	8000	TCP

滚动升级最大不可用比例40%

关闭

Helm chart地址oci://harbor.ctyuncdn.cn/esx-dev/chart-test/http-echo:0.1.3

命名空间dadada

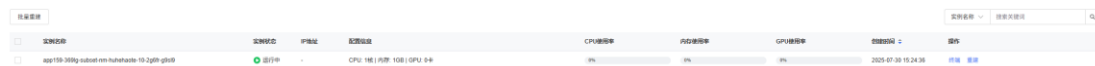
chart values	1 # Default values for lzy-test. 2 # This is a YAML-formatted file. 3 # Declare variables to be passed into your templates. 4 5 # This will set the replicaset count more information can be found 6 replicaCount: 1 7 8 # This sets the container image more information can be found 9 image: 10 repository: harbor.ctyuncdn.cn/esx/hashicorp/http-echo 11 tag: latest 12 13 args: 14 - "-text=Hello from Helm chart!" 15 - "-listen=:8080" 16
--------------	--

超时时间300 秒

关闭

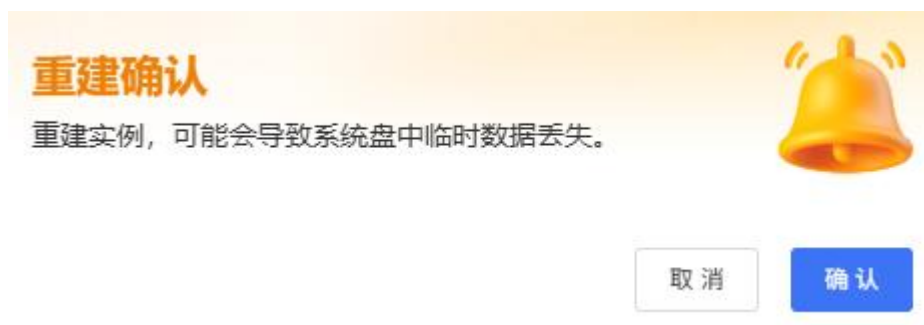
4.3.3 查看资源信息

- 1.登录应用托管控制台，在左侧导航栏单击【应用实例】， 点击应用实例名称进入详情页。
- 2.点击【资源】TAB，可查看该应用实例的部署实例资源信息。

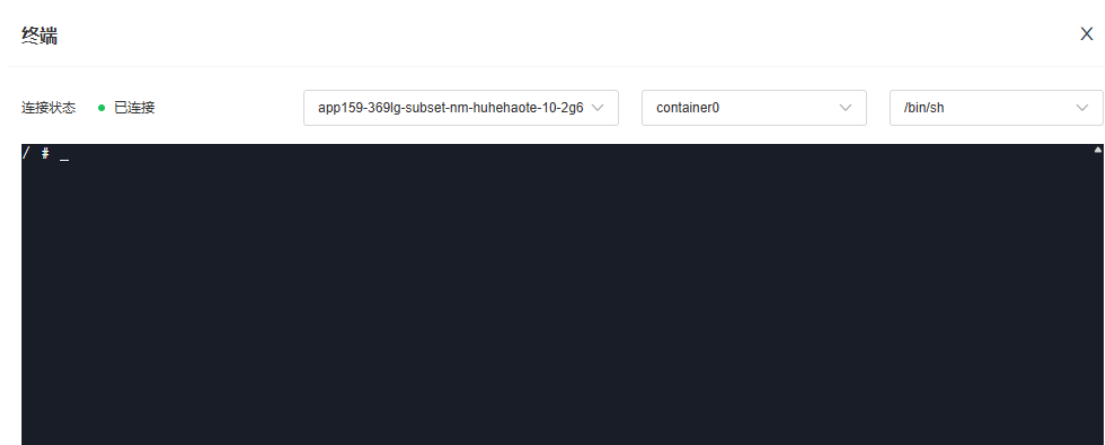


实例名称	实例状态	实例规格	配置信息	CPU使用率	内存使用率	GPU使用率	创建时间	操作
app159-369lg-subset-nm-huhehaote-10-2g6h-gt05	运行中	-	CPU: 1核 内存: 1GB GPU: 0卡	0%	0%	0%	2025-07-28 15:24:38	详情 重建

- 3.点击【重建】或【批量重建】，可手动对实例进行单独或批量的重启操作。

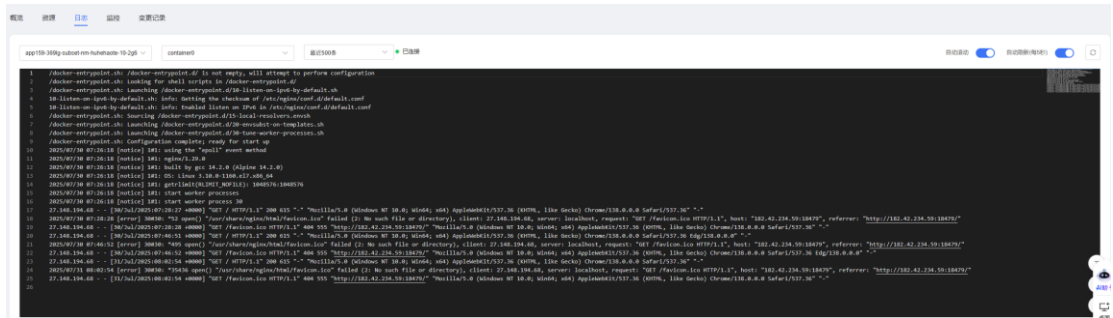
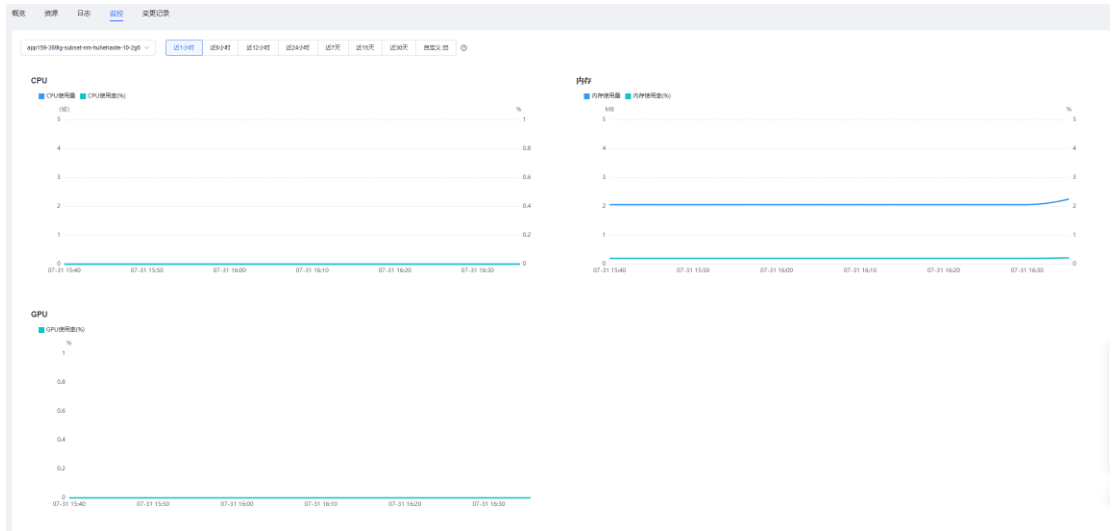


- 4.点击【终端】进入到资源的控制台页面，支持输入命令进行故障的排查。



4.3.4 查看实例监控

- 1.登录应用托管控制台，在左侧导航栏单击【应用实例】， 点击应用实例名称进入详情页。
- 2.点击【监控】TAB，可查看该应用实例的监控信息，支持切换实例，可查询实例的 CPU、内存和显卡数据。



概況詳情日程記錄變更記錄

選擇變更類型

選擇變更狀態

+

變更類型	變更狀態	開始時間	結束時間	執行時長	操作人
新增	成功完成	2025-07-28 19:06:58	2025-07-30 10:07:52	3 秒	
修改	成功完成	2025-07-28 19:19:23	2025-07-28 19:19:26	3 秒	
刪除 (新增)	成功完成	2025-07-24 16:22:25	2025-07-24 16:22:30	5 秒	

共 6 条 < 1 > 20 页 前往

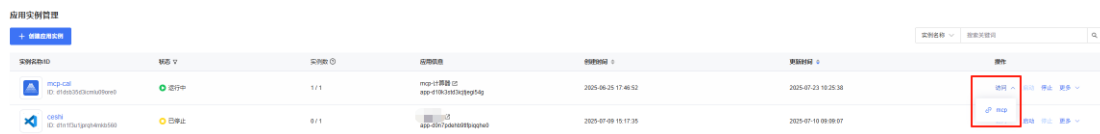
4.4 访问实例

前提条件

确认当前实例状态处于可调用状态，包括：运行中、更新中，且服务端口配置正确。

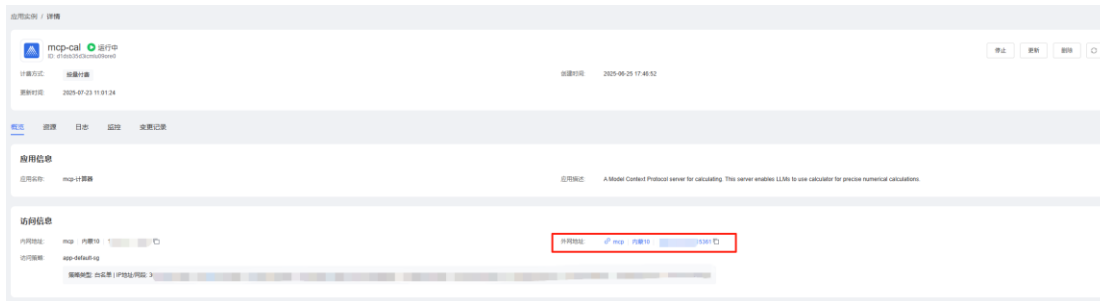
操作步骤

1.登录应用托管控制台，在左侧导航栏单击【应用实例】，在列表页，找到需要访问的实例，在操作栏点击【访问】按钮。



实例名称	状态	实例数	应用名称	创建时间	更新时间	操作
mcp-cal	运行中	1/1	mcp-cal	2025-06-25 17:46:52	2025-07-23 19:25:38	访问
mcp-cal	已停止	0/1	mcp-cal	2025-07-08 15:17:35	2025-07-18 09:09:07	访问

2.可进入实例详情页，可查看访问的域名或地址，复制链接到浏览器中可访问对应服务。



应用名称	应用状态
mcp-cal	运行中

访问信息
公网地址: mcp-cal
公网地址: mcp-cal

4.5 启动服务

前提条件

确认当前服务状态处于可启动状态，包括：已停止。

操作步骤

1.登录应用托管控制台，在左侧导航栏单击【应用实例】，在列表页，找到需要启动的实例，在操作栏点击【启动】按钮。



2.在进入启动服务操作弹窗后，点击【确认】完成启动操作。



4.6 停止服务

前提条件

确认当前实例状态处于可停止状态，包括：运行中。

操作步骤

1.登录应用托管控制台，在左侧导航栏单击【应用实例】，在列表页，找到需要停止的实例，在操作栏点击【停用】按钮。



2.在进入停止服务操作弹窗后，点击【确认】完成停止操作。



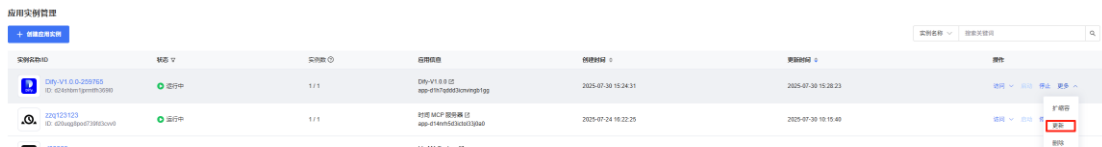
4.7 更新服务

前提条件

确认当前服务状态处于可升级状态，包括：创建中、 运行中 、运行失败、 已停止 。

操作步骤

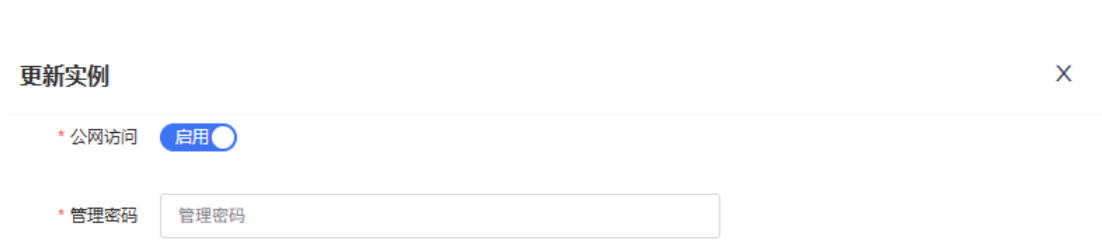
1.登录应用托管控制台，在左侧导航栏单击【应用实例】，在列表页，找到需要升级的服务，在操作栏点击【更新】按钮。



2.在进入升级操作页面后，具体参数区分以下几种情况。

2.1 应用市场/MCP 市场部署的应用

展示部署时定义的部署参数，根据不同应用展示对应的变量参数。



2.2 自定义应用（公共资源）

通过【应用实例】创建的自定义应用，资源类型为公共资源的应用。

参数名称	参数说明
公网访问	是否开启公网访问
镜像	填写镜像地址（私有仓库需提供授权）
启动命令	配置服务的启动命令
环境变量	配置服务的环境变量
网络变量	配置调用端口（默认 8000）、本地监听端口（默认 8000）以及协议信息（http）。使用 80、8080、443、8443 敏感调用端口时，需先备案，再联系运营人员启用端口
滚动升级	配置服务最大不可用百分比，该参数仅影响服务升级。采用逐步用新版本实例替换旧版本实例的升级方式，请求会均衡分布到新旧版本实例上

更新实例X

公网访问 ☒ 启用

镜像类型

镜像地址

harbor.ctyuncdn.cn/esx-dev/mcp/mcp-server-time

☐ 私有化镜像仓库授权

启动命令

运行命令

请输入运行命令

运行参数

☐ 单个填写 ☒ 批量填写(换行进行分隔)

-pass-environment

-port=8001

-sse-host

0.0.0.0

填写示例

运行命令

python /var/rt/mnist/mnist_with_summaries.py

运行参数

-log_dir=/train

-learning_rate=0.01

环境变量

变量键	变量值	操作
暂无数据		
添加		

网络变量

调用端口	本地监听端口	协议	操作
8000	8000	TCP	删除
添加			

添加80、8080、443、8443敏感端口时，需先备案，再由运营后端放通敏感端口。

滚动升级

最大不可用比例

-

40

+

%

滚动更新中不可用实例数占预期实例数的百分比

确定

取消

2.3 自定义应用（独享资源）

通过【应用实例】创建的自定义应用，资源类型为独享资源的应用。

参数名称	参数说明
Helm chart 地址	指定 Helm Chart 的存储位置。如果您的 Helm Chart 存放在私有仓库中，需要进行授权设置
命名空间	制定应用部署的命名空间

参数名称

参数说明

自定义 Helm Chart 的配置参数，根据实际需求调整各个参数，如 chart values

镜像版本、副本数量、服务类型等

超时等待时

定义 Helm 部署过程中拉取 Chart 的最大等待时间

间

更新实例

Helm chart地址

oci://harbor.ctyuncdn.cn/esx-dev/chart-test/http-echo:0.1.3

获取chart values

☐私有仓库授权

命名空间

dadada

chart values

```
1 # Default values for lzy-test.
2 # This is a YAML-formatted file.
3 # Declare variables to be passed into your templates.
4
5 # This will set the replicaset count more information can be found here: https://kubernetes.io/docs/concepts/workloads/controllers/replicaset/
6 replicaCount: 1
7
8 # This sets the container image more information can be found here: https://kubernetes.io/docs/concepts/containers/images/
9 image:
10   repository: harbor.ctyuncdn.cn/esx/hashicorp/http-echo
11   tag: latest
12
13 args:
14   - "-text=Hello from Helm chart!"
15   - "-listen=:8080"
16
```

超时等待时间

-

300

+

秒

Helm chart说明

规范

示例

1.templates模版文件中每一个资源都必须包含metadata, metadata.labels, metadata.namespace

2.helm默认会对资源的metadata.labels添加[include "\$名称.labels". | nindent 4]], 请保留或按第三点添加

3.metadata.labels需包含helm内置的`app.kubernetes.io/instance` 标签, 值为release名称(例:{{ .Release.Name }})

4.metadata.namespace的值需指定为release.namespace(例:{{ .Release.Namespace }})

5.工作负载类资源(deployment statefulset等), 需要在spec.template.metadata中添加`app.kubernetes.io/instance` 标签

3.完成修改后，点击【确认】即可启动更新流程。

4.8 扩缩容

前提条件

确认当前服务状态处于可扩缩容状态，包括：运行中、更新中。

操作步骤



1.登录应用托管控制台，在左侧导航栏单击【应用实例】，在列表页，找到需要扩缩容的实例，在操作栏点击【扩缩容】按钮。

应用实例管理						
实例名称ID	状态	实例数	应用描述	创建时间	更新时间	操作
mcp-cali id: ef76b02420mca00000	运行中	1 / 1	mcp-计算实例 app-ef76b02420mca00000	2025-07-23 17:46:52	2025-07-23 17:51:24	详情 日志 停止 更多
ccali id: ef76b02420mca00000	已停止	0 / 1	mcp-计算实例 app-ef76b02420mca00000	2025-07-23 15:17:35	2025-07-23 15:18:07	详情 日志 更多
mcp-github-rendering-201112 id: ef76b02420mca00000	已停止	0 / 1	mcp-github-rendering app-ef76b02420mca00000	2025-07-23 17:38:18	2025-07-23 17:38:18	详情 日志 更多

2.在进入扩缩容操作页面后，可以调整的字段如下：

2.1 实例数量：增加/减少实例数量，实例数量最小为 1。

2.2 资源池：增加/减少勾选资源池，至少需要有 1 个资源池。

扩缩容

调整当前服务实例数后，系统会按照当前的资源配置进行伸缩。弹性扩缩容会覆盖手动扩缩容的执行结果，如果想要手动管理实例数量，建议保持弹性扩缩容关闭。

资源池	支持加速卡	地域	规格	实例数量
☒ 内蒙10 资源充足	--	华北-内蒙古-呼和浩特	CPU: 1核 内存: 1GB	- 1 +

已选1个

取消

确定

3.完成修改后，点击【确认】即可启动扩缩容流程。在详情中的资源列表可查询扩缩容实例信息。

名称	实例ID	实例状态	IP地址	资源配置	CPU使用率	内存使用率	GPU使用率	创建时间	操作
app-123-12345678901234567890	app-123-12345678901234567890	运行中	-	CPU: 1核 内存: 1GB GPU: 0卡	10%	10%	10%	2025-07-23 10:25:45	详情 删除

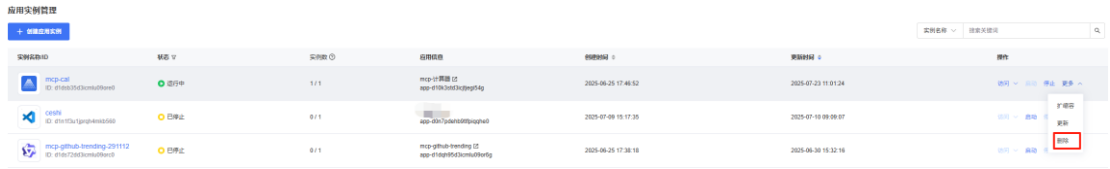
4.9 删除服务

前提条件

确认已完成服务的迁移，删除后该实例所有资源将会清空且不可恢复。

操作步骤

1.登录应用托管控制台，在左侧导航栏单击【应用实例】，在列表页，找到需要删除的实例，在操作栏点击【删除】按钮。



2.二次弹框删除服务，输入"确认删除"，点击【确定】按钮。



5.访问控制

5.1 创建访问控制策略

操作步骤

1.登录应用托管控制台，在左侧导航栏单击【推理】->【应用访问策略】，点击【创建访问策略】按钮。





2.填写基础信息：策略名称、选择策略类型、IP 地址/网段、描述。

创建访问控制策略

* 策略名称

演示策略

* 策略类型②

☒ 白名单

☐ 黑名单

* IP地址/网段

1.0.0.0

描述

✓ 每行一个IP地址或网段，以回车键换行

✓ 每个IP地址或网段都可以用“|”分隔

✓ 描述长度范围是0到255字符，不能包含<>

添加描述 如 “10.0.0.0/24|办公网出口”

0/255

确 认

取 消

注意：访问控制策略对关联的应用生效。

3.完成表单填写后，点击【确认】按钮，完成访问策略创建。

创建访问控制策略

×

* 策略名称

演示策略

* 策略类型②

☒ 白名单 ☐ 黑名单

* IP地址/网段

1.0.0.0

描述

✓ 每行一个IP地址或网段，以回车键换行

✓ 每个IP地址或网段都可以用“|”分隔添加描述 如 “10.0.0.0/24|办公网出口”

✓ 描述长度范围是0到255字符，不能包含<>

0/255

确认

取消

4.您可在访问控制列表查看您创建的访问控制策略，并对访问控制策略进行管理。

访问控制策略

策略名称	策略类型①	包含IP地址/网段	关联服务	描述	创建时间	操作
演示策略	白名单	1.0.0.0	所有服务	—	2025-05-15 11:33:37	修改 删除

共1条 10 < 1 >

5.2 访问控制策略管理

前提条件

已创建访问控制策略。

5.2.1 修改访问控制策略

1.登录应用托管控制台，在左侧导航栏单击【应用】->【应用访问策略】，进入访问控制策略列表页，点击【修改】。



策略名称	IP地址/网站	关联应用组 (运行中/停用)	描述	策略时间	操作
演示策略 白名单	1.1.1.1	00	--	2025-07-30 15:39:30	修改 删除

2.填写基础信息：策略名称、选择策略类型、IP 地址/网段、描述。

修改访问控制策略

×

* 策略名称

演示策略

* 策略类型?

☒ 白名单

☐ 黑名单

* IP地址/网段

1.0.0.0

描述

0/255

确认

取消

3.完成表单填写后，点击【确认】按钮，完成访问策略修改。

修改访问控制策略



* 策略名称 演示策略

* 策略类型 ☒ 白名单 ☐ 黑名单

* IP地址/网段 1.0.0.0

描述

0/255

确认

取消

5.2.2 删除访问控制策略

1. 登录应用托管控制台，在左侧导航栏单击【应用】->【应用访问策略】，进入访问控制策略列表页，单击【删除】。

应用访问策略					
+ 新建访问策略		请选择策略类型		请输入策略名称	
策略名称	IP地址/网段	关联应用数 (运行中/总数)	描述	更新时间	操作
演示策略 白名单	1.1.1	0/0	-	2025-07-30 15:39:30	修改 删除

2.点击弹窗【删除】按钮，即可删除访问控制策略

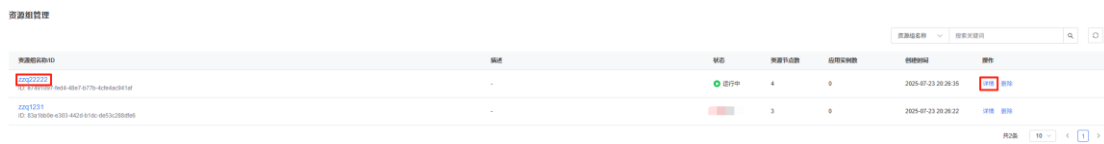


注意：系统默认策略无法删除。有关联应用的访问策略无法删除。

6.资源组管理

6.1 查看资源组

1.登录应用托管控制台，在左侧导航栏单击【资源组管理】，进入资源组列表，点击资源组名称或详情进入详情页。



2.支持查看集群的基础信息、资源配置消息和节点信息

zzq22222运行中

ID: a7491987-8a84-4ba7-b7d7

公网IP:

更新时间: 2025-07-23 20:26:35

应用实例数: 9

资源节点数: 4

创建时间: 2025-07-23 20:26:35

编辑资源组

资源配置

资源组资源概览

CPU (核数)	内存 (GiB)	GPU (型号)	GPU (卡数)	节点可用数/总数
112	1006.95	-	8	1 / 1
8	15.12	-	8	3 / 3

资源组资源总占用量/可用量/容量

CPU (核数)	内存 (GiB)	GPU (卡数)
- / 126 / 126	- / 1551.93 / 1551.93	- / 8 / 8

节点

节点名称 按条件筛选

节点名称

状态

CPU核数 (占用量/总容量)

内存 (占用量/总容量)

GPU卡数

GPU卡数 (占用量/总容量)

IP

创建时间

到期时间

node- 	正常	- / 112	- / 1006.95 (GiB)	-	- / 8		2025-07-16 17:08:59	-
master-4b1 	正常	- / 8	- / 15.93 (GiB)	-	- / 0	10.0.0.1	2025-06-11 14:22:39	-
master-ef8e6d1 	正常	- / 8	- / 15.93 (GiB)	-	- / 0		2025-06-11 14:22:38	-
master-ef8e6d1 	正常	- / 8	- / 15.93 (GiB)	-	- / 0	10.0.0.1	2025-06-11 14:22:14	-

6.2 管理资源组

当前尚未提供自助资源组新建能力，如果您想基于资源组部署独享资源应用，可通过联系客服经理或工单申请。

7.项目空间

7.1 创建项目空间

操作步骤

1.登录应用托管控制台，在左侧导航栏单击【资产】->【项目空间】，点击【创建项目空间】按钮。

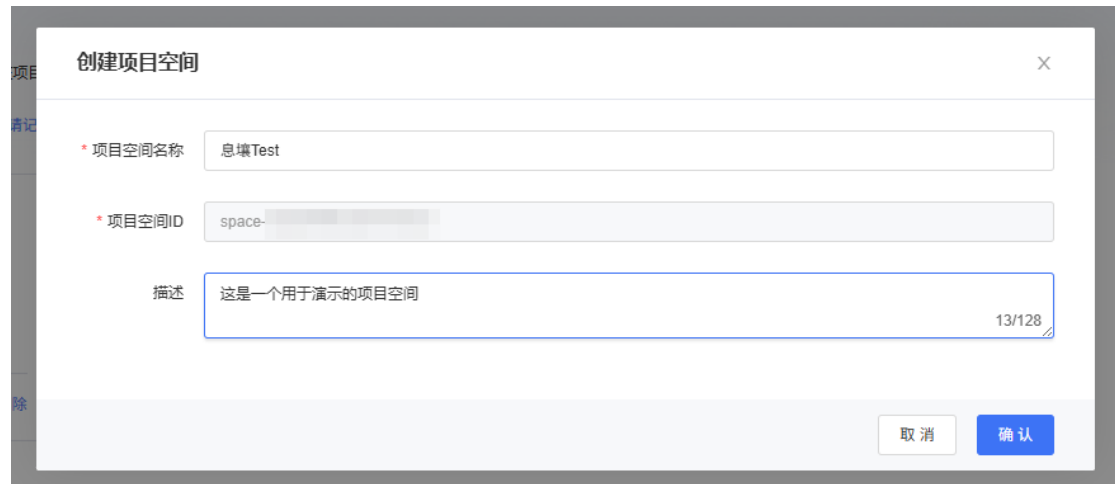


说明

租户首次使用应用托管控制台，平台会自动为用户创建默认空间并初始化。

若项目空间初始化失败导致服务不可用，请刷新页面或重新进入应用托管控制台。

2.填写基础信息：项目空间名称、描述。

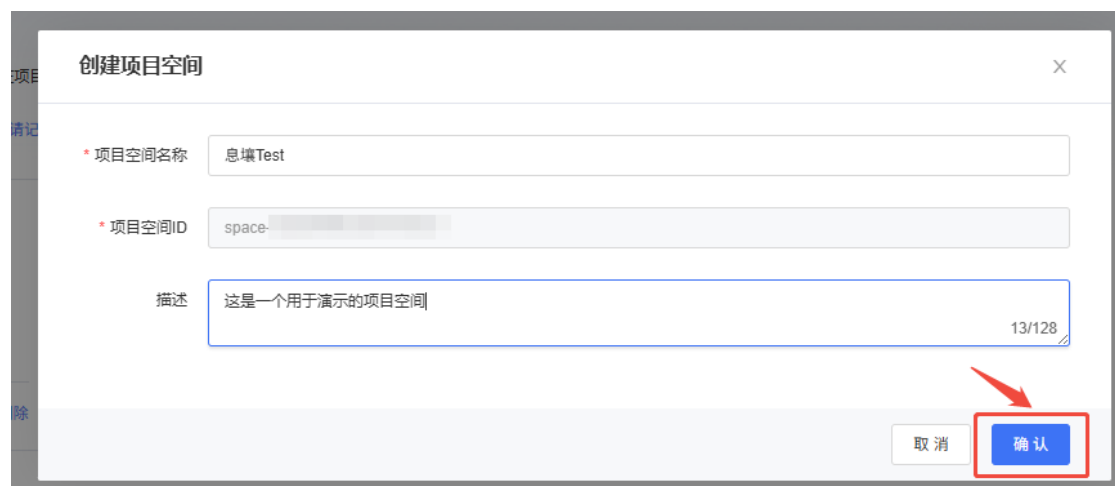
该截图显示了一个名为“创建项目空间”的对话框。对话框顶部有一个关闭按钮（X）。表单包含三个输入项：1. “* 项目空间名称”，输入框内已填写“息壤Test”。2. “* 项目空间ID”，输入框内显示为“space-”后跟一串灰色模糊字符。3. “描述”，输入框内已填写“这是一个用于演示的项目空间”，右侧显示字符限制为“13/128”。对话框底部右侧有两个按钮：“取消”和“确认”。

说明

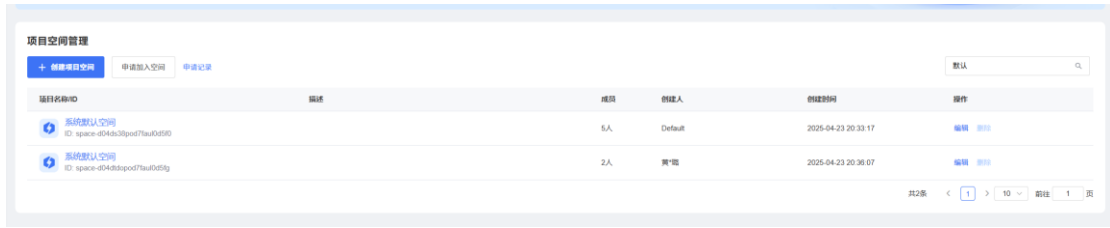
项目空间名称具有租户下唯一性，同一个租户下不可创建同名项目空间。

项目空间 ID 由平台自动生成，项目空间 ID 全局唯一，不可修改。

3.完成表单填写后，点击【确认】按钮，完成项目空间创建。

该截图与上方一致，展示了“创建项目空间”对话框。与上方不同的是，此图中“确认”按钮被一个红色的方框圈出，并且有一个红色的箭头从右侧指向该按钮，以指示用户点击确认以完成创建。

4.您可在项目空间栏目查看您创建或已加入的项目空间，并对项目空间进行管理。



项目名称/ID	描述	成员	创建人	创建时间	操作
系统默认空间 ID: space-d048c38pod7bau005d0		5人	Default	2025-04-23 20:33:17	编辑 删除
系统默认空间 ID: space-d048c38pod7bau005d0		2人	黄**	2025-04-23 20:36:07	编辑 删除

7.2 项目空间管理

前提条件

已开通项目空间并完成项目空间初始化。

7.2.1 查询项目空间列表

1.登录应用托管控制台，在左侧导航栏单击【资产】->【项目空间】，查询项目空间列表。

支持输入项目空间名称关键字进行模糊搜索。



项目名称/ID	描述	成员	创建人	创建时间	操作
系统默认空间 ID: space-d048c38pod7bau005d0		5人	Default	2025-04-23 20:33:17	编辑 删除
系统默认空间 ID: space-d048c38pod7bau005d0		2人	黄**	2025-04-23 20:36:07	编辑 删除

7.2.2 申请加入空间

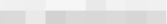
1.点击【项目空间】进入项目空间列表页，点击【申请加入空间】，填写项目空间 ID 和申请备注，点击【确认】即可发送申请。



申请加入空间

X

* 项目空间ID

space-cv! 

申请备注

计算机学院6班XXX

10/128

取消

确认

项目空间 ID 可联系对应的项目空间创建人或管理员获取。

若项目空间未开启【允许自主申请加入】或项目空间 ID 不存在会导致申请失败，申请前请确认项目空间 ID 是否正确及已开启【允许自动申请加入】

2. 点击【申请记录】可查看已提交的项目空间申请及审批状态，支持撤回申请。

创建及管理项目空间

01 创建项目空间

02 管理项目空间

项目空间管理

创建项目空间

申请加入空间

申请记录

项目名称ID	描述
系统默认空间	space-d8a7gjh89tflpqh8g
系统默认空间	space-004318b5a022a8-c79b8

申请记录

空间名称	空间ID	申请时间	审批结果	审批意见	审批人	审批时间	操作
	space-d8a7gjh89tflpqh8g	2025-06-06 15:23:36	已通过	同意审批	*****	2025-06-06 15:30:25	撤回
	space-004318b5a022a8-c79b8	2025-05-20 12:09:17	已通过	同意审批		2025-05-20 12:09:56	撤回

共2条 < 1 > 10 前往 1 页

7.2.3 进入/切换项目空间

1. 在项目空间列表，找到想要使用的项目空间，点击【空间名称】按钮，即可切换到指定工作空间，系统会自动跳转【概览】页面。

创建及管理项目空间

01 创建项目空间

02 管理项目空间

03 监控资源使用情况

项目空间管理

创建项目空间

申请加入空间

申请记录

项目名称ID	描述	成员	创建人	创建时间	操作
系统默认空间	space-d8a7gjh89tflpqh8g	5人	Default	2025-04-23 20:33:17	编辑 撤回
系统默认空间	space-d8a7gjh89tflpqh8g	2人	黄**	2025-04-23 20:36:07	编辑 撤回

2. 可以在菜单切换项目空间，也可以在【概览】页顶部查看当前的项目空间信息及点击下拉列表进行切换。项目空间切换后系统会自动更新【概览】页面数据。

应用托管

系统默认空间

项目管理员

02377a12

a2c979c8

670b8da5

系统默认空间(@Default创建)

ID: space-d04ds38pod7faul0d5f0

374e097



项目



项目空间: 系统默认空间(@Default创建) 空间ID: space-d04ds38pod7faul0d5f0

Hi, 欢迎使用应用托管

为您提供更丰富、更优质的应用产品, 助力您的业务轻松部署与应用便捷管理

找应用

丰富的云应用和工具, 助力企业数字化转型

- LLaMA Factory
- DeepSeek-R1-7B (1卡进阶型)
- DeepSeek-R1知识库

查看全部 >

找MCP

提供优质的MCP应用资源, 赋能AI多元服务

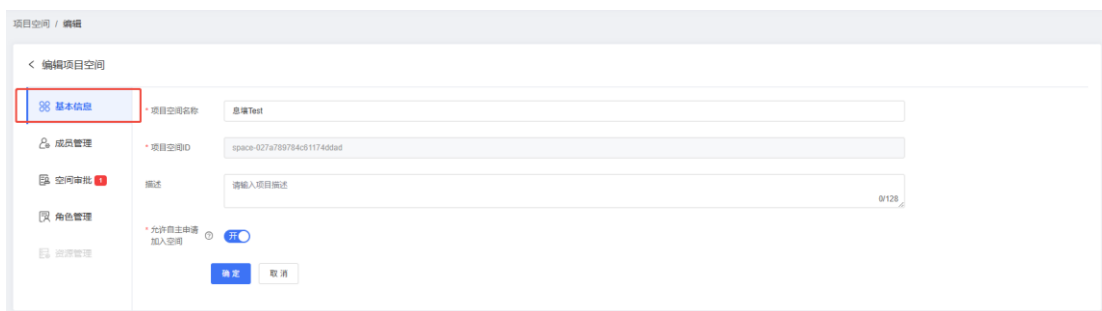
- 中文趋势聚合
- Wikipedia MCP 服务器
- mcp-server-tmdb

查看全部 >

7.2.4 编辑项目空间基本信息

1.在项目空间列表，点击项目空间的【编辑】按钮，进入编辑页面。

2.在【基本信息】栏目，可以修改项目空间名称、项目描述等信息，点击【确定】按钮保存修改。若您希望其他用户可以自主申请加入您的项目空间，可以打开允许自主申请加入项目空间按钮。空间加入申请可在【空间审批】中查看。



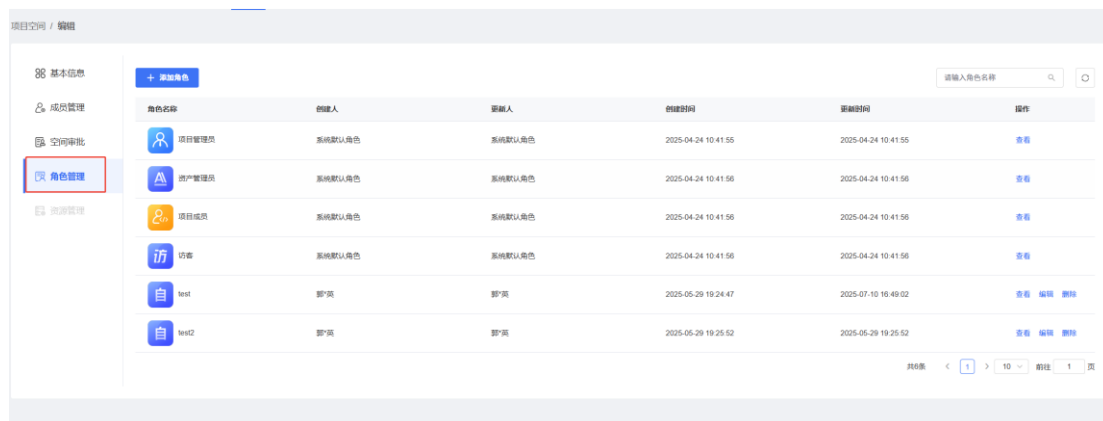
说明

同一个租户下的项目空间不允许重名。

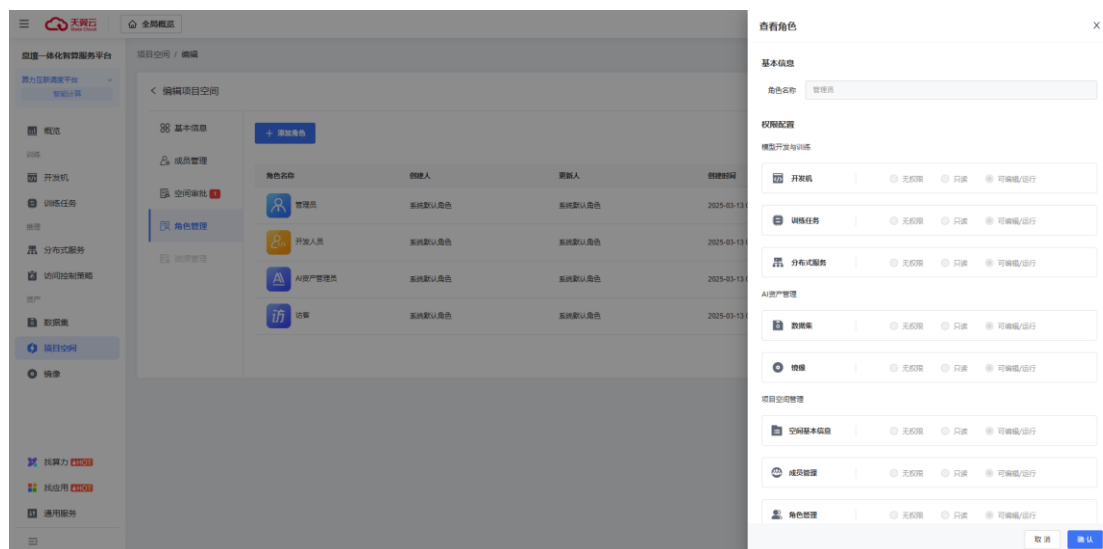
项目空间 ID 由平台自动生成，不可修改。

7.2.5 项目空间角色管理

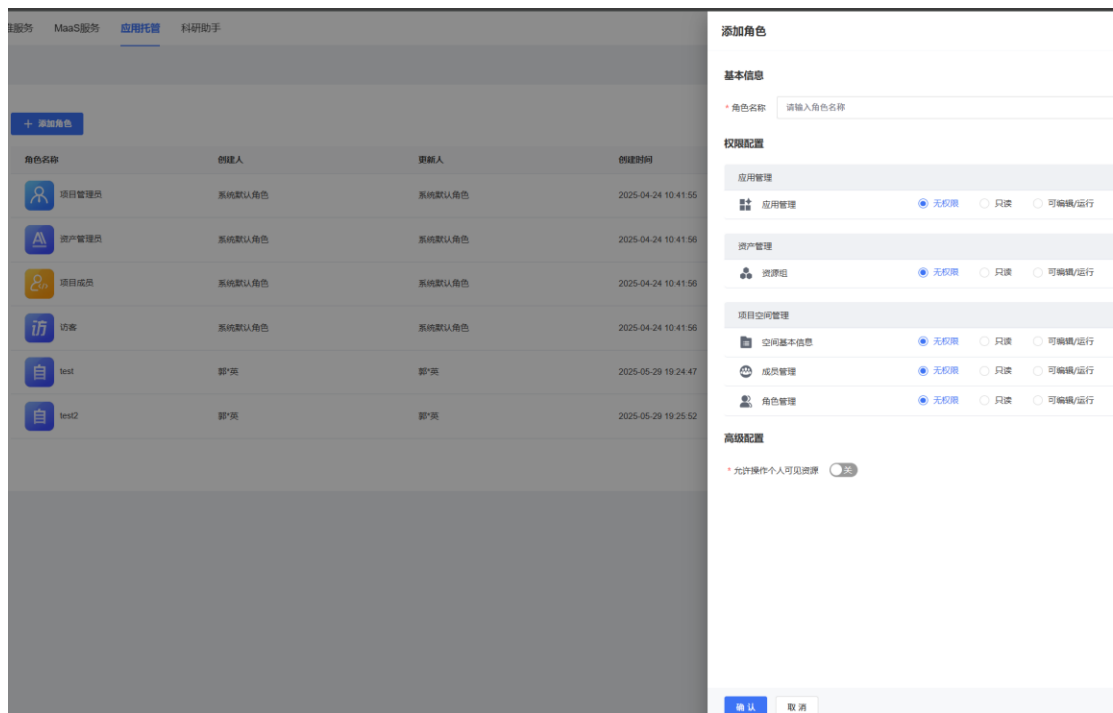
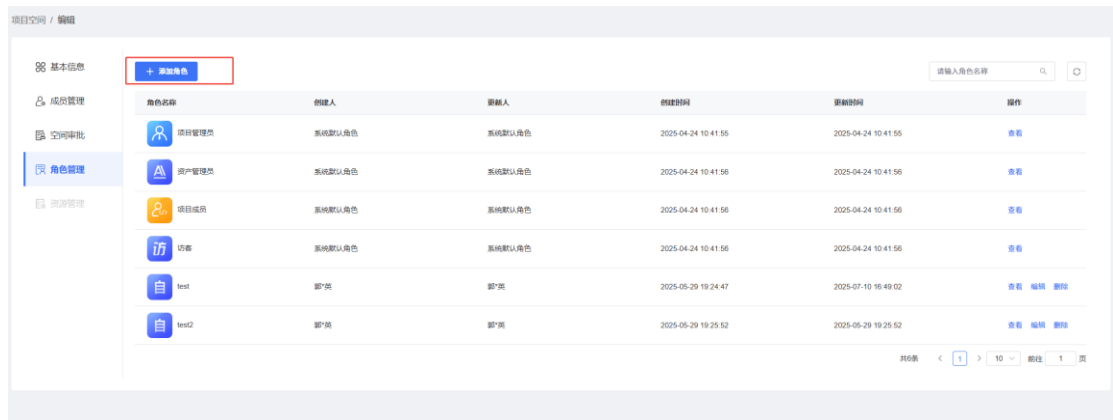
1.在项目空间编辑页面，点击【角色管理】进入项目空间角色管理页面，您可在【角色管理】添加和管理项目空间角色。



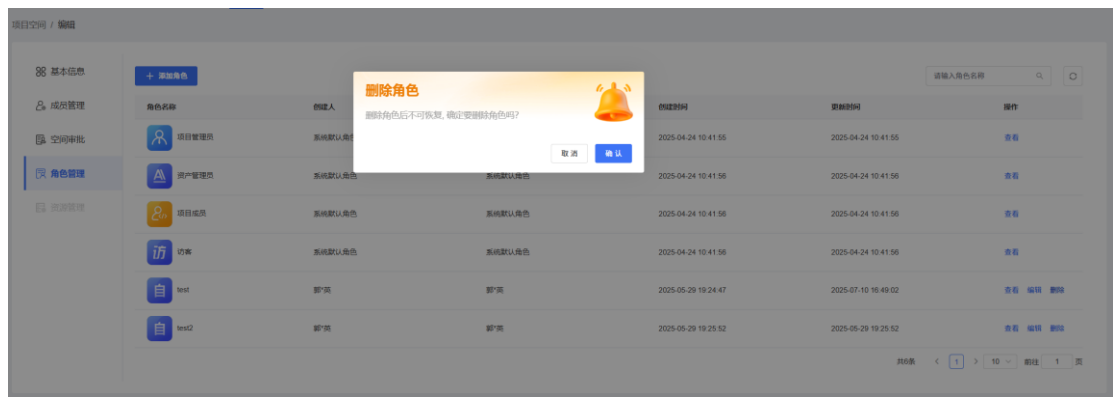
2.系统默认角色：系统会自动为项目空间创建【管理员】【项目成员】【资产管理员】【访客】四种默认角色。可以点击【查看】系统默认角色基本信息和权限信息，系统默认角色不支持编辑和删除。



3.添加角色：支持用户添加自定义角色，点击【添加角色】按钮，编辑角色名称和配置对应权限，点击【确认】即可完成添加自定义角色。自定义角色支持二次编辑和删除。

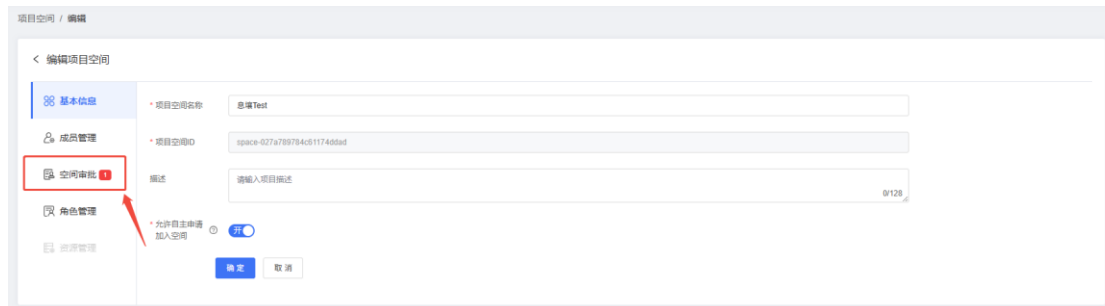


4.删除角色：支持用户删除项目空间自定义角色，点击【删除】按钮并进行二次确认即可删除角色。当角色下存在已关联项目空间成员时不可删除。



7.2.6 空间审批

1.在项目空间编辑页面，点击【空间审批】进入项目审批管理页面，您可在【空间审批】查看其他用户的空间加入申请并进行处理。



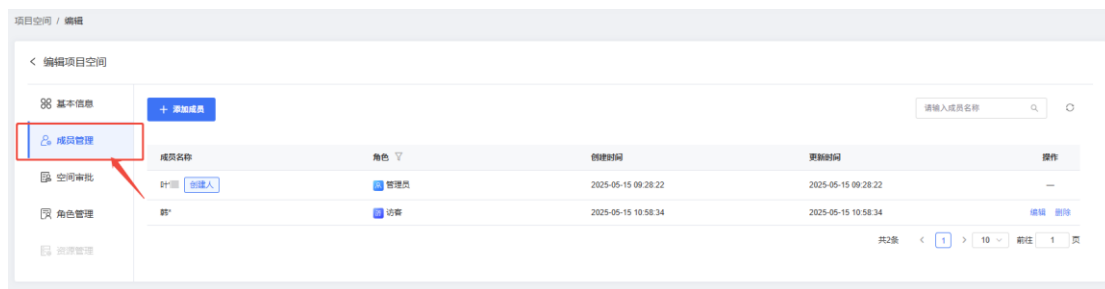
2.在项目空间审批列表，找到需要处理的申请消息，为用户选择项目空间角色，并点击【同意】即可允许该用户加入该项目空间。点击【拒绝】则会驳回该用户的申请。

3.可在【已审批】列表查看已处理的空间申请。



7.2.7 项目空间成员管理

1.在项目空间编辑页面，点击【成员管理】进入空间成员管理页面，您可在【成员管理】查看已加入空间的成员用户并进行管理。



2.添加成员：点击【添加成员】按钮，输入用户邮箱并选择角色，点击【确认】即可邀请指定用户加入空间。



注意

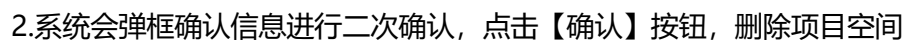
项目空间内的成员可共享资源 and 数据, 请勿随意添加成员并为所有成员设置合理的项目空间权限，避免造成资产损失。

3.编辑成员：点击【编辑】按钮，可修改成员角色。





1.在项目空间列表页，点击项目空间卡片下的【删除】按钮。





- 项目空间删除后不可恢复，请谨慎操作。
- 系统默认空间不允许删除。