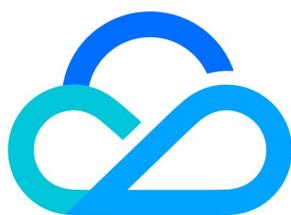


高性能应用服务

快速入门



腾讯云

【 版权声明 】

©2013–2024 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或95716。

文档目录

快速入门

Stable Diffusion 常见操作指南

通过高性能应用服务一键创建应用

快速入门

Stable Diffusion 常见操作指南

最近更新时间：2024-01-19 17:13:41

本文将介绍常用 Stable Diffusion 操作，包括重启 WebUI、导入外部模型、下载插件等。

重启 WebUI

若您在使用过程中遇到 WebUI 界面无法正常打开，可进入 Terminal，通过如下指令重新启动 WebUI 服务。

```
cd stable-diffusion-webui
(python3 -u launch.py --skip-prepare-environment --xformers --listen --enable-insecure-extension-
access --port=6889 >> /var/log/sd_service.log 2>&1 &) && tail -f /var/log/sd_service.log
```

查看 WebUI 运行日志

若您希望查看 WebUI 运行状态，包括模型加载、图片生成、插件下载进度等内容，可进入 Terminal，通过如下指令查看 WebUI 运行日志。

```
cd stable-diffusion-webui
tail -f /var/log/sd_service.log
```

导入外部模型

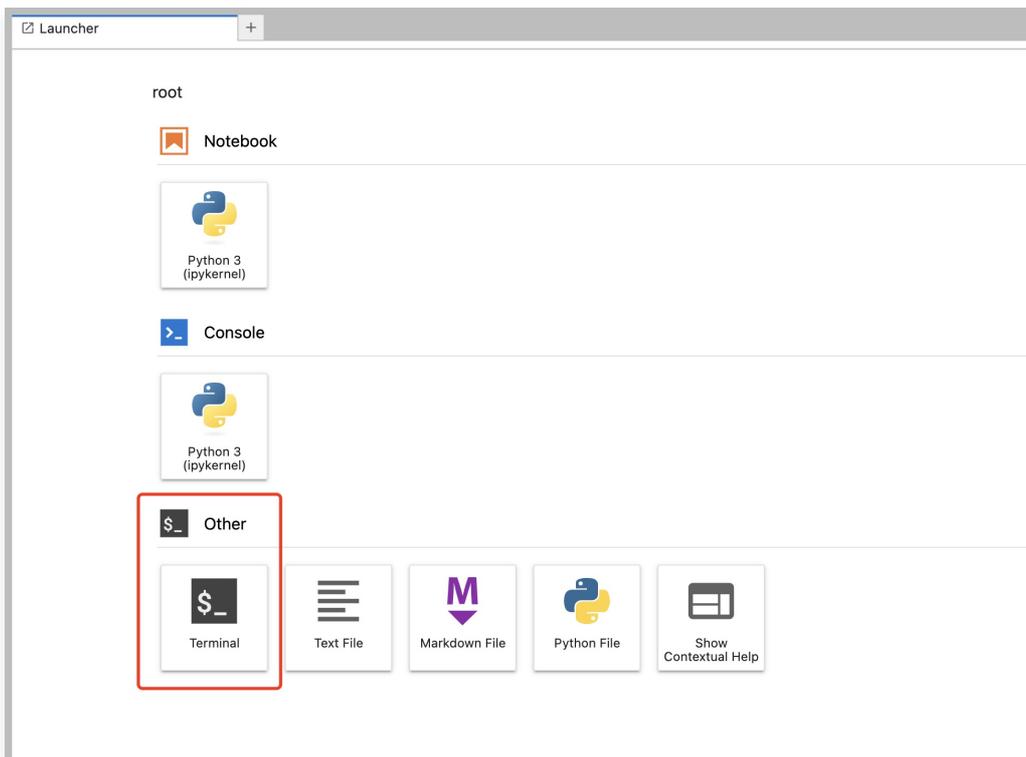
目前支持通过3种路径导入模型：

方式一：直接下载（推荐境外地域实例使用）

1. 登录 [高性能应用服务 HAI](#) 控制台。
2. 选择需要连接的算力，然后单击**算力连接**。在下拉菜单中，单击 **JupyterLab**。



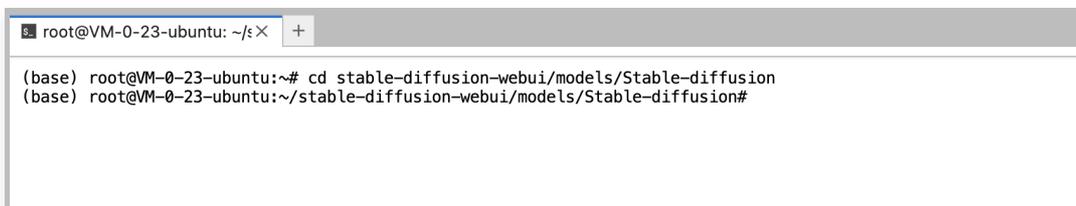
3. 进入 JupyterLab 后，选择 **Other > Terminal**，进入终端界面。



4. 进入后默认处于根目录下，输入如下指令进入 `stable diffusion` 的模型文件夹。

```
cd stable-diffusion-webui/models/Stable-diffusion
```

截图如下：



5. 获取预期使用的模型文件下载地址，在 `stable-diffusion-webui/models/Stable-diffusion` 目录下，将文件使用 `wget` 指令下载到目录中。

```
wget https://xyz.com/api/download/models/xxxx
```

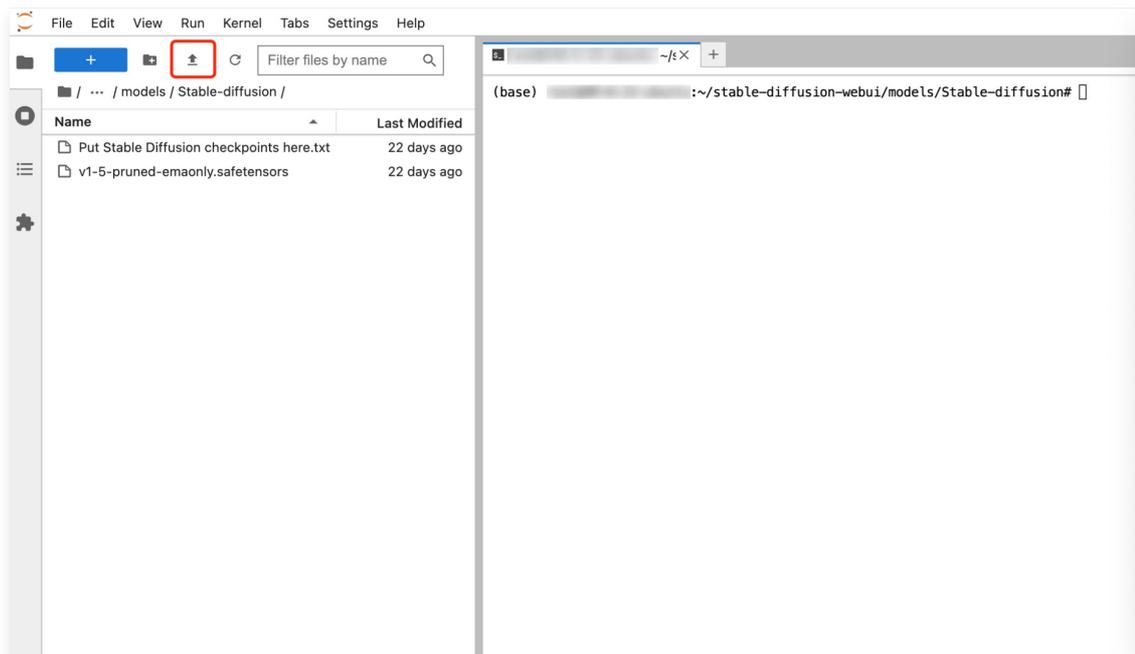
6. 刷新 WebUI，即可使用新模型。

方式二：通过 JupyterLab 上传（适合1GB以内小文件）

1. 将预期使用的模型提前下载到电脑本地。
2. 选择 JupyterLab 进行算力连接，在左侧文件目录处，依次双击进入如下文件夹：

```
root/stable-diffusion-webui/models/Stable-diffusion
```

3. 单击左侧文件目录处的文件上传，选择已下载的模型文件，即可将模型上传至高性能应用服务实例。

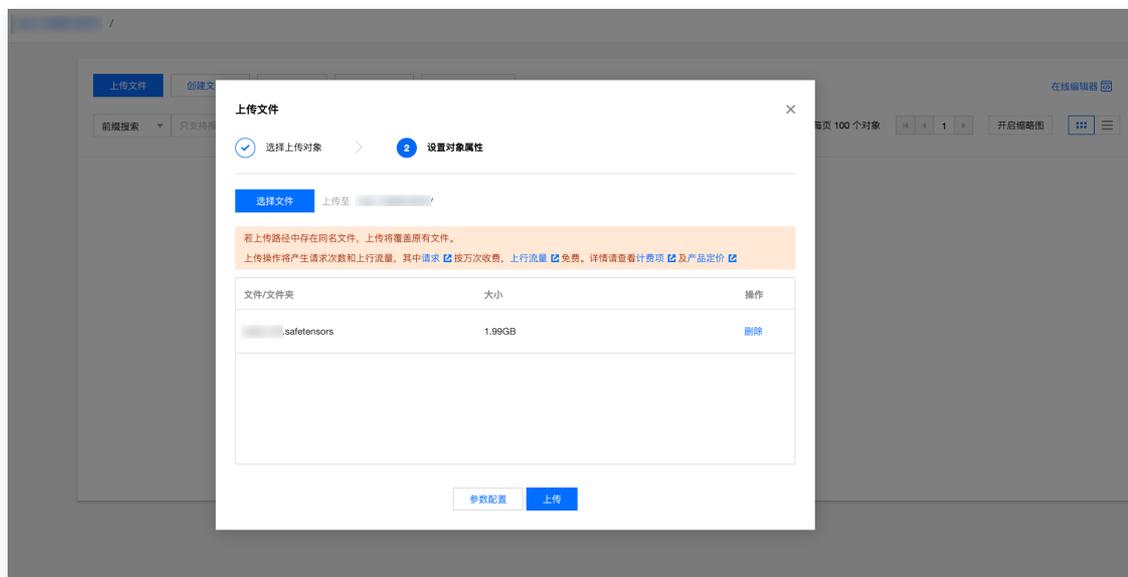


方式三：通过 COS 上传（适合大文件）

1. 将预期使用的模型提前下载到电脑本地。
2. 进入腾讯云 [对象存储控制台](#)（cos），创建新存储桶并将模型文件上传至该存储桶。

⚠ 注意：

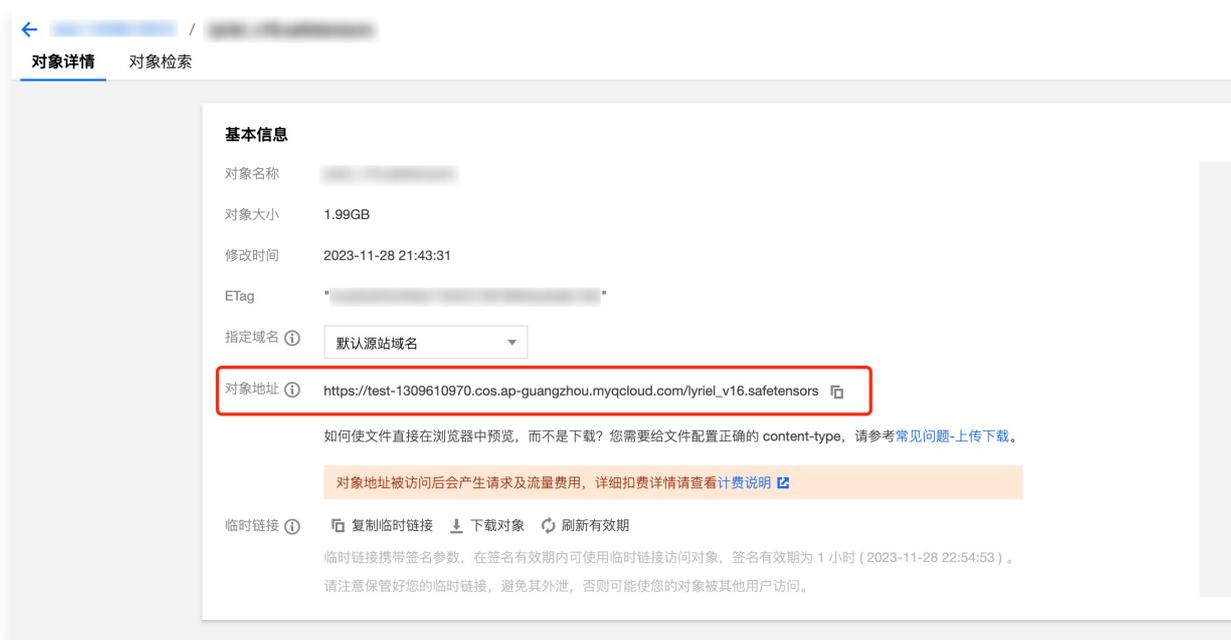
创建 COS 存储桶时，需注意与应用实例相同地域。否则会在拉取文件时产生额外流量费用。



3. 选择 JupyterLab 进行算力连接，进入Terminal后，进入后默认处于根目录下，输入如下指令进入 stable diffusion 的模型文件夹。

```
cd stable-diffusion-webui/models/Stable-diffusion
```

4. 获取在存储桶中的模型文件下载地址，并在该目录下，将文件 wget 到高性能应用服务实例中。



```
wget https://test-xxx.cos.ap-guangzhou.myqcloud.com/lyriel_v16.safetensors
```

5. 刷新 WebUI，即可使用新模型。

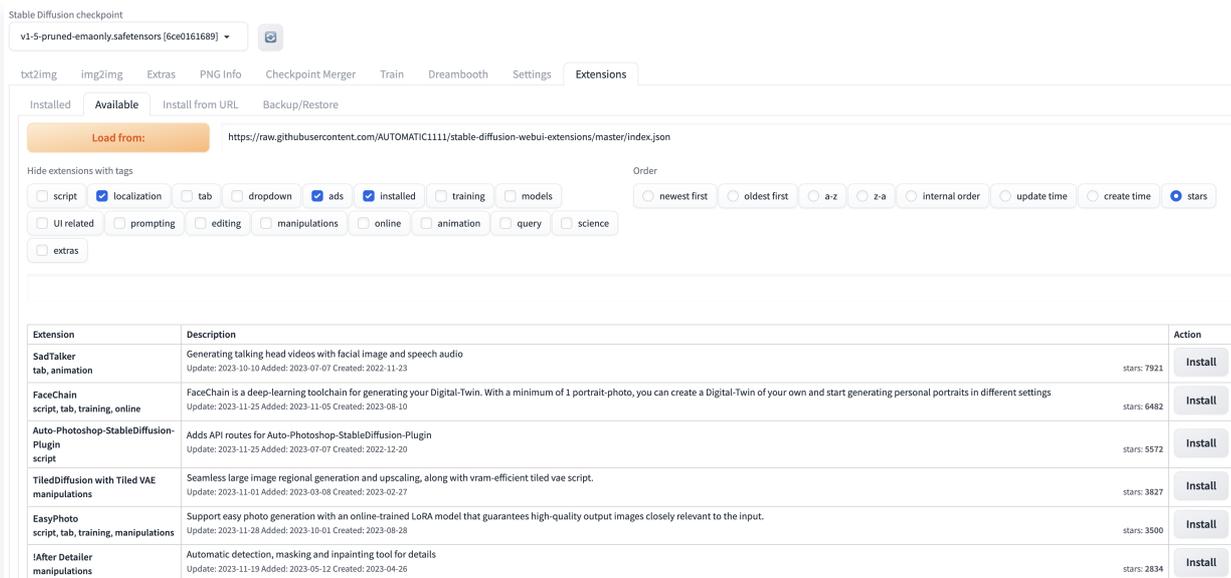
新增插件

直接安装（推荐境外地域实例使用）

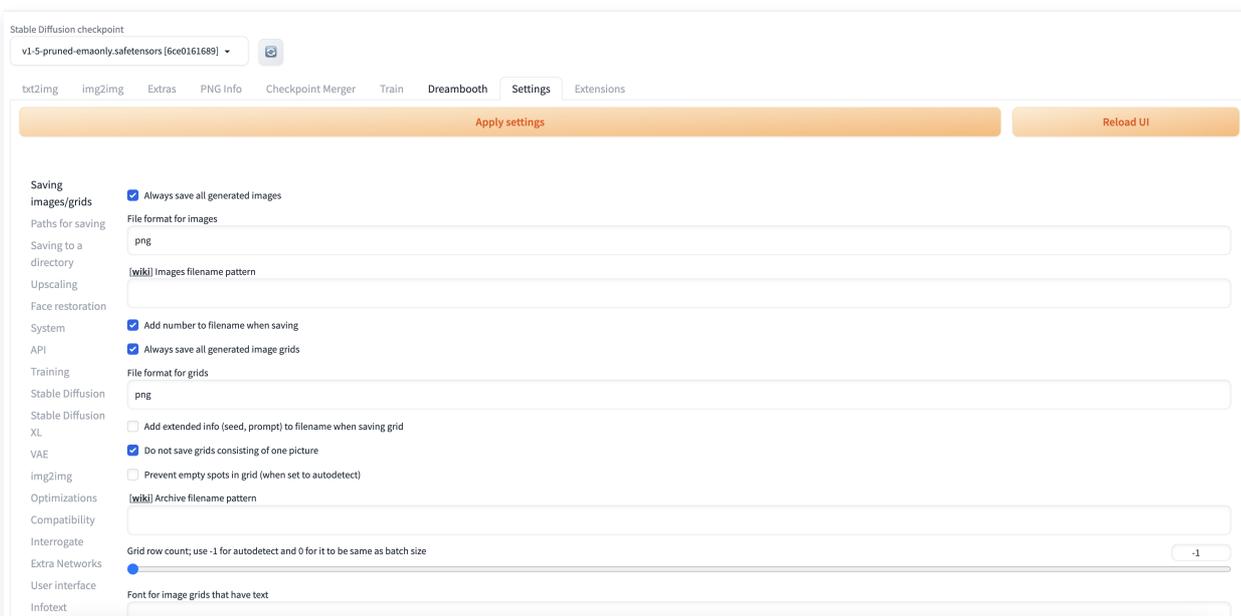
1. 登录 [高性能应用服务 HAI 控制台](#)。
2. 选择对应的算力，单击**算力连接**。在下拉菜单中，单击 **Gradio WebUI**。



3. 进入 WebUI，单击进入 **extensions**，单击 **Load from** 后，选择预期使用的插件，单击 **Install**。



4. 安装完成后，单击进入 **Settings**，单击 **Reload UI** 后，插件即可生效。

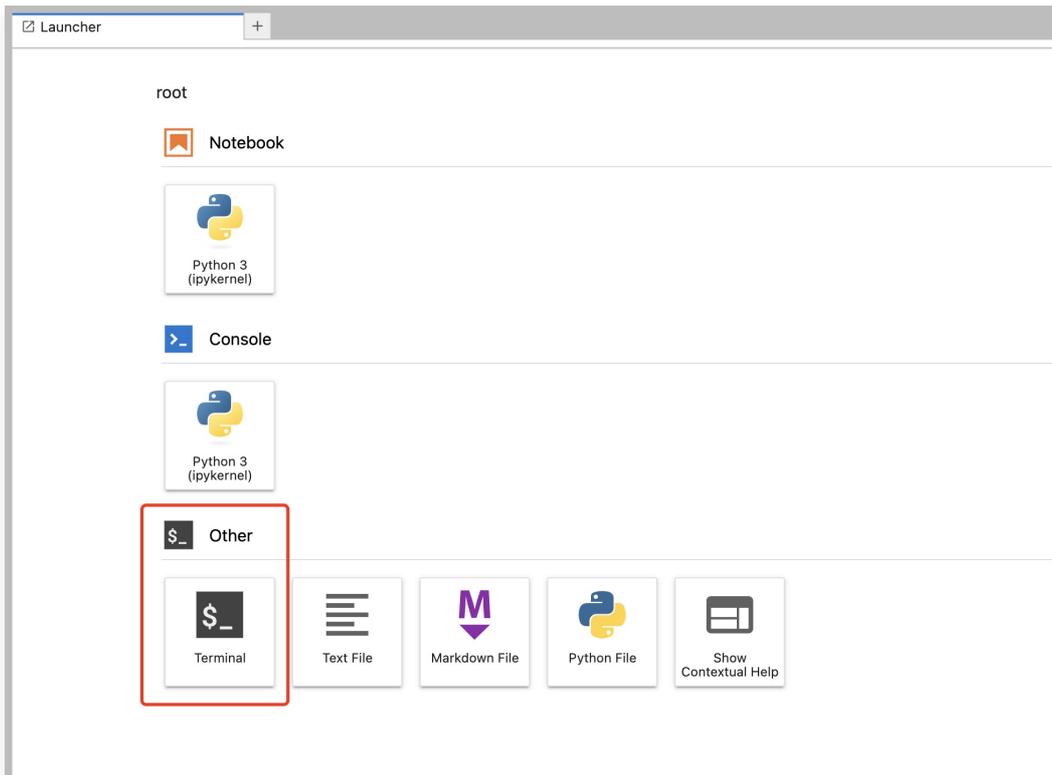


进入插件目录安装

1. 登录 [高性能应用服务 HAI 控制台](#)。
2. 选择需要连接的算力，然后单击**连接算力**。在下拉菜单中，单击 **JupyterLab**。



3. 选择 **JupyterLab** 进行算力连接，进入 Terminal。



4. 进入后默认处于根目录下，输入如下指令进入 `extensions` 文件夹。

```
cd stable-diffusion-webui/extensionts
```

5. 获取预期使用的插件下载地址，在 `stable-diffusion-webui/extensionts` 目录下，将文件使用 `git clone` 指令下载到目录中。

```
git clone https://github.com/OpenTalker/SadTalker
```

6. 刷新 WebUI，即可使用新插件。

通过高性能应用服务一键创建应用

最近更新时间：2024-05-10 16:40:51

步骤1：注册和充值

1. [注册腾讯云账号](#)，并进行实名认证。

如果您已在腾讯云注册，可忽略此步骤。

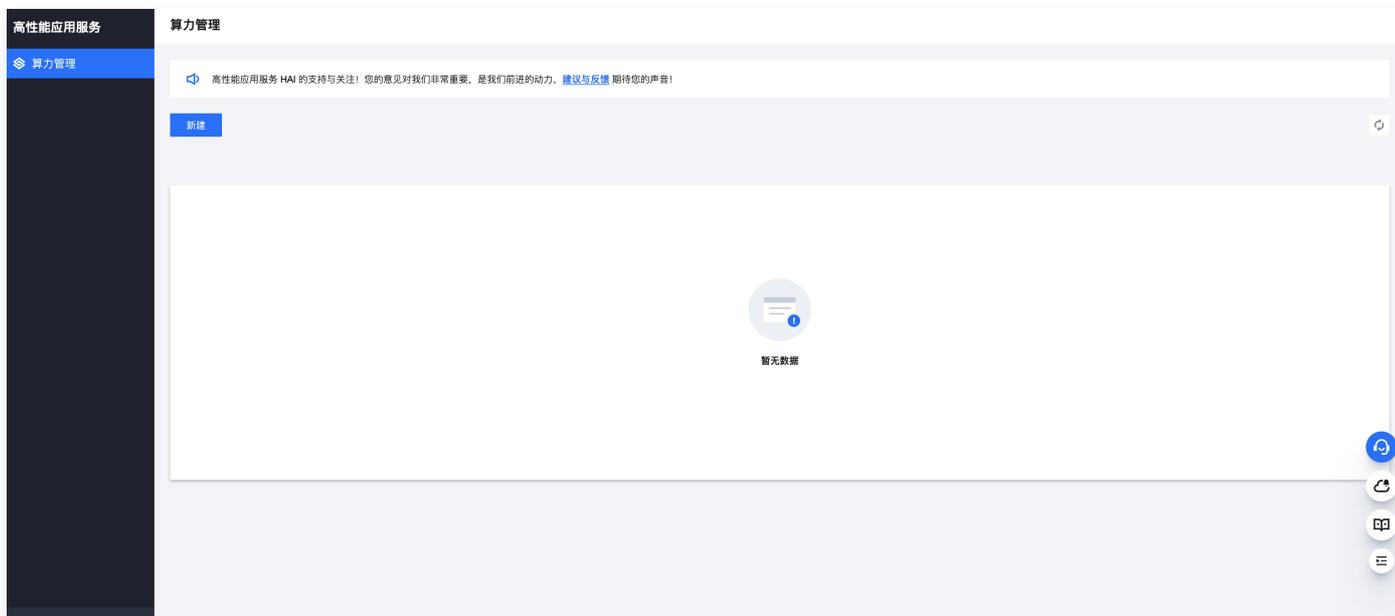
2. [在线充值](#)。

高性能应用服务器以按量计费模式售卖，购买前，需要在账号中进行充值。具体操作请参见 [在线充值](#) 文档。

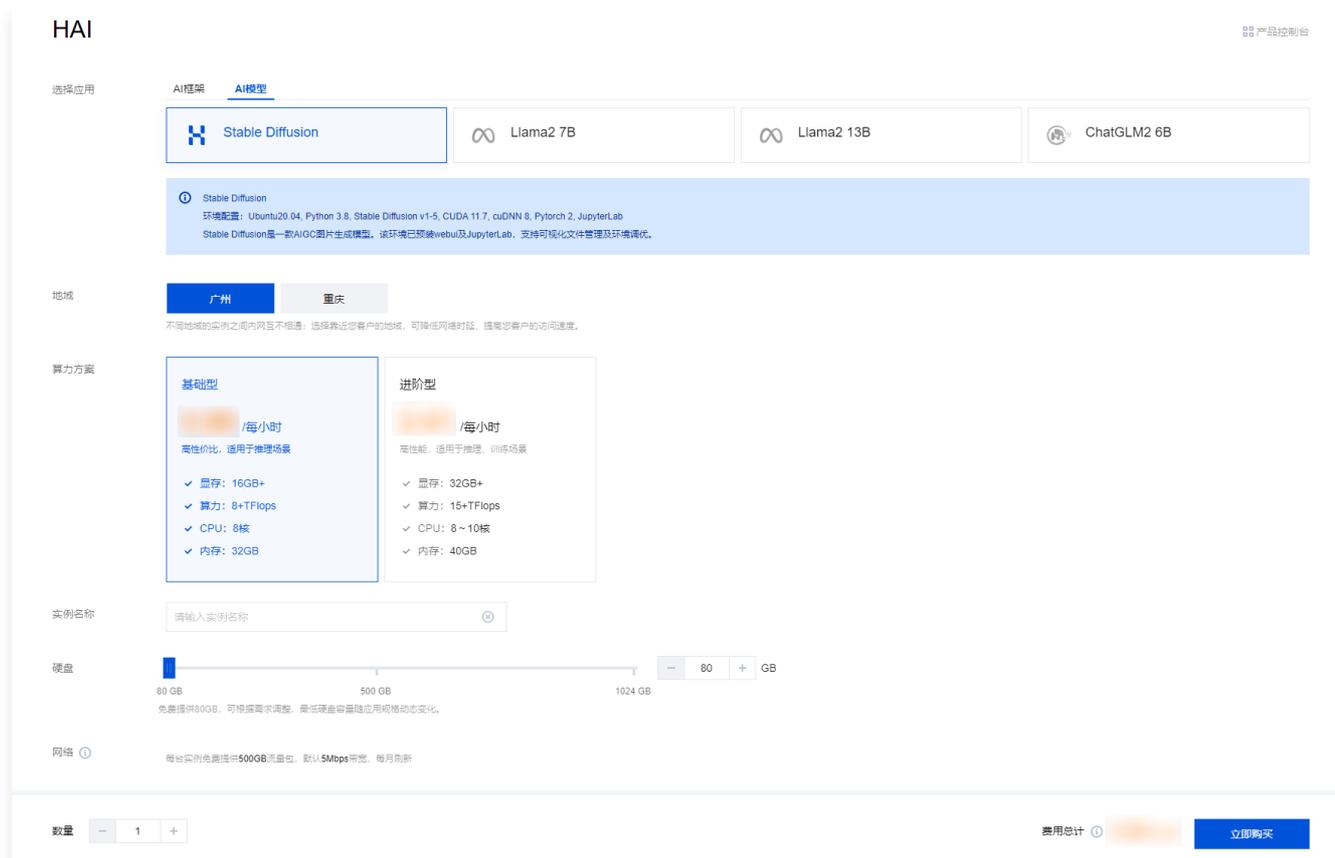
3. 如果您是传统账户，需先进行升级。判断账户类型或了解升级方法，可参见 [账户类型说明](#)。

步骤2：创建高性能应用服务

1. 登录 [高性能应用服务控制台](#)。



2. 单击**新建**，进入 [高性能应用服务购买页面](#)。



- **选择应用：** 目前提供 AI 框架、AI 模型两类应用。单击应用后可预览应用环境配置详情及应用介绍信息。
- **地域：** 建议选择靠近目标客户的地域，降低网络延迟、提高您的客户的访问速度。
- **算力方案：** 支持基础型及进阶型两类算力方案。用户可根据自己所需的配置进行选择。算力套餐对应的显存、算力、CPU、内存信息在算力方案卡片进行展示。
- **实例名称：** 自定义实例名称，若不填则默认使用实例 ID 替代。
- **硬盘：** 默认提供 80GB 免费空间，可根据实际使用需求进行调整。
- **网络：** 每台实例每月免费提供 500GB 流量包，默认 10Mbps 带宽，每月刷新。
- **购买数量：** 默认1台。

3. 单击**立即购买**。

4. 核对配置信息后，单击**提交订单**，并根据页面提示完成支付。

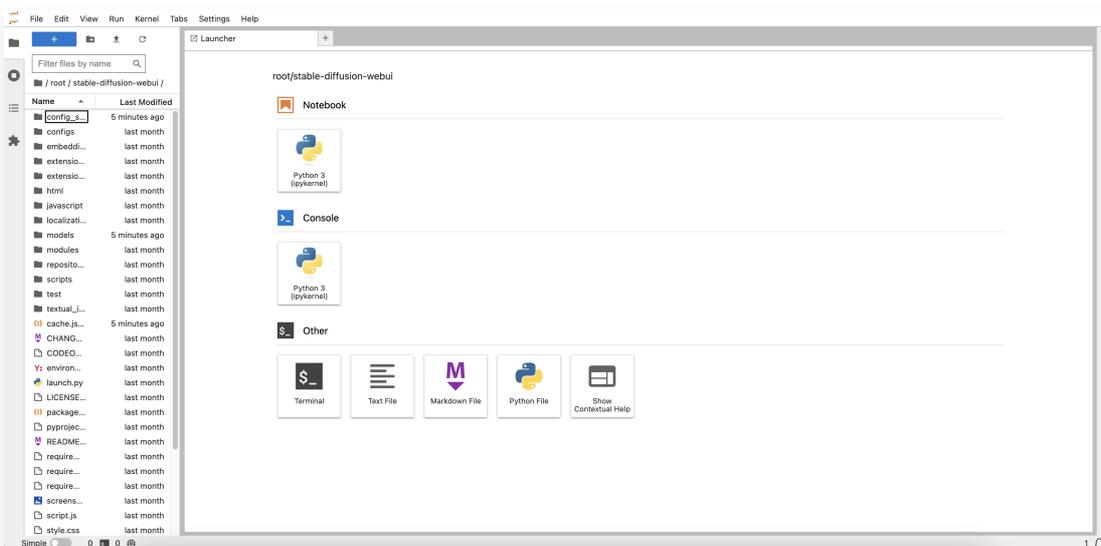
当您付费完成后，即完成了 Stable Diffusion 应用的创建。接下来，您可以登录实例并管理应用。

步骤3：连接实例

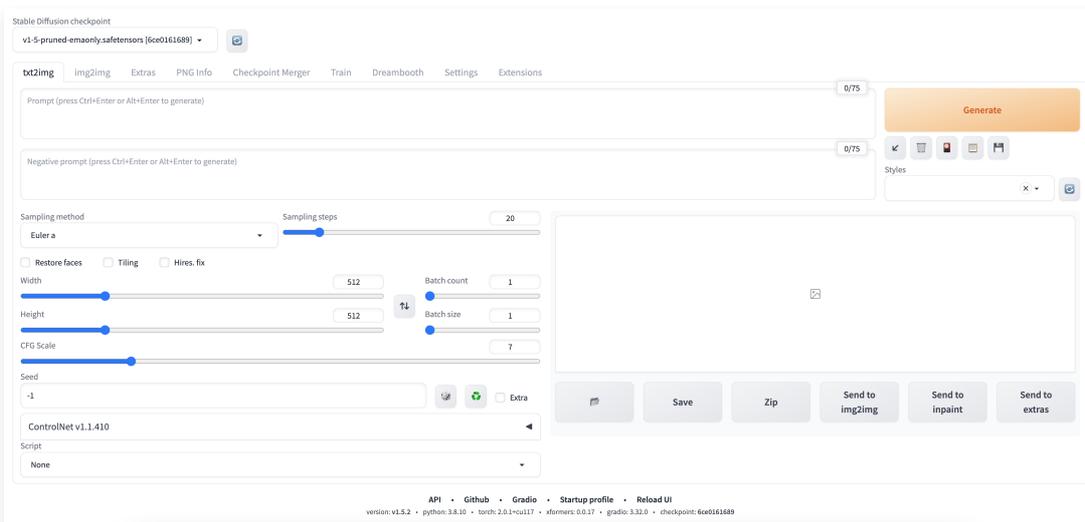
1. 登录 [高性能应用服务控制台](#)，在服务器列表中，单击**算力连接**。进行以下两种登录方式。



- 方式一：选择 JupyterLab 登录方式，进入 JupyterLab 终端。



- 方式二：选择 Gradio WebUI 登录方式，进入 Stable Diffusion checkpoint 可视化交互界面。

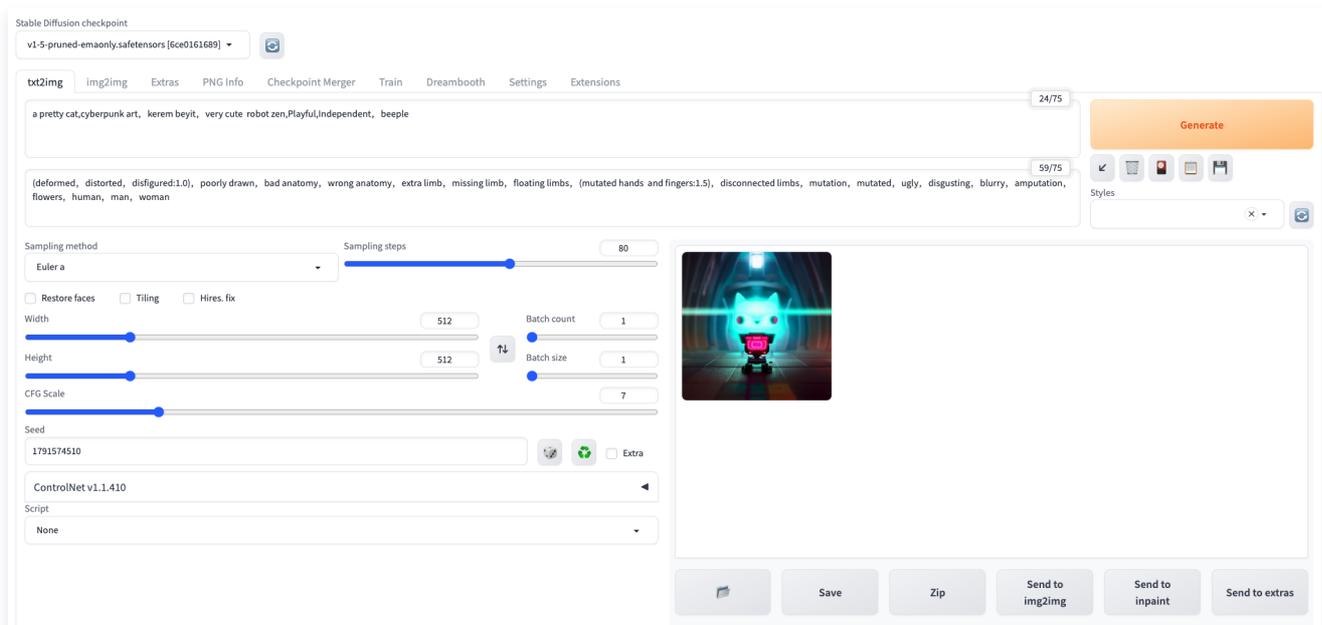


参考以下信息调整参数。

参数名	描述	值
Prompt	主要描述图像，包括内容风格等信息，原始的 WebUI 会对此处有字数的限制，您	a pretty cat,cyberpunk art, kerem beyit, very cute robot zen,Playful,Independent, beepie

	可以通过安装一些插件来突破字数的限制。	
Negative prompt	为了提供给模型，您不需要的风格。	(deformed, distorted, disfigured:1.0), poorly drawn, bad anatomy, wrong anatomy, extra limb, missing limb, floating limbs, (mutated hands and fingers:1.5), disconnected limbs, mutation, mutated, ugly, disgusting, blurry, amputation, flowers, human, man, woman
CFG scale	分类器自由引导尺度，图像与提示符的一致程度。越低产生的结果越有创意，数值越大成图越贴近描述文本。一般设置为7。	7
Sampling method	扩散算法的去噪声采样模式会影响最终效果，不同的采样模式的结果会有很大差异，一般是默认选择 euler。	Euler a
Sampling steps	在使用扩散模型生成图片时所进行的迭代步骤。需要注意的是，更高的迭代步数会消耗更多的计算时间和成本，但并不意味着一定会得到更好的结果。	80
Seed	随机数种子，生成每张图片时的随机种子。	1791574510

单击 **Generate** 即可生成图片，截图如下：



步骤4：销毁高性能应用服务（可选）

在使用结束后，在控制台算力管理页中，单击**更多 > 销毁**，即可销毁高性能应用服务，停止计费，结束使用。



The screenshot displays a service instance card in the Tencent Cloud console. At the top left is a blue icon. To its right, the instance status is shown as '运行中' (Running) in a green box, followed by '基础型' (Basic type) in a grey box, and '8核 32GB' (8 cores, 32GB) in a light blue box. Below this, the specifications are listed: '显存 16GB+ | 算力 8+TFlops | 系统盘 80GB'. The public IP address is shown as '公网IP: [redacted]' with a copy icon. The usage is '已使用: 8天21小时8分钟6秒'. On the right side, there are two dropdown menus: '算力连接' (Compute connection) and '更多' (More). A '销毁' (Destroy) button is located at the bottom right of the card.