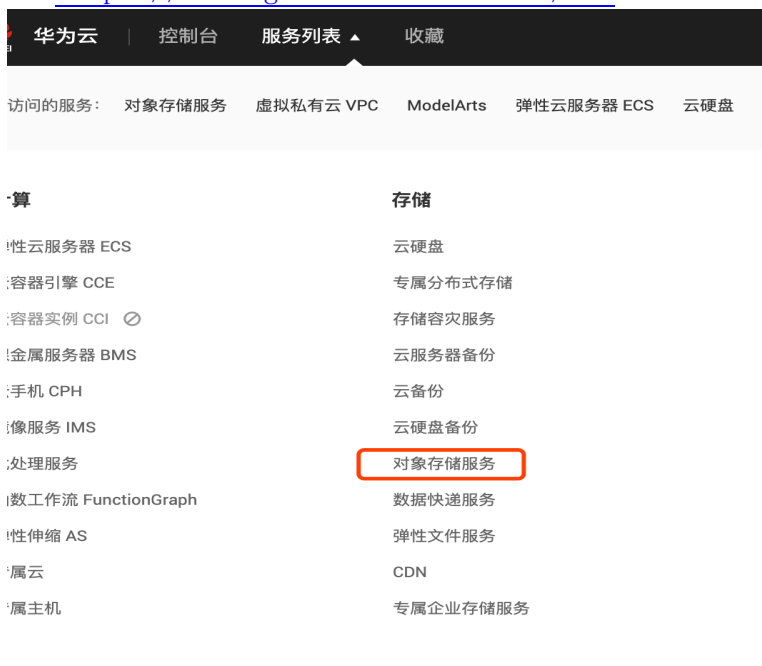


1. 首先创建 OBS 桶(功能类似云盘)，用来存储代码和训练数据。

<https://storage.huaweicloud.com/obs>



2. 跟随提示创建桶即可，就不详细说明。但注意 modelArts 在华北-北京四区，
OBS 桶的位置最好跟随 ModelArts。

创建完成后就要访问 OBS 桶，来上传数据。访问桶既可以如下图里网页上直接访问，也可以本地 PC 使用 OBS Browser 工具(Mac OS)上传数据(类似于 FileZilla 工具 FTP 传输)。建议小文件，即时访问 使用网页。大文件、多文件使用软件工具访问 OBS。



3. 使用 OBS Borwser 工具，会要求 **AK 和 SK 访问密钥**，在如下图 OBS 主页里即可新建 AK、SK 密钥。



我的 OBS Browsers 软件账号设置如下，账号跟华为账号无关，需要 填写 AK、SK，访问路径填桶名：

Tip1: 传输图片数据，建议打包成压缩包再上传。

Tip2: OBS 桶费用低，后面训练费用绝大部分是 ModelArts GPU 服务器费用，所以不必担心。

4. OBS 桶创建并且数据代码上传完毕即可去 ModelArts 创建 Notebook 进行训练。

<https://console.huaweicloud.com/modelarts>



跟随网站指示和自己需要创建 Notebook。但是**注意选择云硬盘**，不建议选择对象存储服务。并且**注意自动停止选项**，按需开启或关闭。

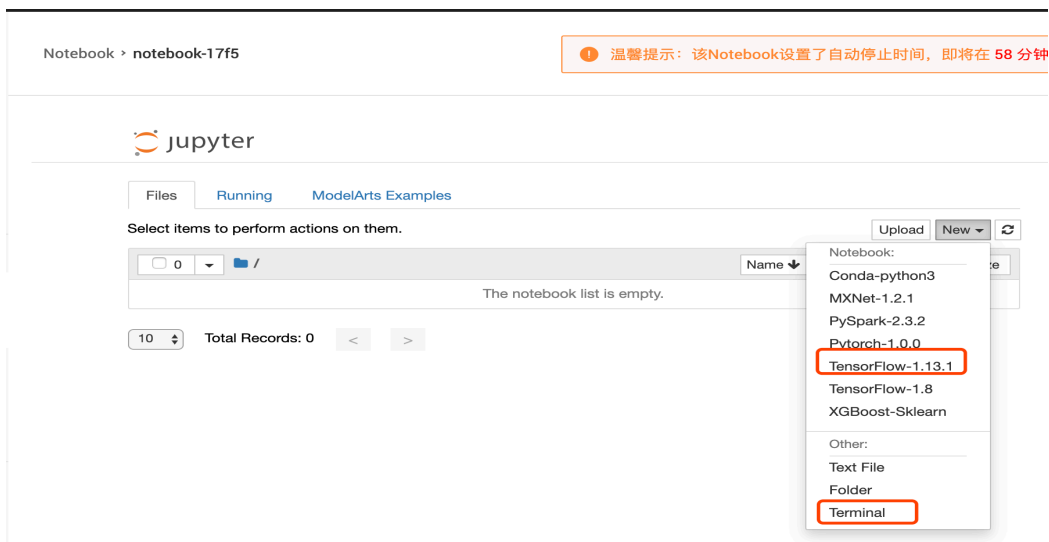
注 1：这里是创建了一个挂载在华为服务器的 Jupyter Notebook，这个服务器有 GPU，所以我们是利用这个服务器来训练数据，Notebook 只是我们用来修改代码的辅助工具。

注 2：之所以选择云硬盘，不选择对象存储服务(直接使用 OBS 桶)，因为我试过直接把 OBS 桶作为存储，有过 Notebook、终端控制台、OBS 桶三方数据不同步的问题(Notebook 中的 Sync 试过也没用)。而且云硬盘不贵，肯定比访问 OBS 桶更快。

5. Notebook 服务器创建成功后，就**需要把数据和代码从 OBS 桶传送到 Notebook 服务器上**。

如下图，notebook 创建标签，使用华为 Moxing 包，进行数据传输。后面有详细说明。

再创建 Terminal 终端来查看服务器情况。



Notebook 使用华为 moxing 包进行数据传输，从 OBS 桶复制数据和代码到 notebook 服务器。

关于 moxing 包传输数据更多技巧看 https://github.com/huaweicloud/ModelArts-Lab/blob/master/docs/moxing_api_doc/MoXing_API_File.md

这里只需要如下图两句即可（从 OBS 桶拿过来、送到 OBS 桶里）。

注意 OBS 桶名：s3://bucket_name/。s3 不变，bucket_name 换成桶名。这个路径是桶根目录。

传输完成后，notebook 会自动显示。Terminal 终端也会自动显示。

```
In [2]: import moxing as mox
import os
mox.file.shift('os', 'mox')

OBS_PATH = 's3://ustc369/hwr_crnn/'
LOCAL_PATH = "/home/ma-user/work/hwr_crnn/"

train_paths = os.listdir(OBS_PATH)
for i, class_path in enumerate(train_paths):
    print(class_path)

## copy parallel这个语句只能复制文件夹
## 从OBS复制文件夹到本地
#mox.file.copy_parallel(OBS_PATH, LOCAL_PATH)

## 不管用，请在控制台解压
#!unzip {DATA_ZIP_DIR+TRAIN_ZIP_FILENAME}

## 上传文件夹到OBS
mox.file.copy_parallel('s3://ustc369/hwr_crnn_1127/', "/home/ma-user/work/hwr_crnn/")

## 把Notebook服务器的数据传输到OBS桶做备份。
#mox.file.copy_parallel("/home/ma-user/work/hwr_crnn/", 's3://ustc369/hwr_crnn_1127/')
```

```
import moxing as mox
import os
mox.file.shift('os', 'mox')
```

```

OBS_PATH = 's3://ustc369/hwr_crnn/'
LOCAL_PATH = "/home/ma-user/work/hwr_crnn/"

train_paths = os.listdir(OBS_PATH)
for i, class_path in enumerate(train_paths):
    print(class_path)

## copy_parallel 这个只能复制文件夹

## 上传文件夹到 OBS
mox.file.copy_parallel('s3://ustc369/hwr_crnn_1127/', "/home/ma-user/work/hwr_crnn/")

## 把 Notebook 服务器的数据传输到 OBS 桶做备份。
#mox.file.copy_parallel("/home/ma-user/work/hwr_crnn/", 's3://ustc369/hwr_crnn_1127/')

```

6. 代码后数据传输到 Notebook 服务器后，就可以自己使用 Terminal 终端训练数据了。下面说几个 tip。

Tip1: 服务器是使用 Anaconda 的，初始时是不在任何环境中，也就是没有 python 和 tensorflow。需要使用命令激活 tensorflow 环境：

```
$ source /home/ma-user/anaconda3/bin/activate TensorFlow-1.13.1
```

Tip2: 跑代码时候使用如下指令，防止关闭终端时自动把正在运行的项目关闭(注意红色标识部分，尤其后面的 2>&1 & 别修改。只需修改.py 文件和 log 文件)：

```
$ nohup python -u run.py > 20191126.log 2>&1 &
```

