

Tensorflow-GPU 集群安装指南

集群中的python模块中并没有安装所有用户所需要的库，对于有进行机器学习或网络训练需求的用户，推荐在自己的节点中安装相应版本的 Anaconda。作为一个成熟的python发行版本，其包含了conda, python 及相关的科学计算包。

Anaconda 的官方网站: <https://www.anaconda.com/>

清华开源镜像站 (Anaconda目录) : <https://mirrors.tuna.tsinghua.edu.cn/anaconda/archive/>

选择Linux版本，进行下载（国内下载安装Anaconda可能会面临下载速度慢的问题，可在清华开源镜像站内选择相应的版本下载地址进行下载）：

输入 `wget anaconda_download_address`

```
[xiesj@hpc-login-gpu01 ~]$ wget https://mirrors.tuna.tsinghua.edu.cn/anaconda/archive/Anaconda3-2019.10-Linux-x86_64.sh
--2020-12-04 14:36:09-- https://mirrors.tuna.tsinghua.edu.cn/anaconda/archive/Anaconda3-2019.10-Linux-x86_64.sh
Resolving mirrors.tuna.tsinghua.edu.cn (mirrors.tuna.tsinghua.edu.cn)... 101.6.8.193, 2402:f000:1:408:8100::1
Connecting to mirrors.tuna.tsinghua.edu.cn (mirrors.tuna.tsinghua.edu.cn)|101.6.8.193|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 530308481 (506M) [application/octet-stream]
Saving to: 'Anaconda3-2019.10-Linux-x86_64.sh'

100%[=====>]
530,308,481 4.02MB/s in 2m 27s

2020-12-04 14:38:37 (3.43 MB/s) - 'Anaconda3-2019.10-Linux-x86_64.sh' saved [530308481/530308481]
```

下载成功后，运行下载脚本进行安装: `bash anaconda_installer`

```
[xiesj@hpc-login-gpu01 ~]$ bash Anaconda3-2019.10-Linux-x86_64.sh
```

根据提示，在相关问题下输入 `yes` 或 `ENTER`，之后看到欢迎语即为安装成功，过程大概三四分钟：

```
Thank you for installing Anaconda3!

=====

Anaconda and JetBrains are working together to bring you Anaconda-powered environments tightly integrated in the PyCharm IDE.

PyCharm for Anaconda is available at:
https://www.anaconda.com/pycharm
```

执行 `source ~/.bashrc` 命令，完成环境变量的更改。

```
[xiesj@hpc-login-gpu01 ~]$ source ~/.bashrc
```

Anaconda 安装成功后，之后就能用 `conda install` 命令安装相关的python工具包。我们用Conda 在集群上安装 tensorflow-GPU, 输入 `conda install tensorflow-gpu` 命令：

```
(base) [xiesj@hpc-login-gpu01 ~]$ conda install tensorflow-gpu
```

约一两分钟安装完成后，启动python 加载tensorflow，若如下使用正常没有报错即安装成功。

```
(base) [xiesj@hpc-login-gpu01 GAN]$ python
Python 3.7.4 (default, Aug 13 2019, 20:35:49)
[GCC 7.3.0] :: Anaconda, Inc. on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> import tensorflow as tf
>>> tf.__version__
'2.0.0'
```

安装成功之后即可在程序中引用tensorflow包，并在集群中运行：

```
(base) [xiesj@hpc-login-gpu01 GAN]$ python GAN.py
Extracting /tmp/data/train-images-idx3-ubyte.gz
Extracting /tmp/data/train-labels-idx1-ubyte.gz
Extracting /tmp/data/t10k-images-idx3-ubyte.gz
Extracting /tmp/data/t10k-labels-idx1-ubyte.gz
```

```
Step 1: Generator Loss: 0.875505, Discriminator Loss: 1.243211
Step 2000: Generator Loss: 4.591123, Discriminator Loss: 0.037789
Step 4000: Generator Loss: 4.309066, Discriminator Loss: 0.031199
Step 6000: Generator Loss: 3.967848, Discriminator Loss: 0.212809
```

注意事项：

1. 其他类似的机器学习框架如: Pytorch 等也可以在Anaconda中用类似的流程进行安装。
2. 在使用Anaconda 时应删除已经加载的集群中的python模块，而是用Anaconda中的python。
3. 如果对于python的科学运算有不同的需求，建议在安装工具库时，先用conda 创建相关的虚拟环境，保证工具库的有序和安全。
4. 对于一些确保运算量非常小的程序demo，可以直接在节点进行运行测试，其他的代码一律需要编写PBS脚本提交运行代码，以确保集群的运行。（若不能确定demo的运算量，不可在节点直接运行）