

算力服务器使用指北

By 徐纪元

一、使用 Torch 和 python 进行深度学习

1 创建环境

首先选择环境：

只需要 torch 就使用左边这个，需要 conda 和 ollama 就使用右边这 2 个

端口转发按需填写，但是也只能校内访问。

2 等待环境运行

本地电脑安装 vs code 和 xftp, 参考

<https://172.21.8.10:32108/zh-CN/user/user-tool/vscode-use.html> 使用远程连接

环境名称	状态	运行时长	剩余时长	资源信息	节点	镜像
2025021 8125356	运行中	3天 1时 5 9分 58秒	-	NVIDIA- A800-80 GB-PCIe: 1 CPU: 48	node5(14 9,77,95)	172.21.8. 10:5000/ other/pth- t2005002 7_25021 1-120050 027:2502 11

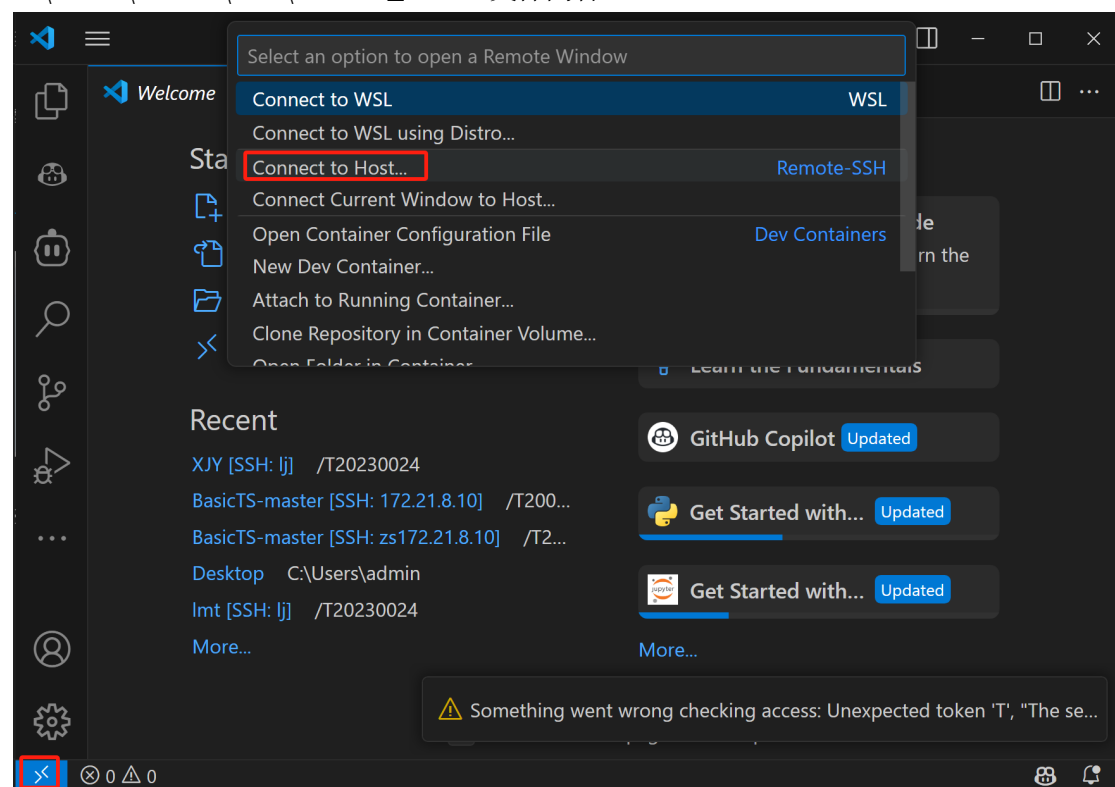
连接命令: `ssh root@172.21.8.10 -p 30511`复制

SSH密码: *****复制

当SSH服务异常或服务没有启动,连接不可用

开发环境 18 12:44: 查看 : 克隆 资源调整 删除

在 vs code 安装 Remote SSH, 注意 windows 用户名为英文, 注意清空 C:\Users\admin\.ssh\known_hosts 文件内容



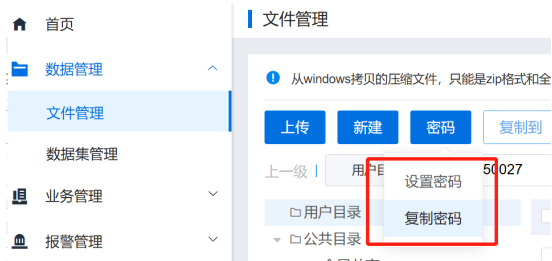
填写对应的连接命令, 选择 linux, 输入密码即可。多个环境如下配置:

```
config
C: > Users > admin > .ssh > config
1 Host abc
2   HostName 172.21.8.10
3   User root
4   Port 30511
5 Host aaabs
6   HostName 172.21.8.10
7   User root
8   Port 30227
9
```

3 连接环境文件和端口

在 xftp 输入下述信息：

The image shows the '算力集群z5属性' (Properties of 算力集群z5) dialog box. It has two tabs: '常规' (General) and '选项' (Options). In the '常规' tab, the following fields are visible: '名称(N):' (Name) is '算力集群z5'; '主机(H):' (Host) is '172.21.8.10'; '协议(R):' (Protocol) is 'SFTP'; '端口号(O):' (Port) is '22'; '代理服务器(X):' (Proxy server) is '<无>' (None); '说明(D):' (Description) is empty. In the '选项' tab, the following options are visible: '登录' (Login) section has '匿名登录(A)' (Anonymous login) unchecked, '使用SSH密钥代理(Xagent)(G)' (Use SSH key agent) unchecked, and '方法(M):' (Method) has 'Password' checked, 'Public Key' unchecked, 'Keyboard Interactive' unchecked, 'GSSAPI' unchecked, 'PKCS11' unchecked, and 'CAPI' unchecked. The '用户名(U):' (Username) is 'T20050027' and the '密码(P):' (Password) is masked with dots. At the bottom, there are buttons for '连接' (Connect), '确定' (OK), and '取消' (Cancel).



密码从平台获取，账号写自己的，就可以上传文件。密码过期就连不上了，重置密码就可以。

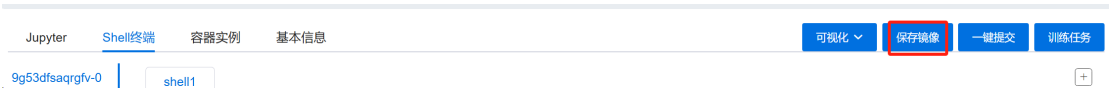
The image shows the '容器实例' (Container Instance) page. At the top, there are tabs for 'Jupyter', 'Shell终端', '容器实例', and '基本信息'. There are buttons for '可视化' (Visualize), '保存镜像' (Save image), '一键提交' (One-click submit), and '训练任务' (Training task). Below the tabs is a table with the following columns: '容器名称' (Container name), '状态' (Status), '节点IP' (Node IP), '容器IP' (Container IP), '加速卡' (Accelerator card), '端口' (Port), and '操作' (Action). The table has one row with the following data: '容器名称' is '9g53dfsagrgfv-0', '状态' is '运行中' (Running), '节点IP' is '172.21.8.4', '容器IP' is '172.21.9.130', '加速卡' is 'node5_1', '端口' is '7100->30100;30000->30000;8085->30185;7200->30200;8081->30181;8080->30180;22->30511;1;8890->30890;8888->48733', and '操作' is 'SSH'.

容器名称	状态	节点IP	容器IP	加速卡	端口	操作
9g53dfsagrgfv-0	运行中	172.21.8.4	172.21.9.130	node5_1	7100->30100;30000->30000;8085->30185;7200->30200;8081->30181;8080->30180;22->30511;1;8890->30890;8888->48733	SSH

端口访问请点击环境，具体看看“节点 ip”是多少，而不是 172.21.8.10，然后访问 <http://172.21.8.4:30180/> 即可，容器内部端口为 8080，需要在创建环境前设置。

4 使用

可以使用 nvidia-smi 监控运行，直接 python 或 pip 使用运行文件。每次重启都是镜像的初始状态，如果需要保存新安装的软件，点击：



期间会无法访问开发环境，下次用新保存的镜像启动即可。

5 纠错

如果遇到驱动问题，如 Segmentation fault(core dumped)之类的，请参考以下命令：

```
sudo find / -name libcuda.so*
```

```
ls /usr/lib/x86_64-linux-gnu/libcuda.so.1 -la
```

```
ls /usr/lib/x86_64-linux-gnu/libcuda.so -la
```

将列出的文件 libcuda.so.565.57.01 填到下面

```
sudo ln -f -s /usr/lib/x86_64-linux-gnu/libcuda.so.565.57.01  
/usr/lib/x86_64-linux-gnu/libcuda.so
```

正确链接驱动即可，一般是 L40s 和 H800 节点采用不同版本驱动造成的。当然也有可能是代码里面指针溢出，检查 dict 多线程取值和重复计算梯度问题。

开放端口运行 ollama 参考：

```
export OLLAMA_HOST=0.0.0.0:11434
```

```
nohup ollama serve &
```

```
nohup ollama pull deepseek-r1:32b &
```

```
ollama pull mxbai-embed-large
```

```
ps -ef |grep ollama
```

运行 openwebui：

```
export LD_LIBRARY_PATH=/usr/lib/x86_64-linux-gnu:$LD_LIBRARY_PATH
```

```
conda create -n ds python=3.11
```

```
conda activate ds
```

```
pip install open-webui
```

```
open-webui serve
```

```
export HF_ENDPOINT="https://hf-mirror.com"
```

```
/bin/bash -c "source ~/anaconda3/bin/activate ds && open-webui serve"
```

```
ps -aux | grep "open-webui serve"
```

临时使用升级库镜像：

```
sudo vi /etc/apt/sources.list
```

```
deb http://mirrors.aliyun.com/ubuntu/ jammy main
```

```
sudo apt update
```

```
sudo apt install libstdc++6
```

```
sudo apt install libcurl4
```