GZU超算中心配置Anaconda环境

by：李浩楠

GZU超算中心配置Anaconda环境

一、使用Xshell连接集群进行登录

一、 ①设置名字，随便写。比如：贵州大学超算中心

二、 ②获取IP。

三、填完后点击连接

四、登陆成功

二、获取Anaconda环境的Linux版本

1.直接用wget获取

2.去到Anaconda官网直接下载后，利用xftp或者超算官网直接传到自己的登录环境

三、安装配置过程

1.安装脚本

2.然后一路回车，同意，等待安装完成。

3.当安装成功后不设置安装路径，则默认安装在：anaconda3这个文件夹中。

4.设置conda环境变量

5.测试是否安装成功

四、配置清华源镜像

1.在使用过程中可能会 碰到以下错误：

2.配置清华源

五、超算中心的一些常用操作

1.sinfo命令

2.squeue命令

3.简单演示

写在前面：禁止在workstation(登陆)节点运行任务，包括【使用conda创建个人env】等占用内存资源的操作。

请先使用salloc、sbatch、srun命令申请资源，再进行下一步操作。

一、使用Xshell连接集群进行登录

> 下载Xshell

- Xshell中文网

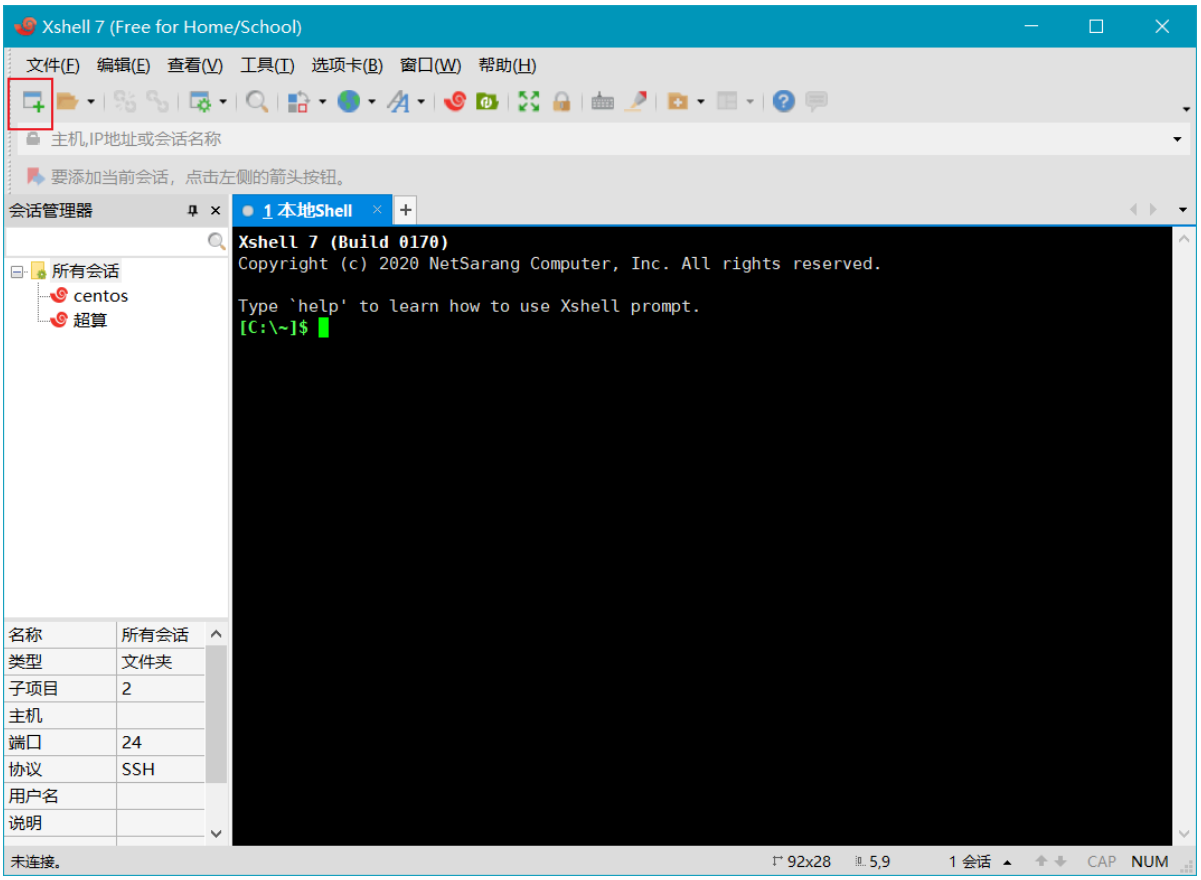
```
https://www.xshell.com/zh/xshell/
```

- Xftp中文网

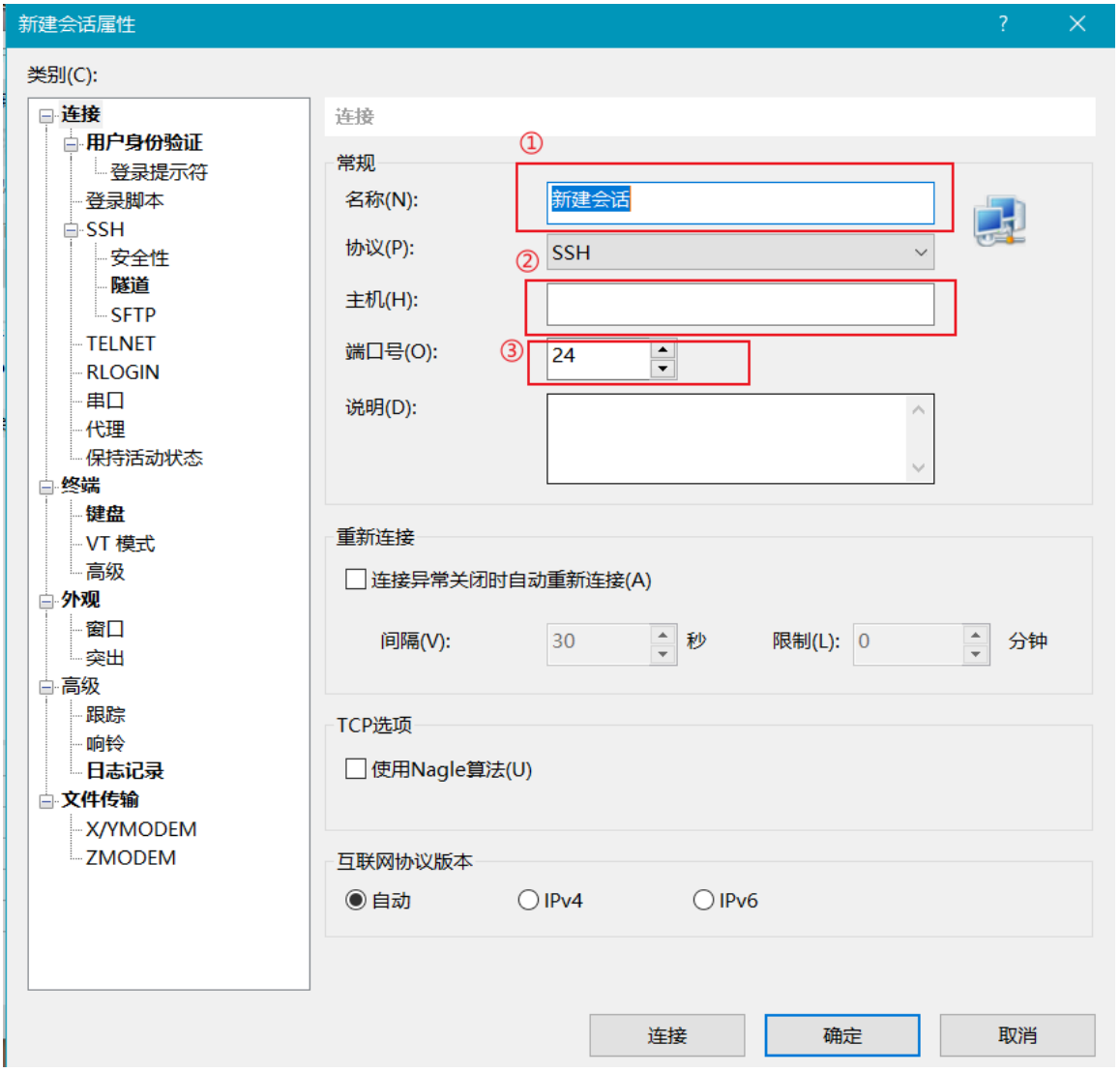
```
https://www.xshell.com/zh/xftp/
```

- 上述两个软件免费，按照指引安装即可。

打开Xshell


Xshell 7 (Free for Home/School)
文件(F) 编辑(E) 查看(V) 工具(T) 选项卡(B) 窗口(W) 帮助(H)
主机,IP地址或会话名称
要添加当前会话，点击左侧的箭头按钮。
会话管理器
所有会话
centos
超算
1 本地Shell
Xshell 7 (Build 0170)
Copyright (c) 2020 NetSarang Computer, Inc. All rights reserved.
Type `help' to learn how to use Xshell prompt.
[C:\~]$
名称 所有会话
类型 文件夹
子项目 2
主机
端口 24
协议 SSH
用户名
说明
未连接。
92x28
5,9
1 会话
CAP
NUM

- 点击左上角红框


新建会话属性
类别(C):
连接
用户身份验证
登录提示符
登录脚本
SSH
安全性
隧道
SFTP
TELNET
RLOGIN
串口
代理
保持活动状态
终端
键盘
VT 模式
高级
外观
窗口
突出
高级
跟踪
响铃
日志记录
文件传输
X/YMODEM
ZMODEM
连接
常规
①
名称(N):
新建会话
协议(P):
②
SSH
主机(H):
端口号(O):
③
24
说明(D):
重新连接
连接异常关闭时自动重新连接(A)
间隔(V):
30
秒
限制(L):
0
分钟
TCP选项
使用Nagle算法(U)
互联网协议版本
自动
IPv4
IPv6
连接
确定
取消

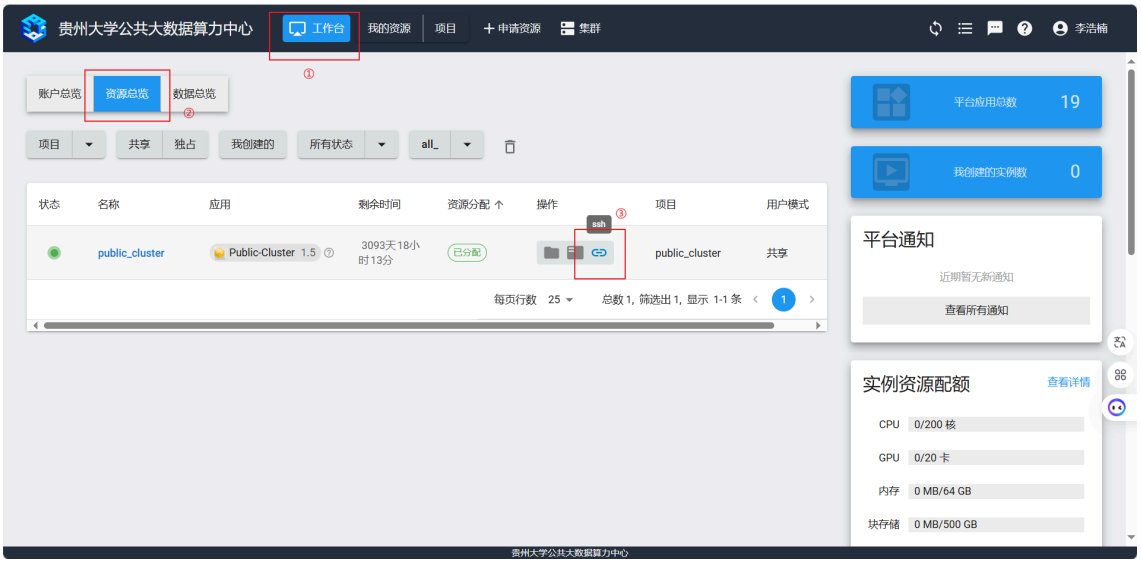
一、 ①设置名字，随便写。比如：贵州大学超算中心

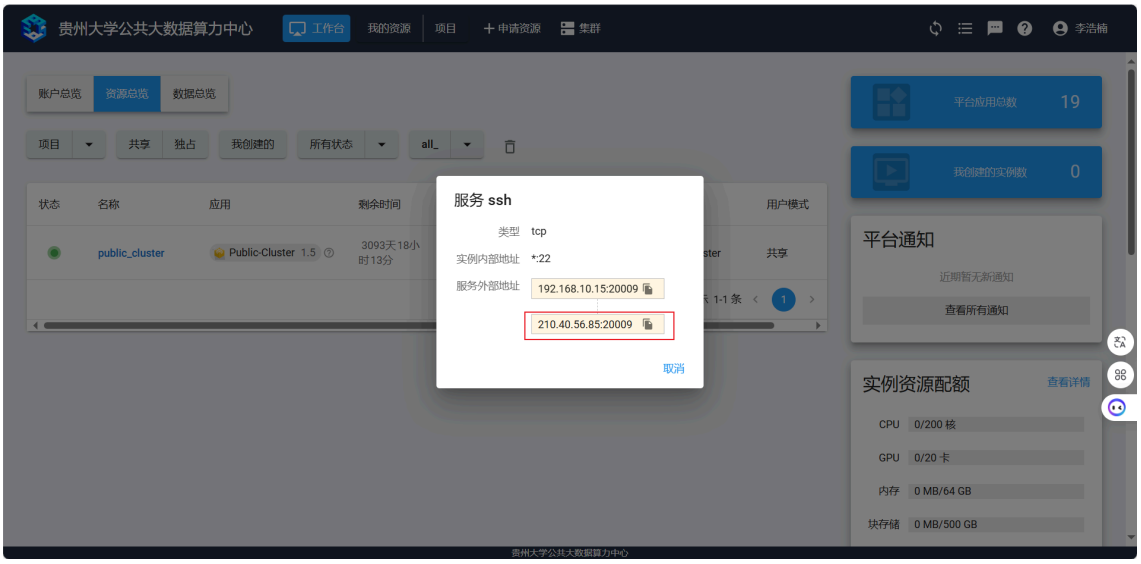
二、 ②获取IP。

1. 访问超算中心

   登录时候不要挂VPN

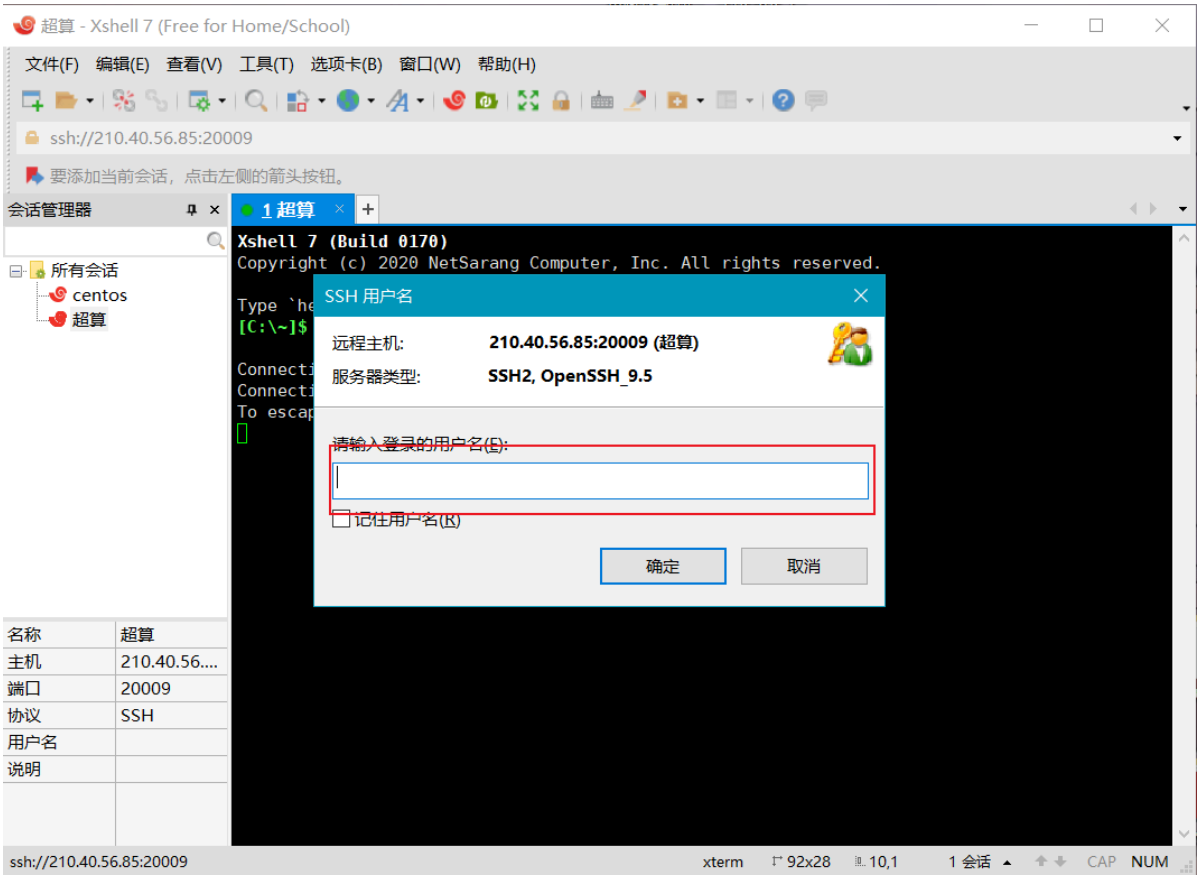
2. 点击工作台----->资源总览----->ssh





3. 将红框中的ip地址以及端口填入会话属性的②和③。

三、填完后点击连接



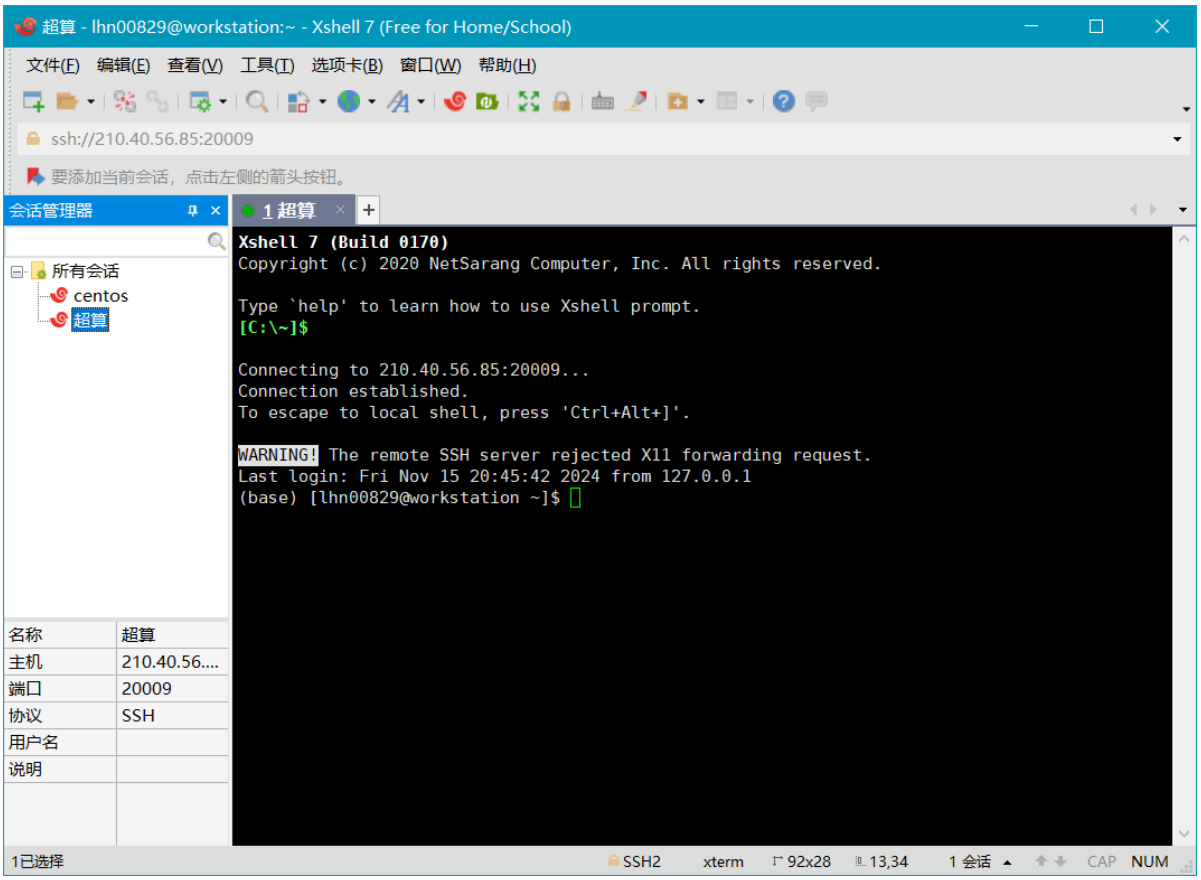
- 红框中输入当时申请时自己起的用户名



- 选择password，然后输入当时注册时的密码

四、登陆成功

成功会显示**[用户名@workstation~]**如下图所示（**base**）是我已经配好环境了，没配是没有显示的，这个配置过程往下看。然后可以配合Xftp将windows文件传入超算，或是通过超算网页的cmd进行传输。



二、获取Anaconda环境的Linux版本

1.直接用wget获取

```
wget https://repo.anaconda.com/archive/Anaconda3-2023.03-1-Linux-x86_64.sh
```

因为官网比较慢，也可以利用镜像下载

```
wget https://mirrors.tuna.tsinghua.edu.cn/anaconda/archive/Anaconda3-2023.03-1-Linux-x86_64.sh
```

2.去到Anaconda官网直接下载后，利用xftp或者超算官网直接传到自己的登录环境

- anaconda官网

```
https://www.anaconda.com/
```

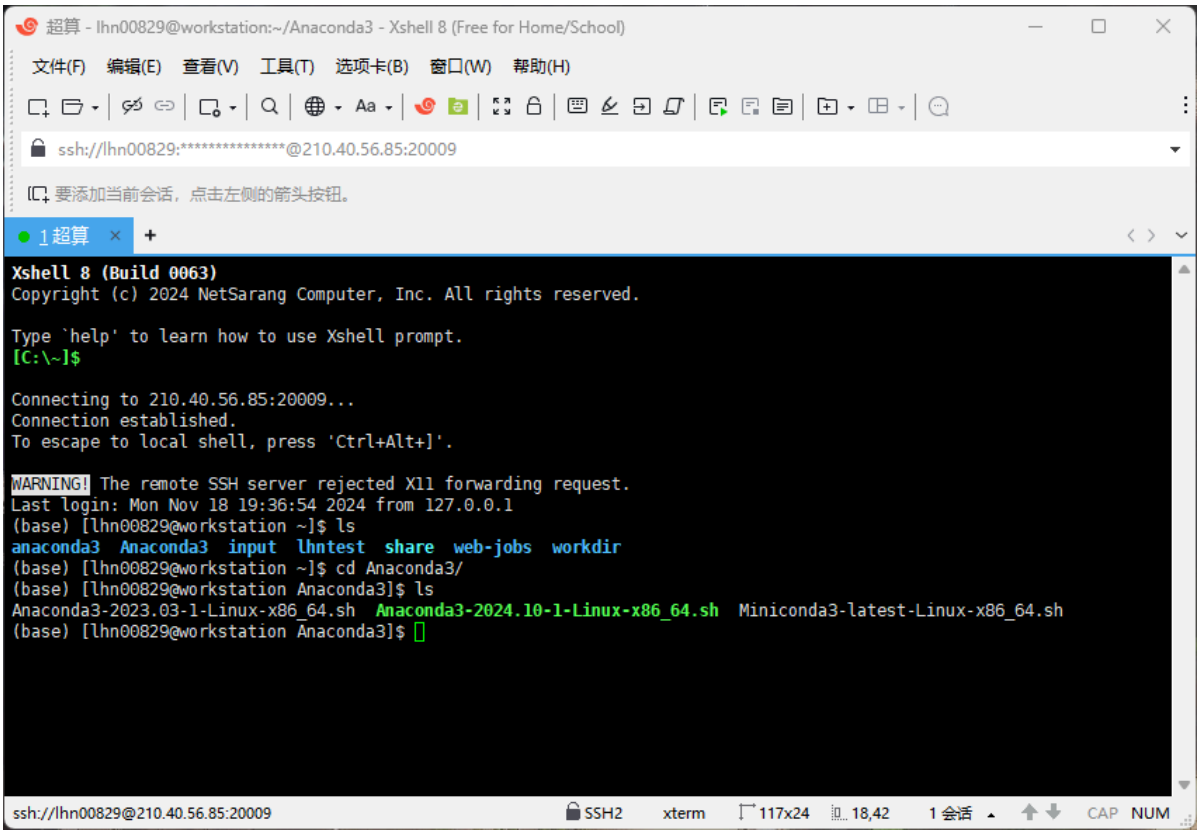
- 所有anaconda版本

```
https://repo.anaconda.com/archive/index.html
```

建议下载 如下版本：

```
Anaconda3-2023.03-1-Linux-x86_64.sh
```

下载后在自己 的workstation中找到：



如上图绿色高亮是其中一个版本，只下载一个的话就只有一个。

三、安装配置过程

1.安装脚本

使用bash命令（按tab可快速填充包名字）：

```
bash Anaconda3-2023.03-1-Linux-x86_64.sh
```

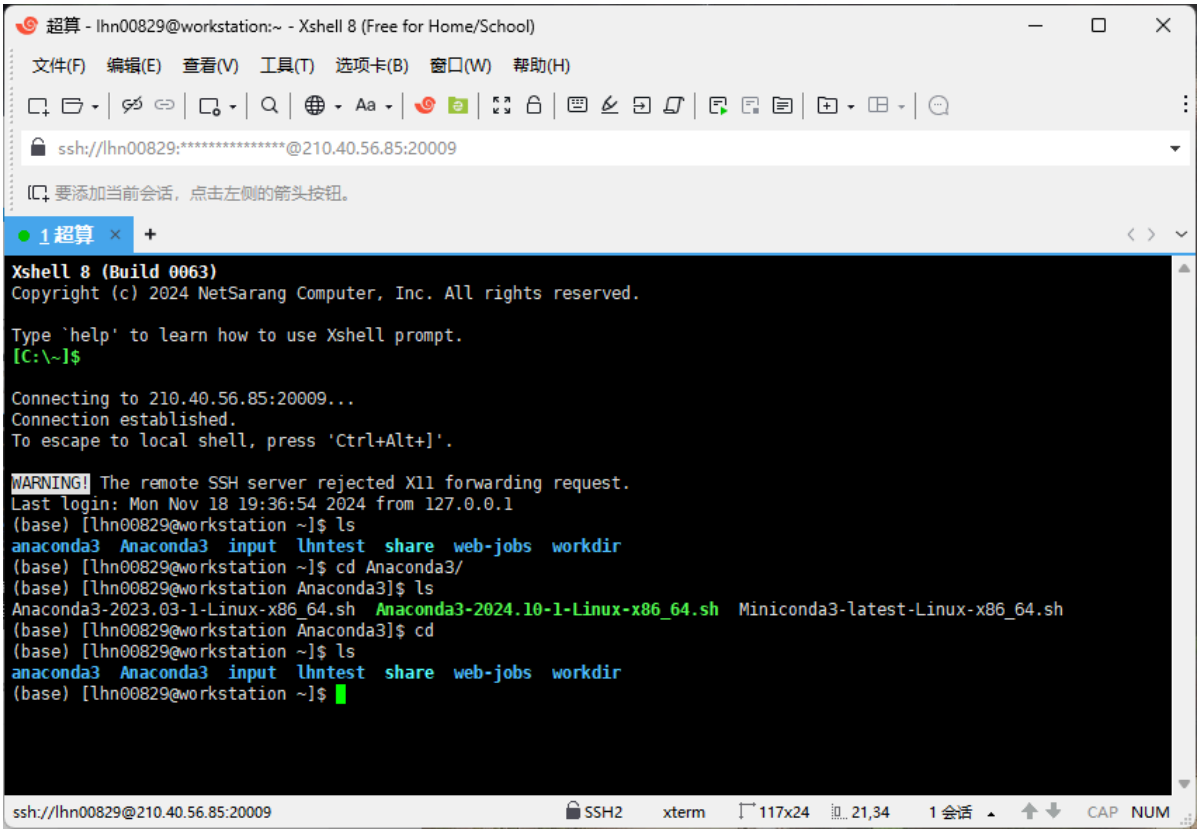
如果报错：未找到AnaAnaconda3-2023.03-1-Linux-x86_64.sh，就重新连接一下。

2.然后一路回车，同意，等待安装完成。

这一步可能会报网络错误，如果出错换一个版本，尽量不要用最新的版本。

碰到问题可以跟我讲，我来更新这个文档

3.当安装成功后不设置安装路径，则默认安装在：anaconda3这个文件夹中。



4.设置conda环境变量

因为超算自己的账户不是root账户，所以不能更改**/etc/profile**中的路径。需要更改用户环境变量**~/.bashrc**

```
vim ~/.bashrc
```

vim常用操作：

- 按i键：输入模式
- 输入完成后按ESC，输入：wq 是保存后退出

加入以下路径变量：

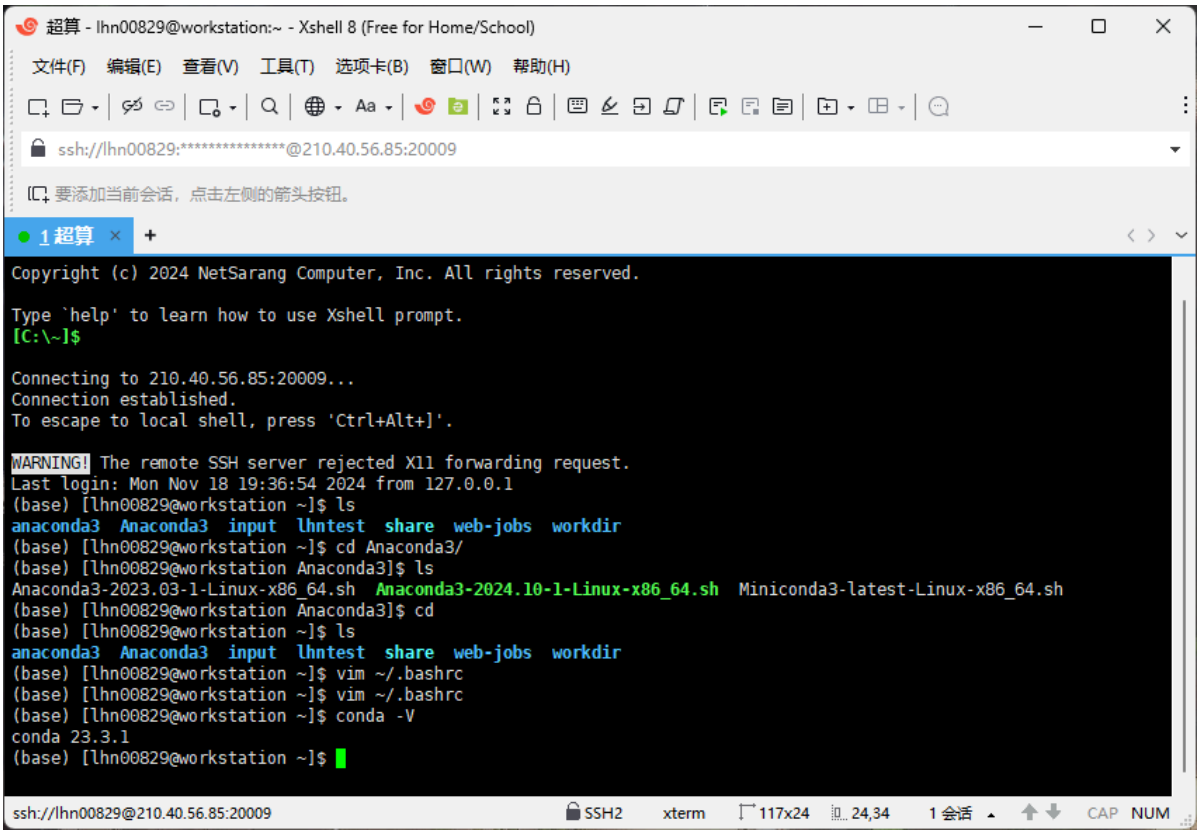
```
export PATH=~/anaconda3/bin:$PATH
```

刷新环境变量

```
source /etc/profile
source ~/.bashrc
```

5.测试是否安装成功

```
conda -V
```



安装成功后会显示 **conda 版本。**

四、配置清华源镜像

先看是否能够正常使用conda install 下载相应的包，能下就不配这个镜像。或者建议使用pip命令进行安装，需要的wheel等可问GPT，镜像有的包会找不到下载。

当conda install报HTTP:000时优先去官网找包和命令

- anaconda官网

```
https://anaconda.org/
```

1.在使用过程中可能会 碰到以下错误：

```
Collecting package metadata (repodata.json): failed

CondaHTTPError: HTTP 000 CONNECTION FAILED for url
<https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/main/linux-
64/repodata.json>
Elapsed: -

An HTTP error occurred when trying to retrieve this URL.
HTTP errors are often intermittent, and a simple retry will get you on your
way.
'https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/main/linux-64'
```

以上问题是因为网络问题，可能清华源镜像挂掉了，尝试更新其他的镜像，或使用VPN。接下来介绍配置清华源的方法，其他镜像类似。

2.配置清华源

- 清华源镜像

```
https://mirror.tuna.tsinghua.edu.cn/help/anaconda/
```

添加镜像：

```
vim ~/.condarc
```

把其中内容更改如下：

```
channels:
  - defaults
show_channel_urls: true
default_channels:
  - https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/main
  - https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/r
  - https://mirrors.tuna.tsinghua.edu.cn/anaconda/pkgs/msys2
custom_channels:
  conda-forge: https://mirrors.tuna.tsinghua.edu.cn/anaconda/cloud
  pytorch: https://mirrors.tuna.tsinghua.edu.cn/anaconda/cloud
```

然后清除索引，并安装常用的包尝试一下

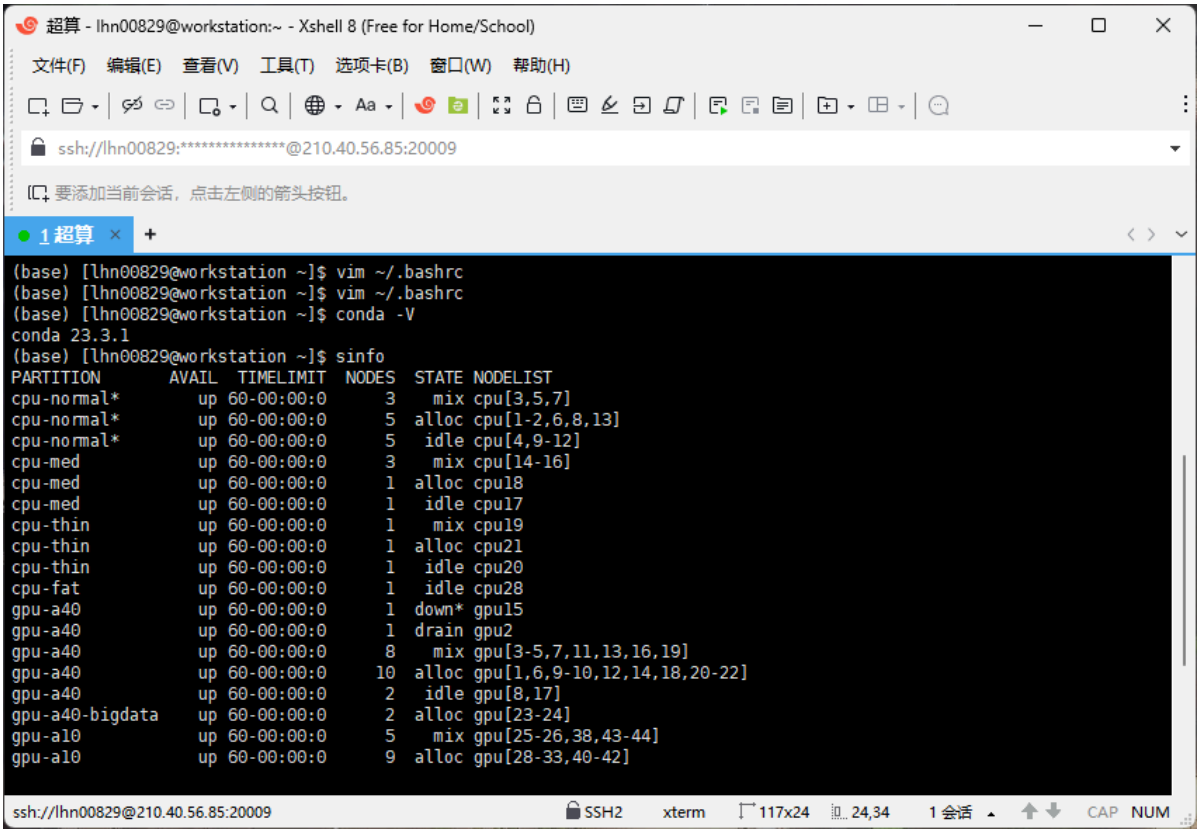
```
conda clean -i
conda create -n [你的环境名字] numpy
```

完成以上操作，conda环境就已经安装完成了，可以根据项目自行配置pytorch环境。

五、超算中心的一些常用操作

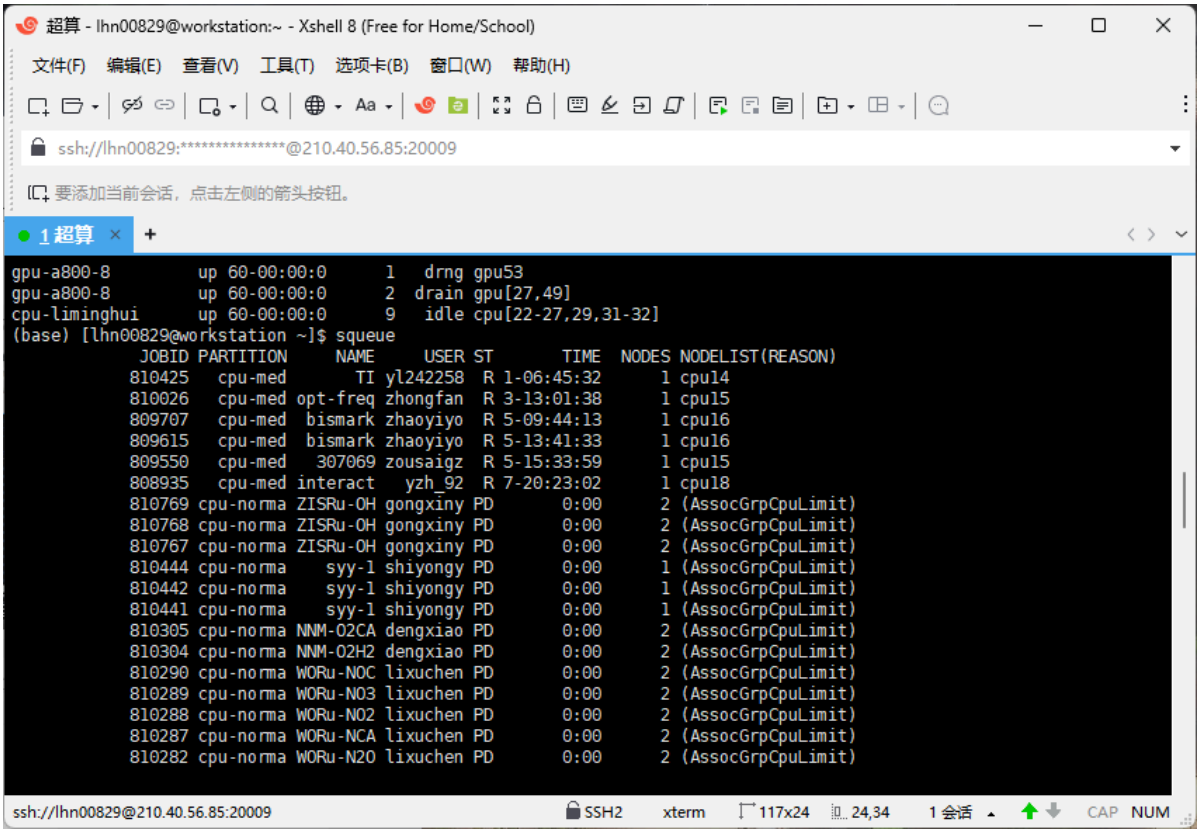
命令	效果	备注
`sinfo`	查看节点与分区状态	
`squeue`	查看队列状态	
`scancel`	取消作业	
`sacct`	查看历史作业信息	
`salloc`	分配资源	
`sbatch`	提交批处理作业	
`scontrol`	系统控制	
`srun`	执行作业	

1.sinfo命令



关键词	含义
`PARTITION`	分区名，对节点的逻辑分组。不同的分区会设置不同权限、资源限制等。
`AVAIL`	可用状态：`up` 可用；`down` 不可用
`TIMELIMIT`	该分区的作业最大运行时长限制，`30:00` 表示30分钟，如果是`2-00:00:00`表示2天，如果是`infinite`表示不限时间
`NODES`	节点数量
`STATE`	状态：`drain`: 排空状态，表示该类节点不再分配到其他；`idle`: 空闲状态；`alloc`: 被分配状态; `mix`:部分被占用，但是仍有可用资源； `down`停机
`NODELIST`	节点列表

2.squeue命令

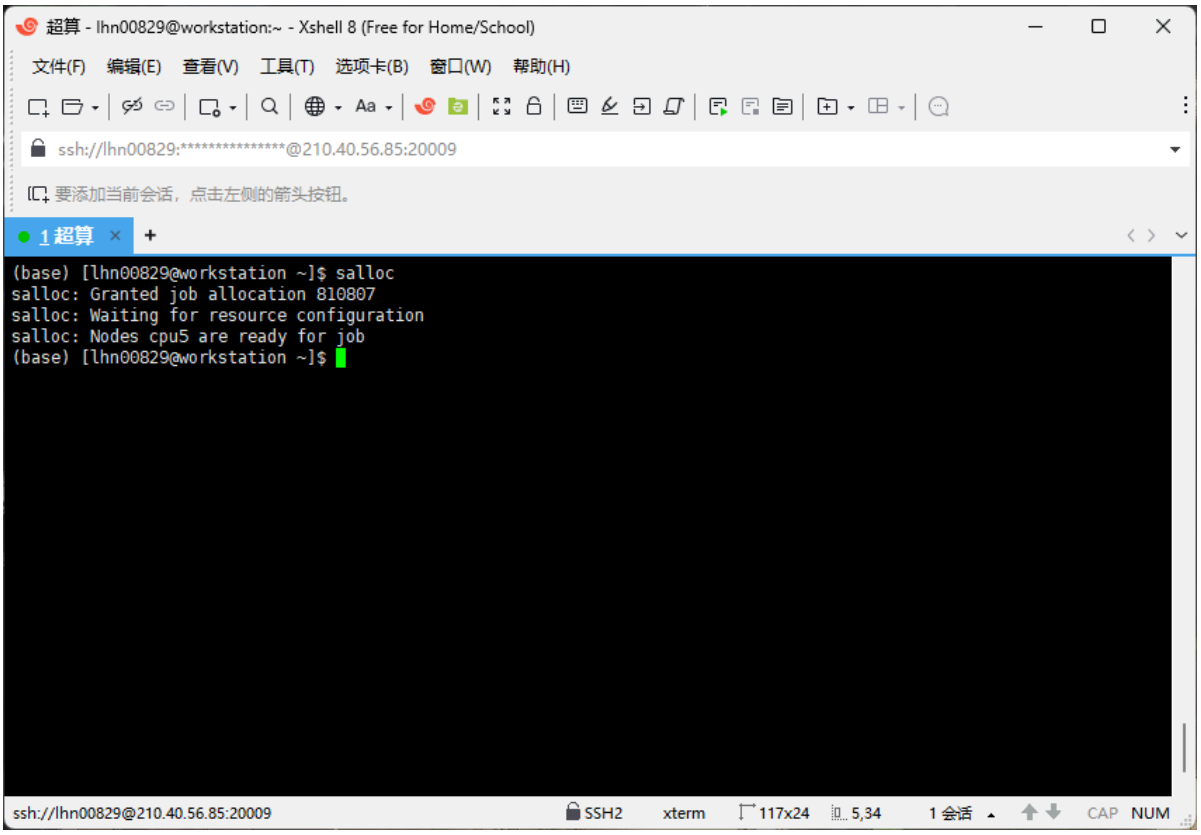


关键词	含义
`JOBID`	作业的id号，每个成功提交的任务都会有唯一的id
`PARTITION`	分区名
`NAME`	作业名称，默认为提交脚本的名称
`USER`	用户名，提交该作业的用户名
`ST`	作业状态：`PD`排队；`R`运行；`S`挂起；`CG`正在退出
`TIME`	作业运行时间
`NODES`	作业占节点数
`NODELIST(REASON)`	作业所占节点列表，如果是排队状态的任务，则会给出排队原因

3.简单演示

下面简单演示一下如何申请资源

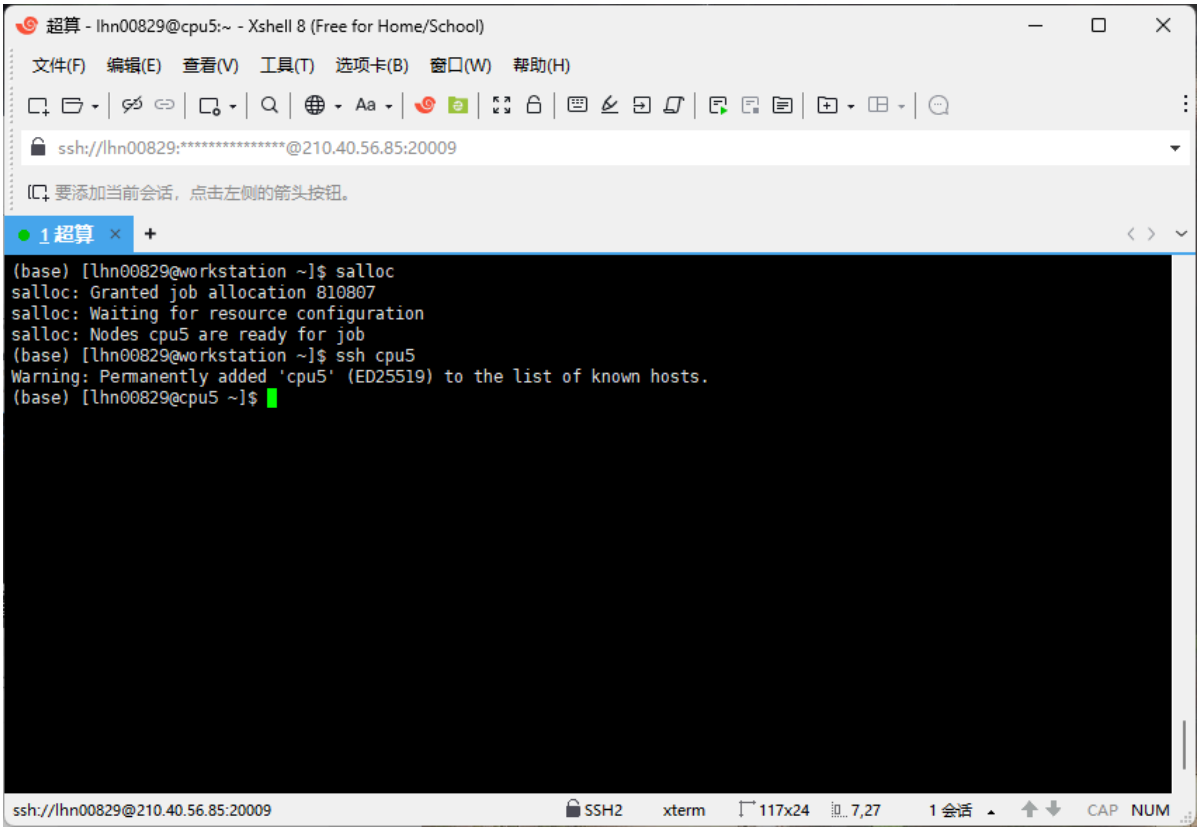
- 使用 salloc
  - 指定 **salloc --gres=gpu -p gpu-a10**
  - 不指定是随机分配



上图**cpu5**就是随机给你分配的计算资源

```
ssh cpu5
```

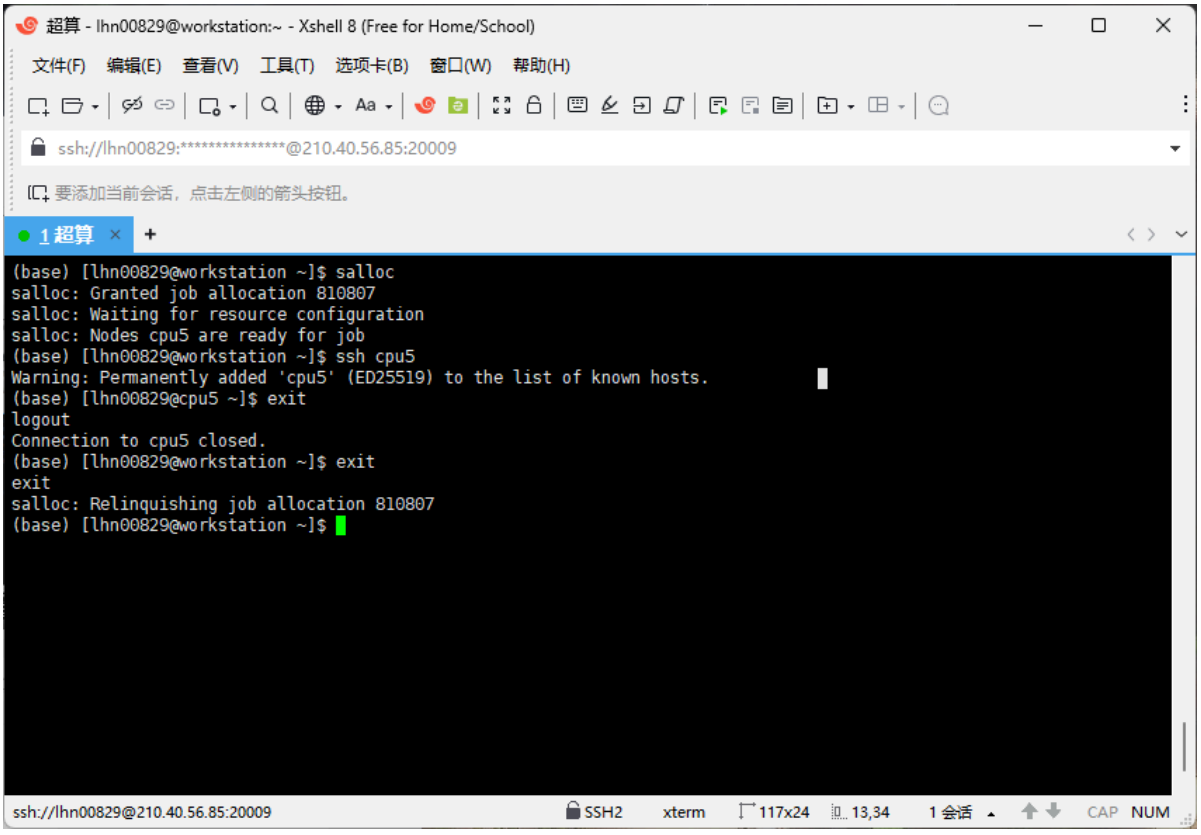
- 使用**ssh**连接这个cpu



如上图所示，你的**@workstation~**以及切换为**@cpu5~**

- 使用**exit**退出
  - 两次退出，第一次释放资源，第二次退出。

- 申请GPU资源的时候不使用不要占用。



- 申请GPU的脚本

```
- salloc --gres=gpu -p gpu-a10
-

 #! /bin/bash
### 表示这是一个bash脚本

#SBATCH --job-name=gpu-example
### 该作业的作业名

#SBATCH --nodes=1
### 该作业需要1个节点

#SBATCH --ntasks=16
### 该作业需要16个CPU

#SBATCH --gres=gpu:1
### 申请1块GPU卡

#SBATCH --comment project_name
### 指定从哪个项目扣费。如果没有这条参数，则从个人账户扣费

source ~/.bashrc
### 初始化环境变量

python test.py
### 程序的执行命令
```