

HPC FOR ZJU 常见问题

BY MY 2022

@ZJU

简介

关于IP地址：

- IP: 10.39.13.147
- Rstudio 采用默认端口：8787
- Jupyternotebook :8789

关于服务器的算力资源使用：

一般没有限制，服务器的可用资源可以通过 top 命令来查看。另外如果是短时间内（不超过半小时）需要大量内存和CPU，可以在群里提前告知下。大家可以错峰运行程序。

请大家提交任务后，多用 **top -u + 用户名** 查看后台资源使用情况

关于 Linux 基础：

如果没有学过Linux，请先自学 (<https://hpc.nih.gov/>)

关于服务器权限：

基本没限制权限。

关于 docker 服务：

服务器未配置 docker OR singularity等容器（你们几乎用不到），如有需要，自行安装。

Q1：服务器SSH登陆

1. ip 地址是什么？

IP地址是由一串数字和“.”号组成的，用来识别每台计算机的数字，也就是说每台电脑都会有一个唯一的IP地址。可以通过以下方法查询自己电脑的IP地址。

IPv4 地址: 169.254.4.74/16

1. 端口是什么?

每一台服务器都会提供 ssh 登录端口和 Rstudio-server 的登录端口

1. ssh是什么?

ssh是一种用于电脑与电脑之间加密传输信息的一种方式，类似于摩斯密码一样，将正常的语言进行加密，避免信息被破译。

具体内容可以参考：<https://mp.weixin.qq.com/s/hcU56Q1xgTlwVZKujQmiw>

服务器的登陆需要在电脑上下载Xshell或finalshell这两个软件(推荐后者)，Mac电脑登陆方式和Windows的又略有不同，下面就分开介绍。

1. mac电脑登录

Mac 电脑自带 ssh 协议，因此登陆的时直接打开 **终端 Terminal** 输入：

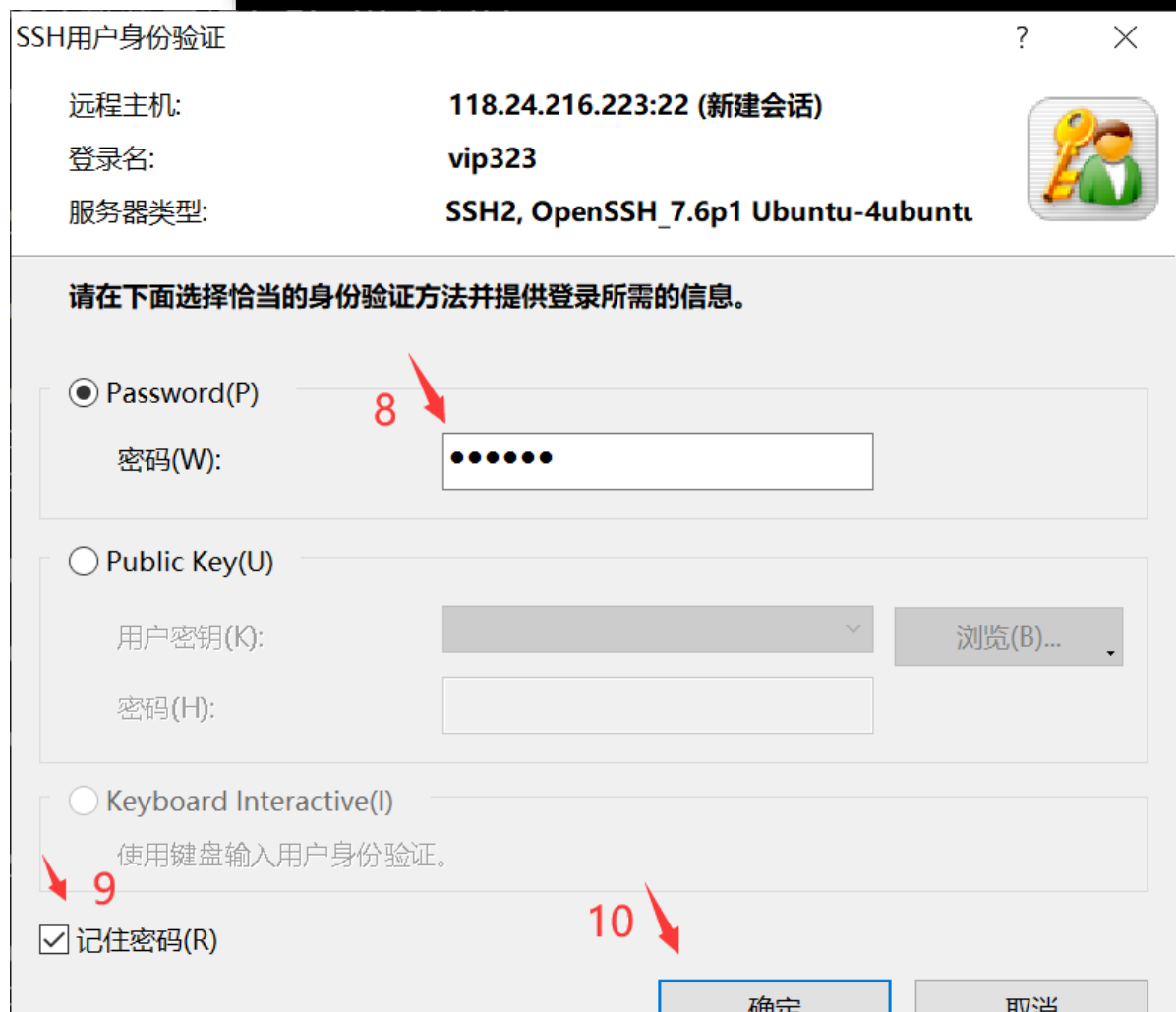
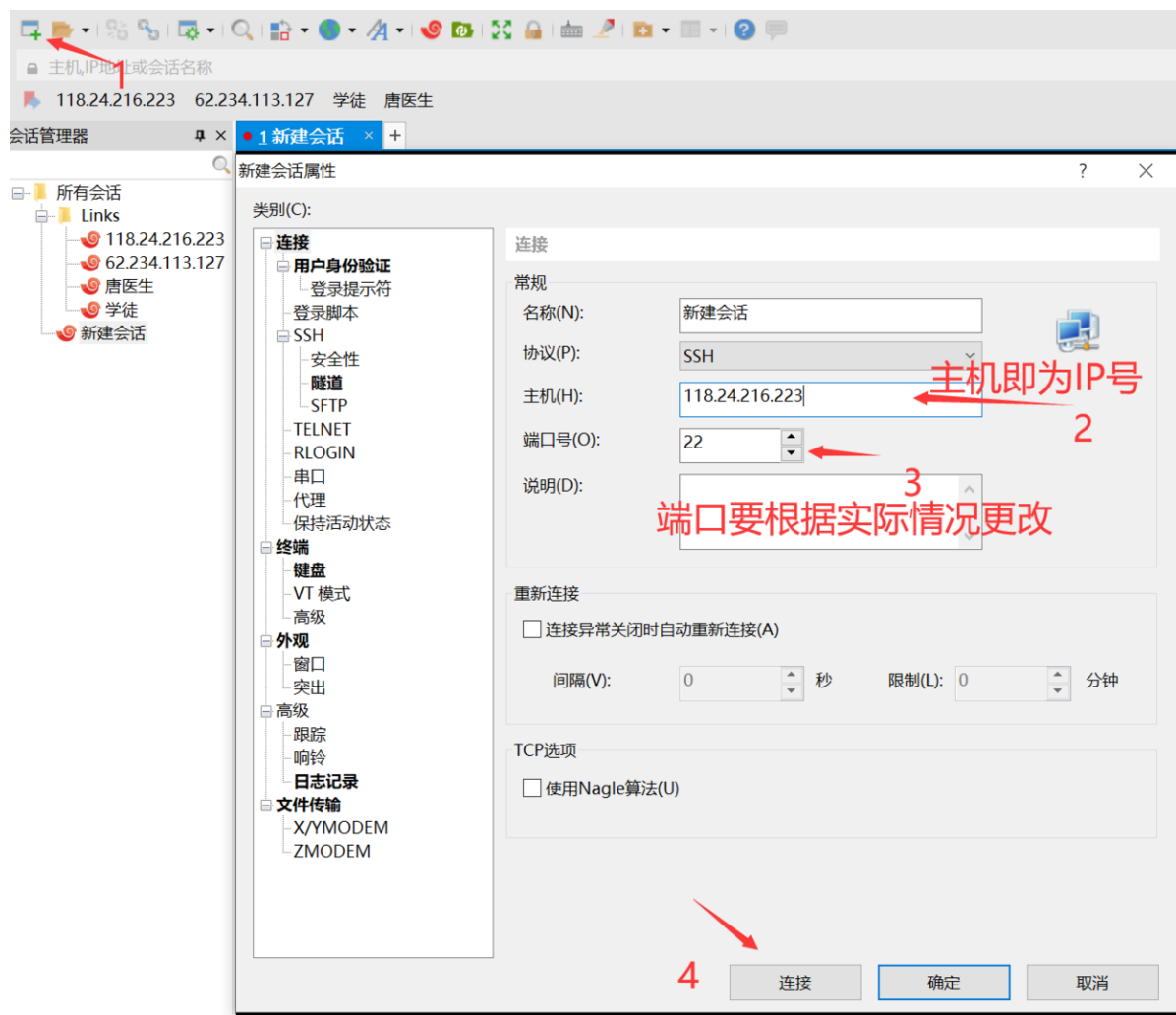
```
ssh 用户名@IP -p 端口号
```

1. Windows电脑登录 (Xshell或finalshell二选一)

5.1 Xshell 登陆

下载安装 xshell 软件：<https://www.netsarang.com/en/free-for-home-school/>

打开 xshell ，按下图操作：注：其中的 IP 地址已经更新为域名：**biotraineer.vip**



5.2 finalshell 登陆

下载安装 finalshell : [finalShell连接](#)

然后按这个教程操作: [如何登陆共享云服务器及Rstudio-server](#)

1. 修改密码


登录后可以使用 passwd 命令修改密码

Q2: Rstudio-server登陆


首先打开浏览器、再在地址栏输入 **IP:端口** (或者**localhost: 8787**) 然后在弹出的页面中输入用户名和密码即可



Sign in to RStudio

Username:  1


eeeeeeeeee

Password:  2

.....

You will automatically be signed out after 60 minutes of inactivity.

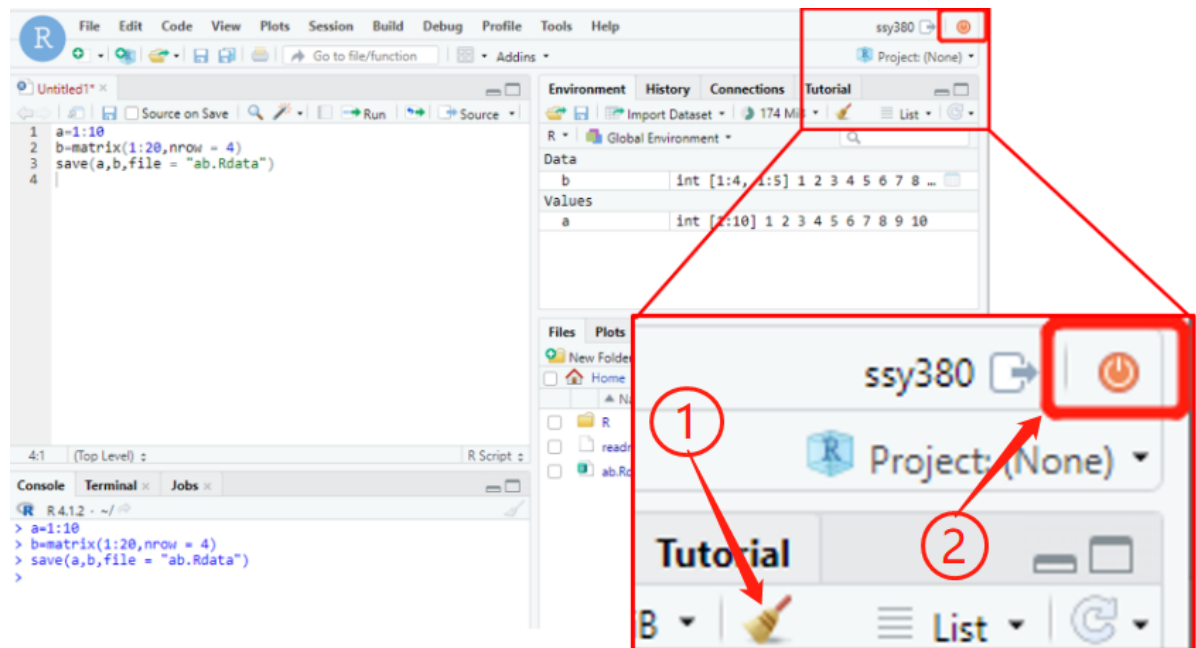
☐ Stay signed in when browser closes

 3

Sign In

Rstudio-server的退出

退出Rstudio-server最好**不要直接关闭网页**，否则有可能导致以后登陆Rstudio-server报错。在退出之前该保存保存好（save函数），然后清理掉变量（扫帚），最后退出（窗口右上角）即可



Q3: 文件上传和下载

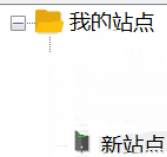
1. FileZilla

文件(F) 编辑(E) 查看(V) 传输(T) 服务器(S) 书签(B) 帮助(H) 有新版本! (N)



站点管理器

选择记录(S):



新站点(N)

新文件夹(F)

新建书签(M)

重命名(R)

删除(D)

复制(I)

常规

高级

传输设置

字符集

协议(T): SFTP - SSH File Transfer Protocol

主机(H): 118.12.23.789 端口(P): 2345

登录类型(L): 正常

用户(U): www

密码(W):

背景颜色(B) 无

注释(M):

连接(C)

确定(O)

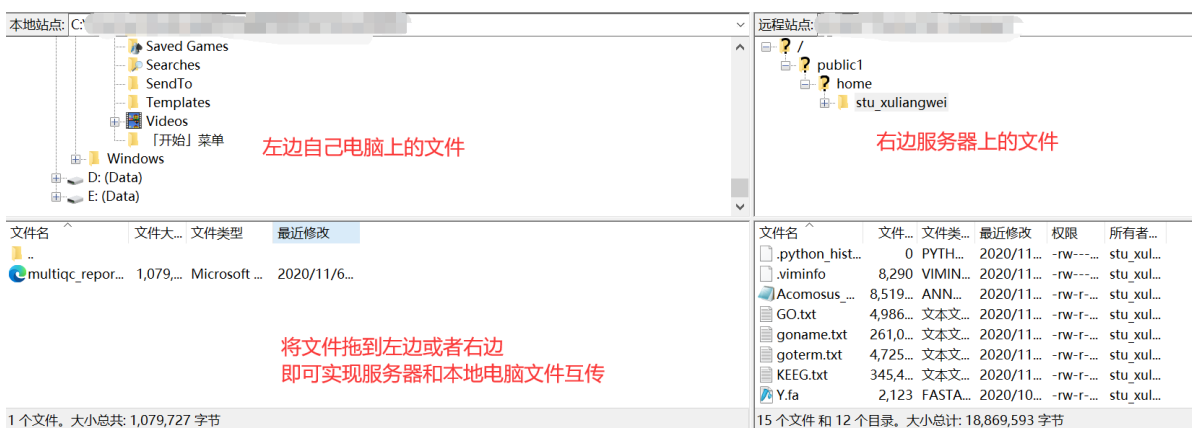
取消

命令输入

文件

命令

文件名	大小	类型	修改时间	权限	用户/用户组
/					
bin		文件夹	2020/09/25 08:28	drwx-----	
boot		文件夹	2020/10/23 13:08	drwxrwxr-x	
cdrom		文件夹	2020/09/29 16:00	drwxr-xr-x	
dev		文件夹	2020/10/22 20:15	drwx-----	
etc		文件夹	2020/11/18 10:53	drwxr-xr-x	
.gnupg		文件夹	2020/10/23 20:22	drwxrwxr-x	
.java		文件夹			
.local		文件夹			
.ncbi		文件夹			
.rstudio		文件夹			
database		文件夹			



1. Xftp

如果使用 xshell 可以登录服务器后双击下面图标打开 Xftp



1. FinalShell

使用FinalShell登陆服务器后，窗口最下方会显示目录结构，点击“↓”即可下载服务器中的文件；点击“↑”即可向服务器上传电脑中的文件

Q4：conda的安装和使用

绝大多数的软件可以通过conda安装，方法参考：<https://mp.weixin.qq.com/s/s5uuf7bTqZMuKvlssz-uWw>

conda 安装包请自行下载

Q5：服务器无法登陆？

1 登陆报错后请仔细查看邮件并比较登陆方法中的 **IP地址**、**用户名**、**端口**是否输错

2 Windows电脑使用Xshell登陆服务器显示以下情况请仔细阅读 服务器登陆方法 **Q1：服务器登陆**

```
[D:\~]$ ssh Aug13@62.234.113.127 -p 22
Invalid port number.
[D:\~]$
[D:\~]$
```

3 换一下网络环境，可以尝试改用其他网络环境，如手机热点。

4 以上三点错误均排除后使用Xshell仍无法登陆，可尝试换用 FinalShell 登陆

链接：https://pan.baidu.com/s/1gZCGUBctF_smdxd3tt16PQ 提取码：pja7

Q6: Rstudio-server无法登陆？ 几种解决方法

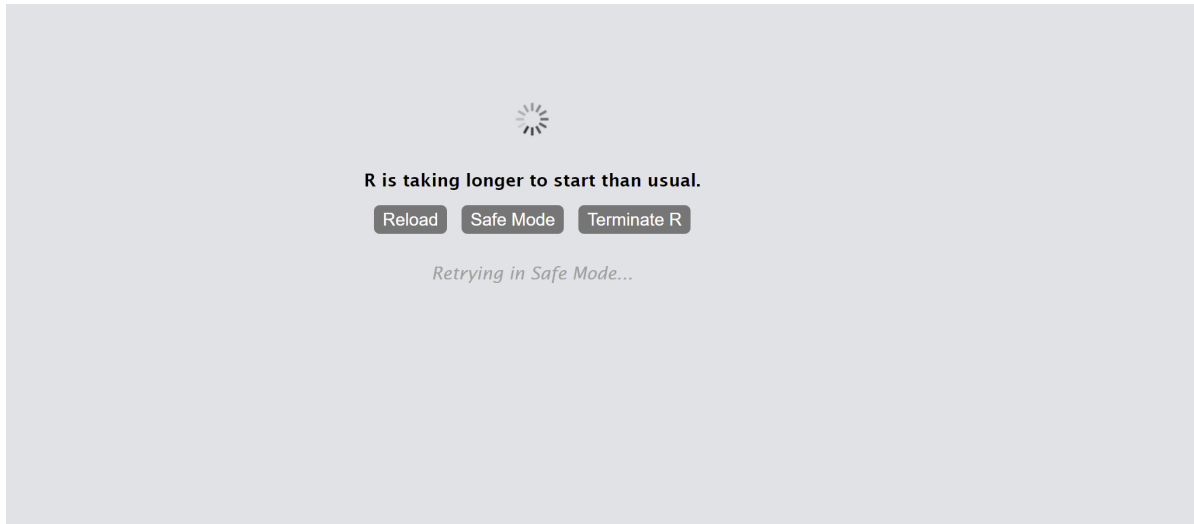
1 检查登陆方式是否正确（IP 端口是否弄错了）

2 查看自己用户名和密码是否输入错误

（大部分无法登陆的原因都是第1点和第2点）

3 清空浏览器cookie缓存，或者换个浏览器登陆，推荐谷歌浏览器

4 如果是出现下面截图界面，试试点击 **Terminate R**



5 ssh登录，然后查看自己的磁盘存储空间是否用尽（方法见Q9），如果是，先清理自己的磁盘空间。每个用户的默认磁盘空间是500G

6 关掉本地电脑的代理、科学上网工具、换个网络环境



7 可能是之前没有正确退出Rstudio，可以先ssh登录，然后kill掉自己的任务

```
ps -ef|grep ${USER} |grep rsession |awk '{print $2}'| xargs kill
```

8 当以上办法都没用时，就可以使用这个终极大招了

首先ssh登陆到服务器中，在家目录中找到并进入 rstudio 这个隐藏目录（有两个，逐一尝试），然后重命名为rstudio-old。最后重新登陆Rstudio-server即可

```
mv ~/.local/share/rstudio ~/.local/share/rstudio-old
mv ~/.rstudio ~/.rstudio-old
```

Q7: 如何调用服务器上的R包?

参考链接: <https://blog.csdn.net/yjiaobani/article/details/89497342>

这里涉及 `.libPaths()` 这个函数和 `.Rprofile` 这个文件。先看下 `.libPaths()` 的作用是什么, 这个函数作用的官方解释如上图。简单来说 `.libPaths()` 是用来搜索R包的路径。例如: 直接在 Rserver 中输入 `.libPaths()` 就会显示调用R包的路径 (如下图)

```
> .libPaths()
[1] "/home/data/liangwei/R/x86_64-pc-linux-gnu-library/4.0"
[2] "/usr/local/lib/R/site-library"
[3] "/usr/local/lib/R/library"
```

如果想在调用服务器中的R包有两种方法, 第一种方法需要每次登陆服务器中都设置一下; 第二种是一次设置长期有效。

1 直接在`.libPaths()`中输入服务器公共R包的路径即可。

结果如下图, 首先调用`ggplot2`, 报错说不存在这个R包。然后`.libPaths()`查看现有R包路径, 这两个路径是默认的, 里面没有安装过`ggplot2`, 所以报错也是应该的。接着`.libPaths("/home/data/refdir/Rlib")`设置公共目录R包的路径, 最后`library(ggplot2)`调用成功

```
> library(ggplot2)
Error in library(ggplot2) : there is no package called 'ggplot2'
> .libPaths()
[1] "/usr/local/lib/R/site-library" "/usr/local/lib/R/library"
> .libPaths("/home/data/refdir/Rlib")
> .libPaths()
[1] "/home/data/refdir/Rlib"          "/usr/local/lib/R/site-library"
[3] "/usr/local/lib/R/library"
> library(ggplot2)
>
```

2 登陆服务器创建一个`.Rprofile`然后在这个文件中输入你要调用R包的位置即可。

代码如下:

```
echo ".libPaths(c('/home/data/vipxx/R/x86_64-pc-linux-gnu-  
library/4.1','/home/data/refdir/Rlib','/usr/local/lib/R/library','/usr/local/lib  
/R/site-library'))" >> ~/.Rprofile
```

```
cat .Rprofile
```

```
$ cat .Rprofile  
.libPaths(c("/home/data/refdir/Rlib", "/home/liangwei/R/x86_64-p  
c-linux-gnu-library/4.0", "/usr/local/lib/R/library", "/usr/local  
/lib/R/site-library"))
```

注意：这里的 vipxx 要换成自己的用户名，另一点值得注意的是，创建了 .Rprofile 添加了路径后可能不会立即生效，可能需要过一天才行。

Q8：如何给自己的家目录设置权限？

用 chmod 命令，一般不需要修改自己家目录的权限，默认就是 700

```
chmod -R 700 用户家目录
```

Q9：如何查看自己的磁盘存储空间？

大家可以用 quota -uvs 命令查看自己的硬盘容量，默认是 500G，用 du -sh ~ 查看已使用的空间。

```
quota -uvs ${USER}  
du -sh ~
```

```
Disk quotas for user vip12 (uid 5015):  
  Filesystem  space  quota  limit  grace  files  quota  limit  grace  
/dev/nvme0n1p2 105M    0K     0K           80      0      0
```

第一列space表示你现在用了多少存储空间，而不是剩余多少空间。

第二列quota表示服务器对你软限制空间是多少。

第三列limit表示服务器对你的硬限制是多少。

第四列grace time表示当你使用的内存介于软限制和硬限制之间时，服务器对你的宽限时间，即在这段限制的时间内你需要管理你的存储空间，使其容量低于软限制。

第五列files表示你现有的文件个数。

第六列quota表示服务器对你文件个数的软限制。

第七列limit表示服务器对你文件个数的硬限制。

Q10: conda install 无法安装软件?

```
CondaHTTPError: HTTP 000 CONNECTION FAILED for url
<https://mirrors.ustc.edu.cn/anaconda/cloud/linux-64/current_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.ustc.edu.cn/anaconda/cloud/linux-64'
```

网络较慢，换个时间试试

或者换个镜像试试（以下3个镜像【不要用清华镜像】选择一个即可）：

```
# 下面四行配置北京外国语大学的conda的channel地址
conda config --add channels https://mirrors.bfsu.edu.cn/anaconda/pkgs/main/
conda config --add channels https://mirrors.bfsu.edu.cn/anaconda/cloud/conda-
forge/
conda config --add channels https://mirrors.bfsu.edu.cn/anaconda/cloud/bioconda/

conda config --set show_channel_urls yes


# 下面配置阿里云的conda的channel地址
conda config --add channels https://mirrors.aliyun.com/anaconda/cloud/msys2
conda config --add channels https://mirrors.aliyun.com/anaconda/cloud/bioconda
conda config --add channels https://mirrors.aliyun.com/anaconda/pkgs/main
conda config --add channels https://mirrors.aliyun.com/anaconda/pkgs/r
conda config --set show_channel_urls yes
```

配置好镜像之后，清空一下环境中的缓存：

```
conda clean -i
```

如果还是不行，试试把 ~/.condarc 中的 https 改成 http

或者删除conda后重新下载一个

如果是出现下面截图报错，则 删除掉 .condarc 中的 defaults

```
(rna) hcguo 13:40:23 ~
$ cat .condarc
channels:
- https://mirrors.aliyun.com/anaconda/pkg/r
- https://mirrors.aliyun.com/anaconda/pkg/main
- https://mirrors.aliyun.com/anaconda/cloud/bioconda
- https://mirrors.aliyun.com/anaconda/cloud/msys2
- https://mirrors.aliyun.com/anaconda/cloud/conda-forge
- defaults
show_channel_urls: true
(rna) hcguo 13:40:28 ~
$ conda install hisat2
Collecting package metadata (current_repodata.json): failed

CondaHTTPError: HTTP 000 CONNECTION FAILED for url <https://repo.anaconda.com/pkg/main/linux-64/current_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.

If your current network has https://www.anaconda.com blocked, please file
a support request with your network engineering team.

'https://repo.anaconda.com/pkg/main/linux-64'

(rna) hcguo 13:43:02 ~
$ vim .condarc
(rna) hcguo 13:43:18 ~
$ cat .condarc
channels:
- https://mirrors.aliyun.com/anaconda/pkg/r
- https://mirrors.aliyun.com/anaconda/pkg/main
- https://mirrors.aliyun.com/anaconda/cloud/bioconda
- https://mirrors.aliyun.com/anaconda/cloud/msys2
- https://mirrors.aliyun.com/anaconda/cloud/conda-forge
show_channel_urls: true
(rna) hcguo 13:43:21 ~
$ conda install hisat2
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

environment location: /home/hcguo/miniconda3/envs/rna

added / updated specs:
- hisat2

The following packages will be downloaded:

package | build
-----|-----
ca-certificates-2022.2.1 | h06a4308_0 122 KB https://mirrors.aliyun.com/anaconda/pk
```

有 defaults

删除掉 defaults

Q11：关闭电脑后怎么继续在服务器上运行程序？

<https://mp.weixin.qq.com/s/DbzYmGSrINbyZcEOl-qQPg>

后台运行 nohup . . . &

Q12: Rserver中怎么调用自己电脑中的数据?

请问用rstudio的话是也需要把自己要处理的文件用finalshell上传吗? setwd应该怎么设置在哪里? 晕晕的[苦涩]

这里面有3个关键词: 服务器、自己的数据、Rserver。

服务器本身就是一台电脑, 只不过是一台没有显示器但内存和存储空间十分大的且十分笨重的一台大型电脑。

Rserver是搭建在服务器上的类似于Rstudio一样的编程界面, Rserver的存在就像Rstudio一样方便在服务器中编程。

因此, 服务器就相当于个人的电脑, Rserver就好比个人电脑上的Rstudio。所以在Rserver中设置工作目录和在Rstudio中没差别, 需要调用自己电脑中的数据就把数据上传到服务器中即可。

图一是服务器中显示家目录中的文件, 图二是Rserver中显示工作目录下的文件。

```
$ ls
backup
database
families.txt
GCF_001540865.1_ASM154086v1_genomic.fna.gz
Homo_sapiens.GRCh38.cdna.all.fa.gz
index.html
miniconda3
nohup.out
num2.txt
num.sh
num.txt
project
R
result.txt
samtools-1.9.tar.bz2
test.sh
liangwei 22:19:55 ~
$ pwd
/home/liangwei
```



```
> dir(".")
[1] "backup"
[2] "database"
[3] "families.txt"
[4] "GCF_001540865.1_ASM154086v1_genomic.fna.gz"
[5] "Homo_sapiens.GRCh38.cdna.all.fa.gz"
[6] "index.html"
[7] "miniconda3"
[8] "nohup.out"
[9] "num.sh"
[10] "num.txt"
[11] "num2.txt"
[12] "project"
[13] "R"
[14] "result.txt"
[15] "samtools-1.9.tar.bz2"
[16] "test.sh"
> getwd()
[1] "/home/data/liangwei"
> setdir("database/genome/")
Error in setdir("database/genome/") : could not fi
> setwd("database/genome/")
> getwd()
[1] "/home/data/liangwei/database/genome"
```



Q13: 怎么查看自己家目录使用了多少存储空间

直接在服务器中输入 ncdu

```
--- /home/data/liangwei -----
229.2 GiB [#####] /project
  5.3 GiB [          ] /miniconda3
114.5 MiB [          ] GCF_001540865.1_ASM154086v1_genomic.fna.gz
 24.6 MiB [          ] /R
 16.7 MiB [          ] /database
 13.6 MiB [          ] /.rstudio
 11.2 MiB [          ] Homo_sapiens.GRCh38.cdna.all.fa.gz
   7.4 MiB [          ] /.cache
   6.4 MiB [          ] families.txt
   4.2 MiB [          ] samtools-1.9.tar.bz2
292.0 KiB [          ] /.local
 88.0 KiB [          ] /.java
 64.0 KiB [          ] /.config
 40.0 KiB [          ] .bash_history
 12.0 KiB [          ] .viminfo
 12.0 KiB [          ] /.conda
   8.0 KiB [          ] /groupg
   8.0 KiB [          ] /.ncbi
Total disk usage: 234.7 GiB Apparent size: 234.7 GiB Items: 158801
```

Q14：如和查看自己是否在docker组中

通常来说新用户是不在docker组中，要加入docker组需要@唐医生说明自己要加入docker组。然后通过 **docker ps** 命令查看自己是否加入了

```
$ docker ps
Got permission denied while trying to connect to the Docker daemon socket at unix:///var/run/docker.sock: Get http://%2Fvar%2Frun%2Fdocker.sock/v1.24/containers/json: dial unix /var/run/docker.sock: connect: permission denied
```

以上显示是未加入docker组

```
vip48@tpm7-desktop:~$ docker ps
```

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
--------------	-------	---------	---------	--------	-------	-------

以上显示是已经成功加入了docker组

Q15：filezilla 为啥链接不上去

红字说的很清楚，把FTP协议改成SFTP协议就可以了



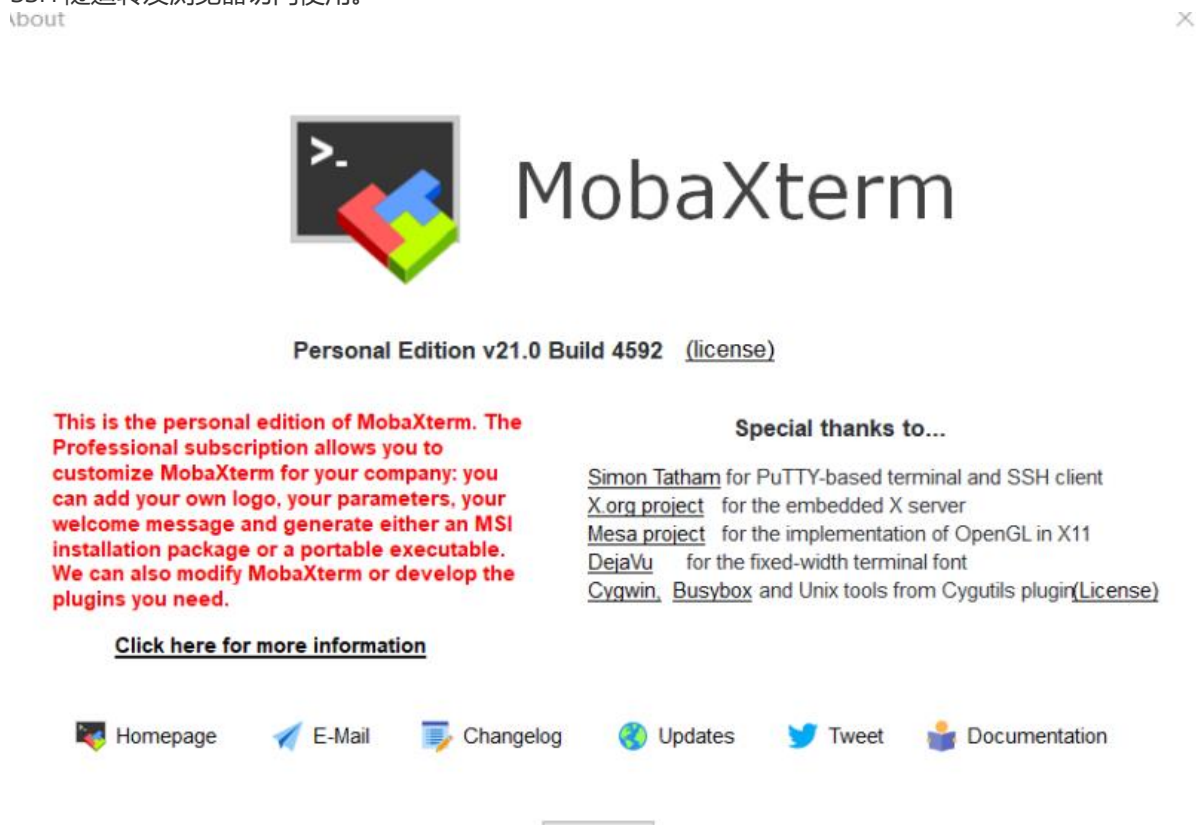
Q16: 遇到别的错误怎么办?

微信搜索、必应谷歌搜索等搜索

Q17: 安装 jupyter

服务器已经配置了 Rstudio-server 和 Jupyterhub。服务器只配置了 Rstudio-server，没有配置 jupyter，如果有需要，请使用普通用户的安装方法，然后进行隧道转发，一个参考的隧道转发方法是：

需要使用服务器 jupyter notebook 的用户，请自己下载 MobaXterm 客户端。不需要做什么设置即可 SSH 隧道转发浏览器访问使用。



xshell putty finalshell 等等其实也可以 就是需要设置ssh tunnel（隧道）转发。网上也有相关教程。设置相对繁琐点。

Q18：如何看自己使用的资源有没有超过【48 CPU或200G内存】？

如果是在 Linux 提交任务，一般命令会有线程或核心、内存相关的设置参数，查看其帮助文档了解相关参数。可以使用 `top -u ${USER}` 或者 `htop -u ${USER}` 查看一下**CPU 列**和**MEM 列**的信息

Q19：conda、pip 或者 R语言 镜像

服务器在国内，在使用 conda、pip 或者 R语言安装包的时候，需要设置国内镜像。

（不要用清华）（不要用清华）（不要用清华）

【1】conda镜像：

在 Q10 已经说得很详细了。

【2】pip 镜像：

阿里云： <http://mirrors.aliyun.com/pypi/simple/>
中国科技大学 <https://pypi.mirrors.ustc.edu.cn/simple/>
华中理工大学： <http://pypi.hustunique.com/>
山东理工大学： <http://pypi.sdutlinux.org/>
豆瓣： <http://pypi.douban.com/simple/>

示例：

```
pip install cellphonedb -i https://mirrors.aliyun.com/pypi/simple/
```

【3】R语言镜像：

CRAN： <https://cran.r-project.org/mirrors.html>

```
# 中科大
options("repos" =c(CRAN="https://mirrors.ustc.edu.cn/CRAN/"))
# 南京大学
options("repos" =c(CRAN="https://mirrors.nju.edu.cn/CRAN/"))
# 腾讯云
options("repos" =c(CRAN="http://mirrors.cloud.tencent.com/CRAN/"))
# 阿里云
options("repos" =c(CRAN="https://mirrors.aliyun.com/CRAN/"))
```

Bioconductor: <https://www.bioconductor.org/about/mirrors/>

```
# 中科大
options(Bioc_mirror = "https://mirrors.ustc.edu.cn/bioc/")
# 南京大学
options(Bioc_mirror = "https://mirrors.nju.edu.cn/bioconductor/")
```

Q20: 安装 github 包

如果想在服务器上安装 github 上面的包，请先安装 y ulab.utils 并了解这个包的用法