

中央财经大学高性能服务器使用指南

(试行)

申请高性能服务器账户

- 高性能服务器账户的老师需要给李丰<feng.li@cufe.edu.cn>发邮件申请，请提供以下信息
 - 姓名：
 - 工号：
 - 电子邮箱：
 - 所在部门：
 - 手机号码：
 - 使用用途：（科研项目需要提供项目名称和编号）
- 暂时不为学生提供账户。

高性能服务器软硬件配置简介

- 基本硬件
 - 一台登录节点: `nod-0`, 两颗 8 核 CPU, 128G 内存
 - 四台 I/O 节点提供分布式文件存储
 - 五台高性能计算节点: 每台两颗 8 核 CPU, 128G 内存
- 软件部署
 - 登录节点和计算节点安装 Debian Linux 8 版本, 无图形界面
 - 已部署软件
 - 常用统计和计算软件: Python, R, Open MPI, BLAS, GCC, Julia
 - 数据库软件: MySQL
 - 大数据分布式计算框架: Hadoop, Hbase, Hive, Mahout, Spark
 - 并行计算排队以及资源分配作业系统: Slurm
- 访问方式
 - 校内 IP 通过 SSH 访问
 - 校外 IP 通过高性能共享桌面 Putty 程序 SSH 访问

登录到高性能服务器

- 中央财经大学校内登录

服务器采用 ssh 方式登录。服务器校内地址为 192.168.113.164，端口为 22。服务器不对校外开放。

- Windows 用户请安装 SSH 程序 Putty (www.putty.org) 和文件传输程序 FileZilla (filezilla-project.org)。
- Linux 和 Mac 用户直接在终端输入

```
ssh yourusername@192.168.113.164
```

- 中央财经大学校外登录
 - 请联系中央财经大学教学技术服务中心申请高性能计算共享桌面，从共享桌面 Putty 程序可以登录到高性能服务器。

新建一个 BATCH 任务

- Windows 用户需要利用 FileZilla 上传你的已经调试好的程序（假设 `myRprogram.R`）、相应数据到服务器的某个文件夹内（假设该文件夹名为 `myProgram`）。Linux 用户可以使用 `scp` 或者 `rsync` 命令上传，请参考具体命令使用手册。
- 如果你使用 R 软件中非自带的 R 包，应该首先登录到服务器，启动 R 程序，安装所需 R 包。

```
ssh yourusername@192.168.113.164
yourusername@nod-0:~$ R
> install.packages("mvtnorm")
> q("no")
```

- 在服务器上代码所在的相同文件夹内建一个以 `.sh` 为扩展名的文件，这里假设为

`slurm.sh` (为了方便起见, `slurm.sh` 应该和你的程序 `myRprogram.R` 在 `myProgram` 文件夹下)

- 切换到 `myProgram` 文件夹内

```
yourusername@nod-0:~$ cd myProgram/
```

- 用文本编辑器创建一个 `slurm.sh` 文件

```
yourusername@nod-0:~/myProgram$ nano slurm.sh
```

- 编辑该文件如下所示 (注意所有的空格和短线。替换 `yourRprogram` 为你所对应的程序名称。替换 `yourname@mail.com` 为你自己的 Email 地址)

```
#!/bin/bash -l
#SBATCH -n 1 # Number of cores
#SBATCH -t 1-00:00 # Runtime in D-HH:MM
#SBATCH -o yourRprogram.out # Output File
#SBATCH -e yourRprogram.err # Error File
#SBATCH --mail-type=ALL
#SBATCH --mail-user=yourname@mail.com # Your email address
srun R CMD BATCH yourRprogram.R
```

- 按 **Ctrl+X** 保存并退出编辑
- 注意事项:
 - 程序运行成功或者中途失败, 系统会发邮件给你指定的邮件地址。
 - 如果你的程序中途停止, 你应该查看你的 `yourRprogram.err` 文件
 - 程序运行成功结果输出会存在 `yourRprogram.out` 文件。
 - 如果你不做并行计算, 在 `slurm.sh` 文件指定一个 Core 就够了:
 - `#SBATCH -n 1 # Number of cores`

- 你需要指定本次任务系统最多需要运行多长时间。如果你运行超过你的指定时间，系统会杀死你的程序。

- `#SBATCH -t 1-00:00 # Runtime in D-HH:MM`

提交 BATCH 任务

- 如果你提交一个任务，请在 `nod-0` 下的 `myProgram` 文件夹下运行以下命令

```
yourusername@nod-0:~$ cd myProgram/  
yourusername@nod-0:~/myProgram$ sbatch slurm.sh
```

- 查看已经提交的任务以及 `JOBID`

```
yourusername@nod-0:~$ squeue -u yourusername
```

- 如果你希望取消一个已经提交的任务

```
yourusername@nod-0:~$ scancel JOBID
```

- 退出登录节点

```
yourusername@nod-0:~$ exit
```

- 注意事项

- 登录节点 `nod-0` 主要是用来维护和提交任务的，不能在上面直接做计算。所有的任务都要通过 `sbatch` 命令来提交。
- 你提交完任务你就可以退出系统了。程序运行结束系统会邮件通知你，你也可以同时提交多个 **BATCH** 任务，系统会根据计算机资源排队安排计算的。

获取帮助

- Linux 基本命令:
 - <http://sourceforge.net/projects/linuxcommand/files/TLCL/13.07/TLCL-13.07.pdf/download>
- SLURM 任务管理系统:
 - <http://slurm.schedmd.com/sbatch.html>
- 李丰 <feng.li@cufe.edu.cn>