

HABANERO 'tutorial' — Feb 28 2020

BEFORE GOING THROUGH THIS TUTORIAL, READ THE FILE *'main steps to interact with Habanero.rtf'*

In what context do I use the HABANERO HPC cluster?

Most of the times because I want to run a script (on R) that is slowing down my computer extremely, or that my computer just can't handle.

So that's the usage this tutorial is going over: running an R script. Below are the SSH commands corresponding to the different steps.

To follow this tutorial you will need the files:

- `test.R` : the R script we want to run on the cluster
- `test.sh` : the corresponding .sh file (containing the cluster resources we're asking for, and then calling `test.R`)
- `test_input.csv` : the input dataset my R script is using
- `install_Rpackages_HABANERO.R` : optional (see below)

Steps:

- 0. Preliminary: creating your free account
- 1. Accessing your directory on the server
- 2. Writing the .R script and .sh file
- 3. Installing R packages 'locally' (in my home directory on the server)
- 4. Transferring my dataset & .R & .sh files to the server
- 5. Running it, opening files with nano to look at the result
- 6. Transferring my output dataset to my local directory

All the corresponding SSH commands:

- 0. Preliminary: creating your free account

You must have requested to have a free account created for you on the cluster. This is done by submitting a request form (see *'main steps to interact with Habanero.rtf'*). Once created, the path to this account will be: [/rigel/home/<UNI>](#)

- 1. Accessing your directory on the server

```
$ ssh <UNI>@habanero.rcs.columbia.edu
```

```
$ type my Columbia password
```

```
$ pwd
```

```
$ mkdir input
```

```
$ mkdir output
```

```
$ mkdir rpackages
```

```
$ ls
```

```
$ exit
```

Check what is the “present working directory” according to the Terminal (for me, on a Mac, it is by default ‘/Users/<myusername>’). You want to have all the files you’re going to upload to the server in that directory. So eventually change the working directory to ‘/Users/<myname>/Documents’, and put your files there.

```
$ pwd
```

```
$ cd Documents
```

— 2. Writing the .R script and .sh file

Open **test.R** and **test.sh**. In both, replace all the “<UNI>” with your own personal <UNI> Save both in your local directory

— 3. Installing R packages 'locally' (in my home directory on the server)

EITHER

```
$ scp install_Rpackages_HABANERO.R
```

```
<UNI>@habanero.rcs.columbia.edu:/rigel/home/<UNI>
```

```
$ ssh <UNI>@habanero.rcs.columbia.edu
```

```
$ Rscript install_Rpackages_HABANERO.R
```

OR

```
$ ssh <UNI>@habanero.rcs.columbia.edu
```

```
$ module R
```

```
$ R
```

```
install.packages("hms", lib = "/rigel/home/<UNI>/rpackages/")
```

```
$ q()
```

```
$ cd rpackages
```

```
$ ls
```

```
$ exit
```

— 4. Transferring my dataset & .R & .sh files to the server

```
$ scp test.R test.sh <UNI>@habanero.rcs.columbia.edu:/rigel/home/<UNI>  
$ scp test_input.csv <UNI>@habanero.rcs.columbia.edu:/rigel/home/<UNI>/input
```

— 5. Running it, opening files with nano to look at the result

```
$ ssh <UNI>@habanero.rcs.columbia.edu  
$ sbatch test.sh  
(You should be receiving emails saying your job has begun and then that your job has ended)  
$ ls  
$ nano routput  
  ^X (to exit the file)
```

— 6. Transferring my output dataset to my local directory

```
$ exit  
$ scp <UNI>@habanero.rcs.columbia.edu:/rigel/home/<UNI>/output/<myFile.ext>  
/Users/<myusername>/Documents/
```