Computing on Clark and More Parallel Computing

J.D. Peiffer

2/7/2021

Section 1: Connecting to Clark with mobaXterm

1. Visit https://missouri.app.box.com/v/rcss-mobaxterm-new and download both files into your downloads folder.
2. Run the .exe file, accept all defaults for installation.
3. Open mobaXterm and click "Start Local Terminal" This will open a command interface for you to connect to Clark with.
4. Generate a ssh key:

    i) In your terminal type `ssh-keygen`.
    ii) Accept all defaults by pushing enter. Do not enter a password.
    iii) You only need to do this once!

5. Connect to Clark

    i) In your terminal, type `ssh <username>@clark.rnet.missouri.edu`
    Replace `<username>` with your pawprint.

    ```
     11/02/2021    13:47.59    /home/mobaxterm   ssh jdp6n8@clark.rnet.missouri.edu
    Enter passphrase for key '/home/mobaxterm/.ssh/id_rsa':
    ```

    ii) Do not include brackets, for example `ssh jdp6n8@clark.rnet.missouri.edu`
    iii) When prompted for your password, enter your university password. **NOTE THAT WHAT YOU TYPE DOES NOT SHOW UP IN THE TERMINAL** You should then see something like this:

    ```
    * *********************************************************************
    * Welcome to clark-r630-login-node907.cobalt.lan, CentOS, 7.9.2009
    * Archive of news is available in /etc/motd-archive
    * University of Missouri Division of IT Research Computing Support Services
    *
    * Use of this system is governed by the rules and regulations of the
    * University of Missouri and the University of Missouri System and also
    * by requirements of various granting agencies.
    *
    * Users must be familiar with and abide by the UM System acceptable use
    * policy (CRR 110.005) and the UM System Data Classification System (DCL).
    * http://umsystem.edu/ums/rules/collected_rules/facilities/ch110
    * http://umsystem.edu/ums/is/infosec/classification

    * Helpful Links
    * News: http://doit.missouri.edu/research/research-computing-news
    * Support: rcss-support@missouri.edu
    * Announcement list: https://po.missouri.edu/cgi-bin/wa?A0=rcss-announce-l
    * Training: http://docs.rnet.missouri.edu/training
    * Documentation: http://docs.rnet.missouri.edu/

    * *********************************************************************
    [jdp6n8@clark-r630-login-node907 ~]$
    ```

Section 2: Linux commands on Clark

Need to add data copy

Watch **this video** if you are not familiar with bash/linux commands. These will be useful in mobaXterm and Clark.

Question 2.1 In your own words, provide a description of what each of the following bash terminal commands do (google is your friend):
- `pwd`
- `ls`
- `cd`
- `touch`
- `rm`
- `cat`
- `cp`
- `mv`
- `mkdir`

Question 2.2 In your home directory, create a folder named `hw4` and in it create a file name `test.txt`. Screenshot the output of `ls` in this directory (Hint: windows users hold `windows button + shift + s`)

When you log into Clark, you are logging into part of a computer cluster called a *node*. Every time, you start out in the *login node*. Here, you can perform basic tasks like making folders and files, but more data intensive tasks, like running R or NCL, must be done on a compute node.

To request a node for this task, type this into your terminal: `srun -n 1 --mem 4G --pty /bin/bash`

Notice the node name at your command line changes!

To load the rstudio software, type `module load rstudio/1.1.456`
Next, type `rstudio`

Section 3: Load and Characterize Data

Make sure your working directory contains the file "allprcp.Rdata".

Load the variable into your workspace with

Question 3.1 What are the dimensions of this dataset? What is represented on the x and y dimension? (Hint: similiar to hw3).
Question 3.2 What is the beginning and end date of this time series?
Question 3.3 What is the column range of the data? (Names begin with US....) Use this information to make a new data frame called rain.

```
data_start_column = #Enter your answer here
data_stop_column = #Enter your answer here

rain = allprcp[, data_start_column:data_stop_column]
```

Question 3.4 How many missing data enteies are there in the entirety of this dataset? (Hint: `is.na()`)