Instructions for running on the RIS from a Mac for A&S faculty (ADM)

Overview

There are several ways to interact with the RIS infrastructure ranging in difficulty from running it in a browser to interacting with it directly through terminal commands in Linux. This document outlines the simplest approach to interacting with these resources called Open OnDemand.

Advantages:

- *Ease of use*. Once you get it started, it looks just like a standard application on your desktop computer.
- *No Docker*. These technical issues have been handled in the background.

Disadvantages:

- *Limited resources*. For instance, in the R example below the maximum number of CPUs we can allocate is 4.
- *Limited scope*. The applications currently running are Jupyter Notebook, Matlab, RStudio, Rellion3, and Stata. There is also a Compute RIS Desktop application. Note that even within these apps there are limitations. For instance, the RStudio App we use below is only available for R versions 4.0.3 and 4.0.2

To use this resource from a Mac, you perform the following steps.

1. Log into the RIS using ssh through the terminal.
2. Navigate to https://compute1-ood-1.ris.wustl.edu/pun/sys/dashboard and choose RStudio.
3. Fill out the form to request resources and specify your configuration and queue.
4. Start your instance on the RIS resources.
5. Move any files you need to your storage on RIS. This assumes you have already mapped your storage on the cluster to your computer.
6. Start running things on the RIS!

The document below provides additional details for each step.

Preliminiaries

Getting an allocation on the RIS. If you do not already have access to the RIS, fill out the form by clicking here.

Just fill the the form and ask for access to the cluster. Be SURE to note you are A&S faculty and are requesting access to the A&S condo. All A&S faculty should be allowed to have access to this resource. Be aware that it may take a day or two for this process to be completed.

Mapping your storage on the RIS Instructions on how to map the drive to your computer should have been sent to you when you first given access to the cluster. A summary of them is as follows:

- In the Finder app, use the Go menu and then select Connect to Server.

- In the dialogue, enter the server address. Part of this address will include your uniquely identifying group for the cluster. So mine would be smb://storage1.ris.wustl.edu/jacob.montgomery. If your name is not Jacob Montgomery then you will need to change that last part.
- You may be prompted to confirm you are connecting to storage1.ris.wustl.edu. Press "Connect" to continue.
- You will be prompted for a password. Be sure to use your WUSTL key and password.
- Click "Connect". From now on, this folder should be visible through the Finder app.

Step 1) Log into the RIS

The next step is to log into the RIS through the terminal.

1. Open the terminal app. This can be done by searching for "Terminal" in Spotlight or by launching the Terminal application from the **Applications > Utilities folder**.
2. Connect to the RIS using the following

```
ssh username@compute1-client-1.ris.wustl.edu
```

```
ssh jacob.montgomery@compute1-client-1.ris.wustl.edu
```

Again, if your name is not Jacob Montgomery you are going to need to change that first part to include your RIS username.

3. You should now be able to log in using your WUSTL key username and password.

Step 2) Open OnDemand

Navigate to https://compute1-ood-1.ris.wustl.edu/pun/sys/dashboard

On the left hand side, choose RStudio from the list of options.

Step 3) Fill in the form.

You will now see a form where you need to include information to configure your session. Most of these are obvious, but some are not. We will walk you through all of them just in case.

Mounts

This is a field that connects your session to your storage. You will need this later when you set your working directory in R.

The field takes two arguments arg1: arg2. The first argument is the path for your storage. The second is what you want to *call* the path for your storage in your session. we recommend that you use the same path structure for both arguments to keep from getting confused. Here is mine:

/storage1/fs1/jacob.montgomery:/storage1/fs1/jacob.montgomery

Once again, you will need to substitute in your own username to replace mine.

Version

Choose the version of R you want to run.

Job Group

If you're logged in, this should automatically populate with your information. By default, it will display as "\jacob.montgomery\ood". However, you can modify it if you prefer to group your jobs differently. If you're not doing anything too complicated, it's best to leave it as is.

User Group

All A&S faculty should put in `compute-artsci` to use the dedicated condo.

Queue

When you submit a job, you'll need to specify the queue you want it to be in. Essentially, your task will be placed in a line for execution, and it will run as resources become available.

A&S faculty members have two options: artsci or artsci-interactive. The main difference between these queues is the maximum amount of time allotted for job execution (as well as the relative busyness of each queue). Jobs in the artsci-interactive queue will automatically terminate after 24 hours, while jobs in the artsci queue can run for as long as 28 days

.[1]

Memory

The memory parameter specifies the amount of RAM (in gigabytes) that your job will require. It's important to note that this is different from storage capacity. To choose the appropriate amount of memory, aim for slightly more than what you think you'll need, but be mindful not to request too much. It's important to be considerate of other users and avoid hogging resources that you won't use.

Keep in mind that requesting excessive amounts of memory can also push your job further back in the queue, resulting in a longer wait time before execution.

Number of hours

Put in the number of hours your task requires. The maximum amount here is 28 days. Once your time expires, your task will be automatically terminated so don't forget to set your alarm!

GPUs to Allocate

GPUs are ideal for some kinds of models like neural networks. But if you do not have an application like this, just leave this at zero. GPUs are always in high demand so your queue times will be much longer if you request these unnecessarily.

Number of processors

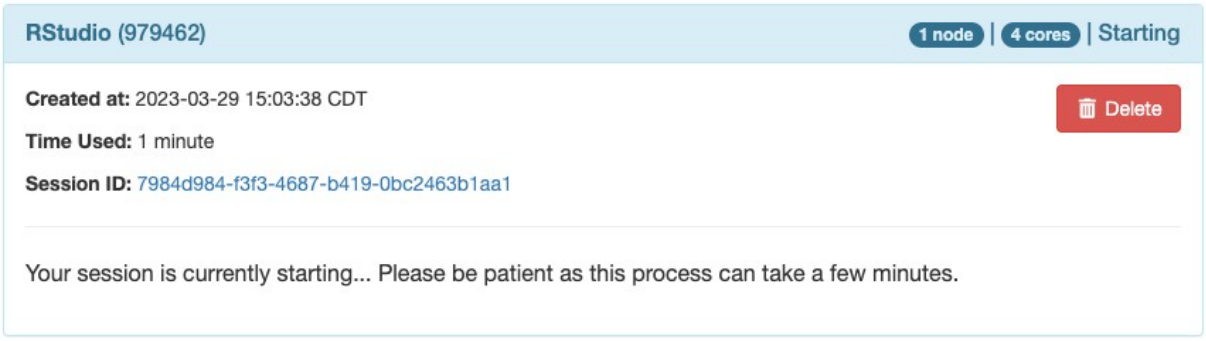
Choose the number of processors you need. But if your code is not multithread and/or parallel there is no point in requesting more than one.

Once you are done filling in the form, select the "Launch" button at the bottom of the screen.

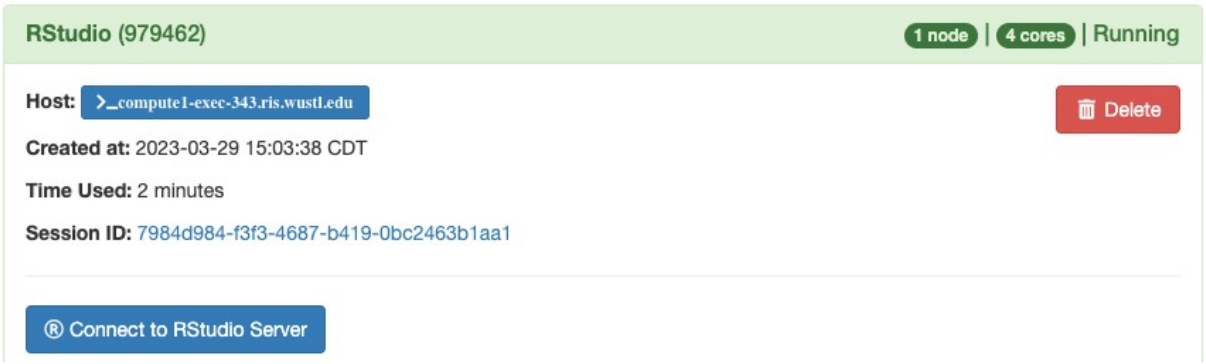
Step 4) Start your RIS instance

Your requested session will now be entered into the queue and then your personal session (with your requested configuration) needs to be spun up. This may take a few minutes to get going. While it is getting set up, it will look like this.

[1] There is no way to have a task run longer than 28 days on teh RIS at this time.



Once it is ready, it will look like this.



Hit the “Connect to RStudio Server” button.

Step 5) Move files and set your directory

You can now use the mapped server in the Finder window to move any files you need for your R session. **NOTE** that all files need to be moved into the `Active` subdirectory. That is the only place that files you use in your task can be stored.

You also need to change your working directory to be pointed to folder you set up in your form. You can do this easily using the `setwd` command in R.

```
setwd("/storage1/fs1/jacob.montgomery/Active")
```

Again, you will need to customize this for your own username.

The working directory will also be where any files that you save are stored (unless your code specified otherwise).

Step 6) Go for it!

With these steps completed, you're now ready to use RStudio as you normally would. Any files you save in your Active Directory will eventually appear in your mapped drive, which you can access through the Finder app.
It's important to note that this is a new session, so you'll need to install any necessary R packages using the "install.packages" command in R.