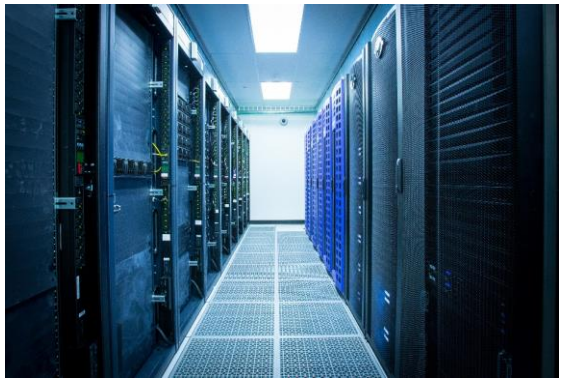
INTRODUCTION TO RIVANNA

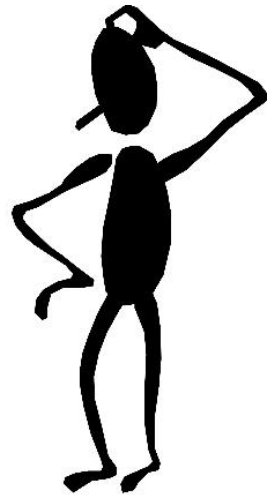
25 September 2020
UVA Research Computing

Rivanna

Rivanna is the university's primary resource for high-performance computation for non-sensitive data. It provides a platform for computationally-intensive research across disciplines.

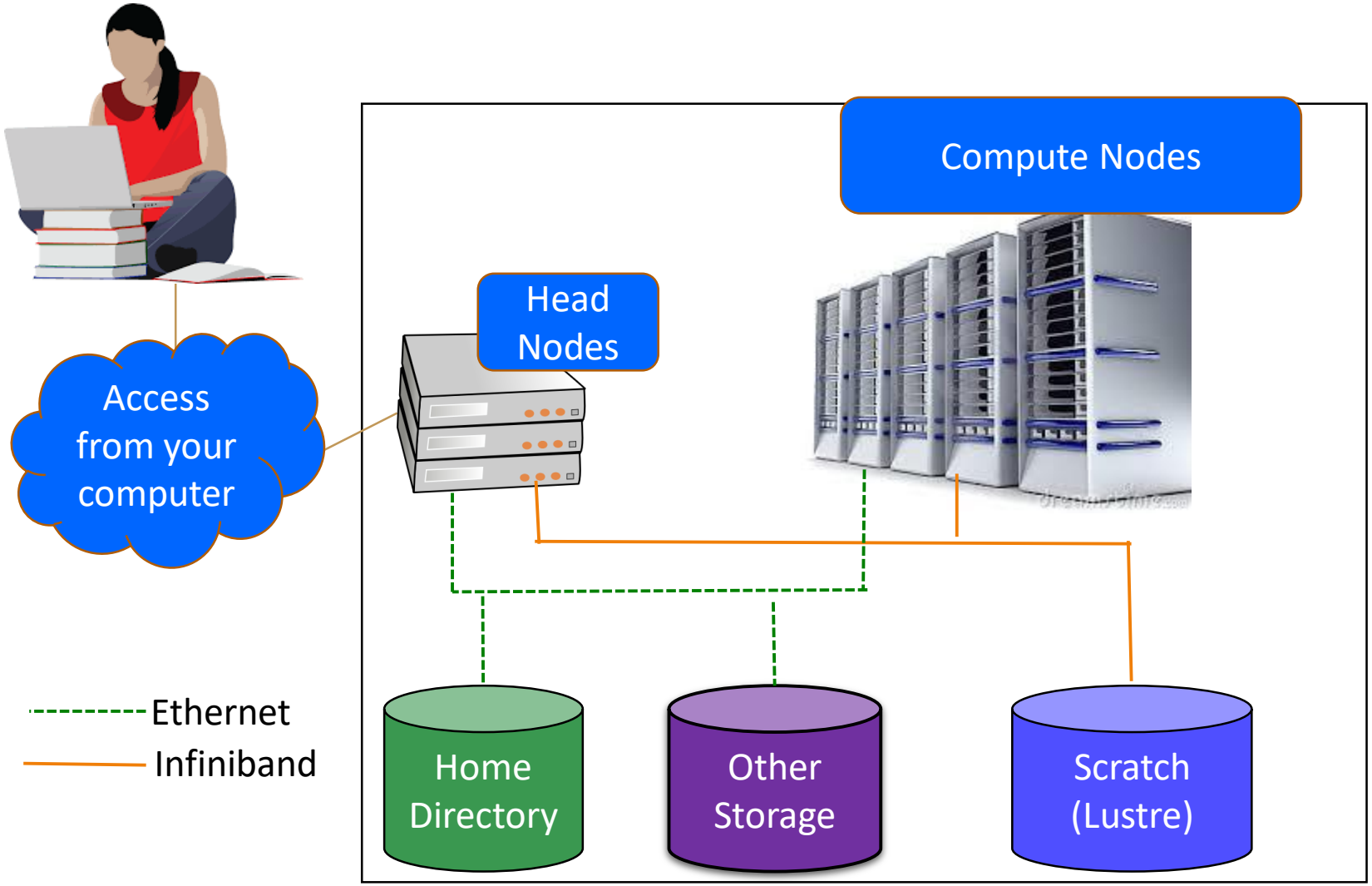


Why would you need to use Rivanna?



- More memory
  - Nodes on Rivanna have many more gigabytes of RAM than your laptop.
  - The larger your data set, the more memory you will need.
- More cores, including GPUs
  - Multiple cores allows parallel processing, which can speed up the computations
  - GPUs are necessary for TensorFlow, Deep Learning models.
- Move computations off of your laptop.
  - If you get to the point where your computations will take hours, you will not want to tie up your laptop for that long.
  - When you are ready, we can teach you how to submit jobs to run “in the background” on Rivanna allowing you to log off or do other work.

Rivanna in More Detail



Terminology

Node – A basic building block of a cluster; Usually a specialized computer

Head Node – computer used for logging on and submitting jobs

Compute Node – computer that does most of the work

Core – an individual processor on a computer

Storage – a location for files, data, and folders

CONNECTING TO RIVANNA

Connecting through the Web Portal

Open onDemand Dashboard

Jupyter Lab

RStudio Server

Connecting through the Web Portal

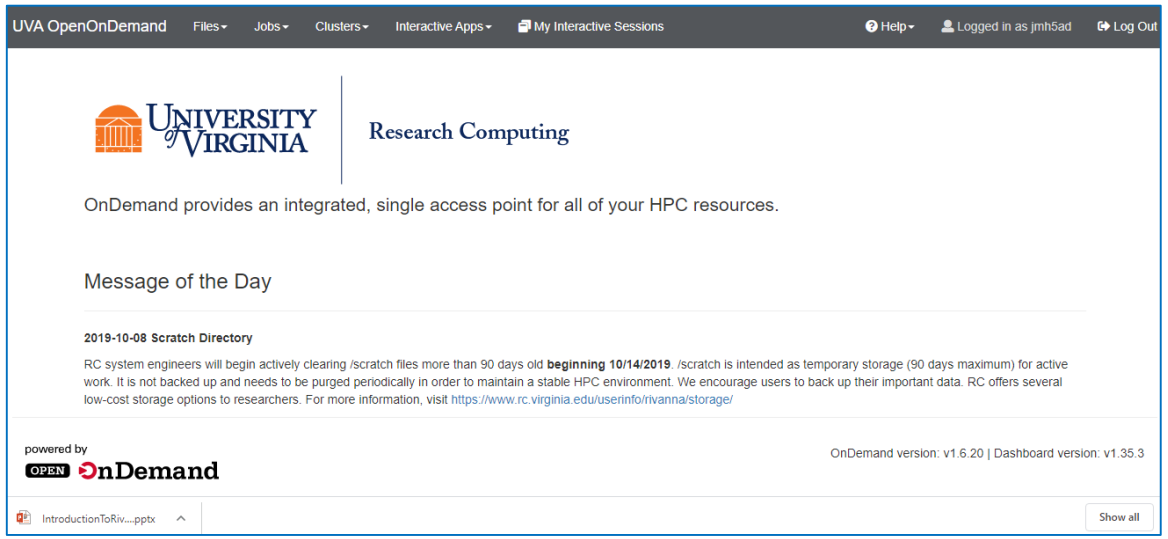
- There are multiple ways to connect to Rivanna
- We are going to focus on connecting through our web portal

Open a web browser on your laptop and type in the URL:
https://rivanna-portal.hpc.virginia.edu

- You will be asked to type in your Netbadge credentials and to confirm through Duo.

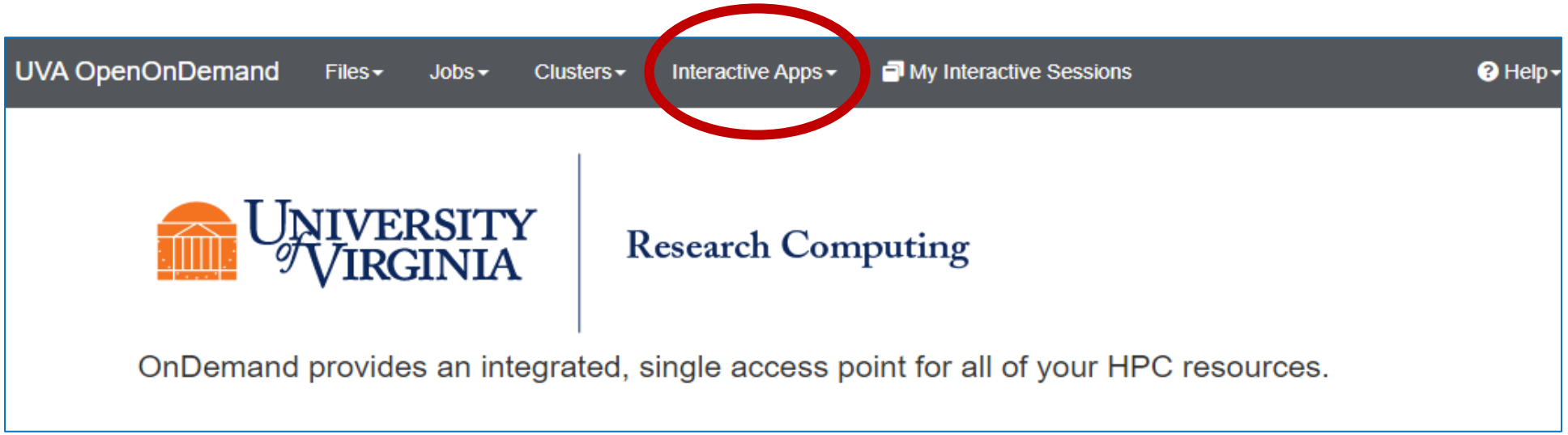
Open onDemand Dashboard

- Open onDemand is the software that we use to create our web portal to Rivanna.
- As soon as you see the Dashboard (shown below), you are connected to Rivanna.

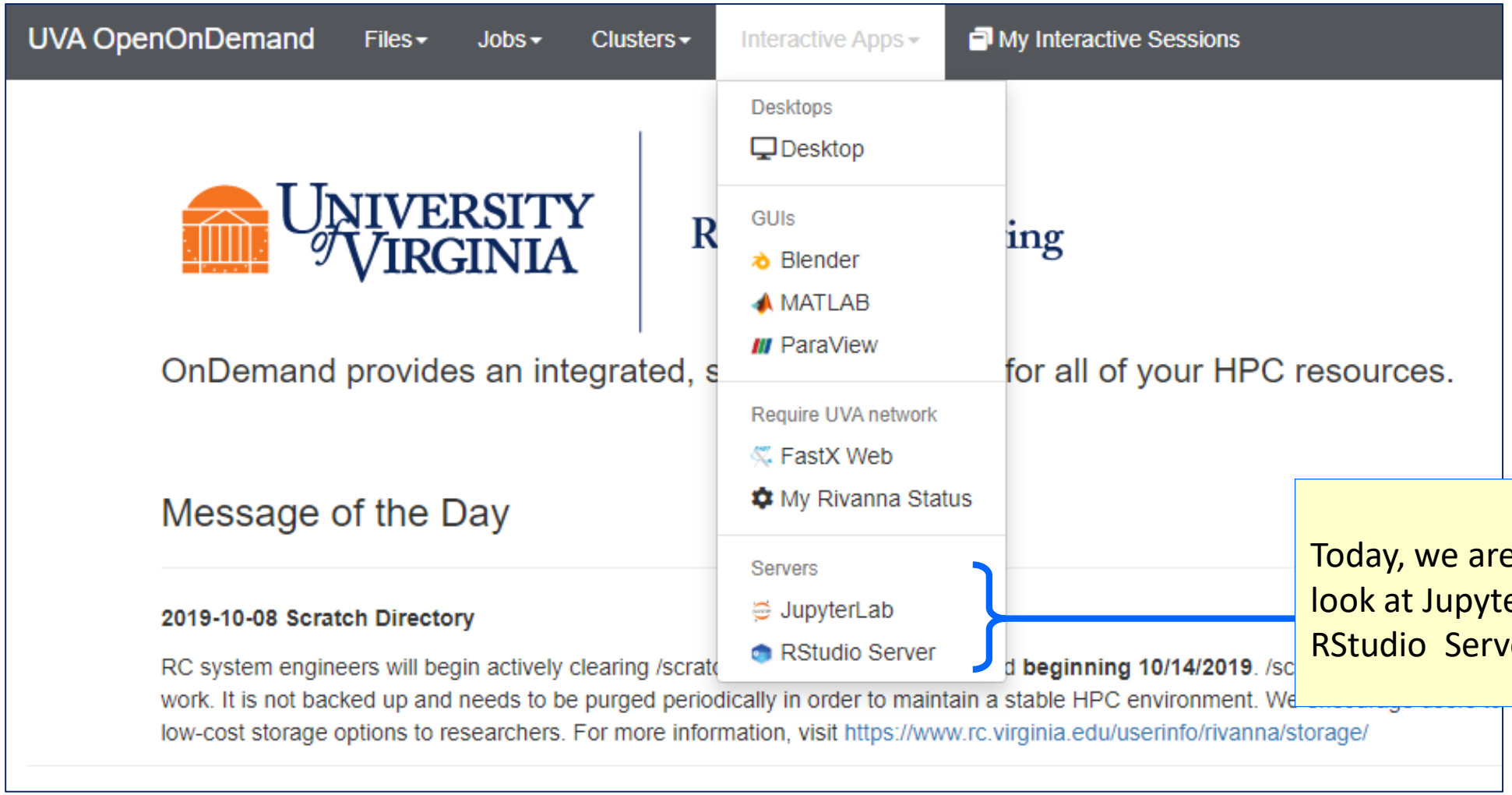


Open onDemand Dashboard

- The Dashboard gives you links to various applications.
- To see the links, click on “Interactive Apps” on the menu bar.

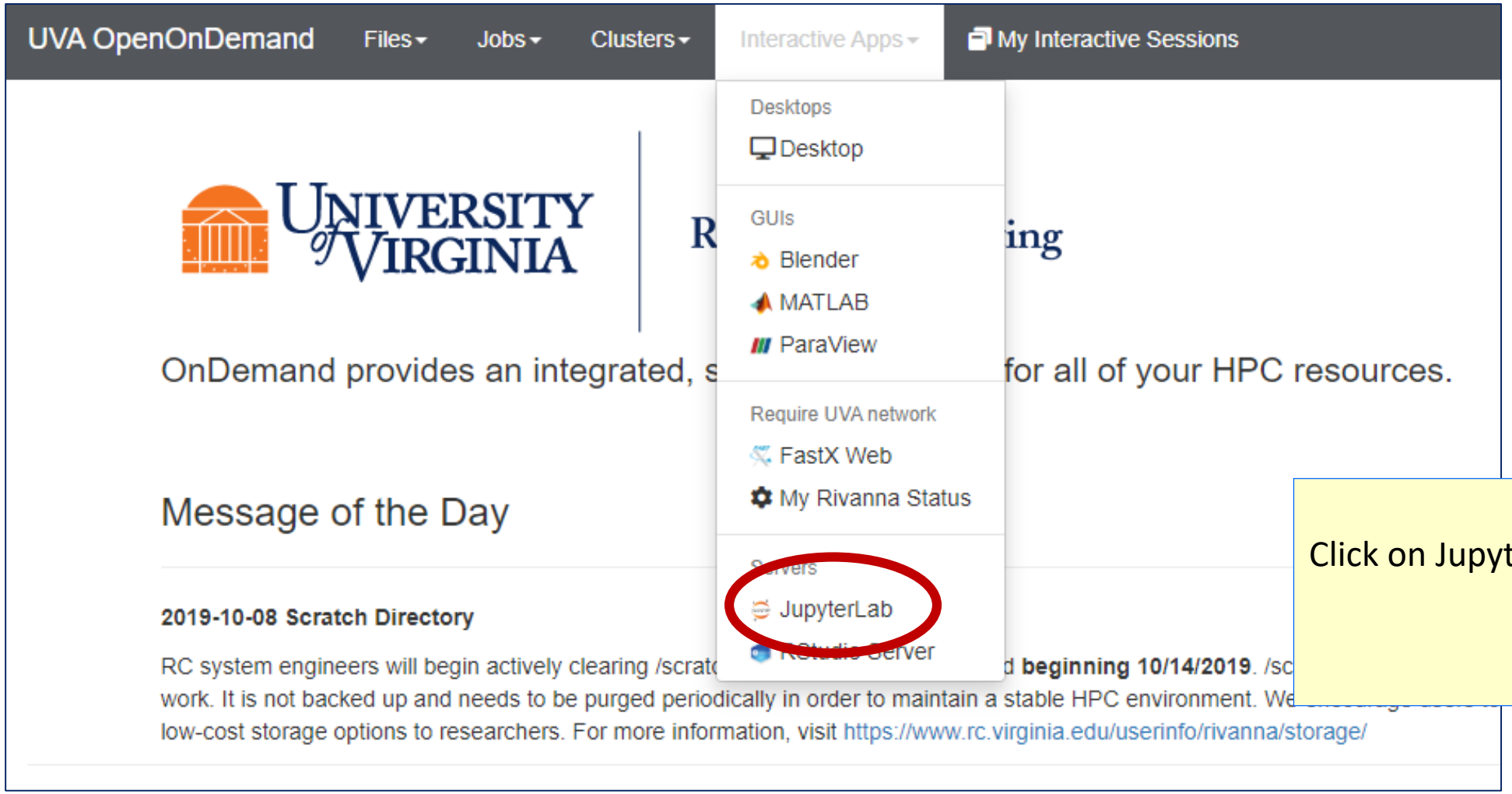


Links to Applications



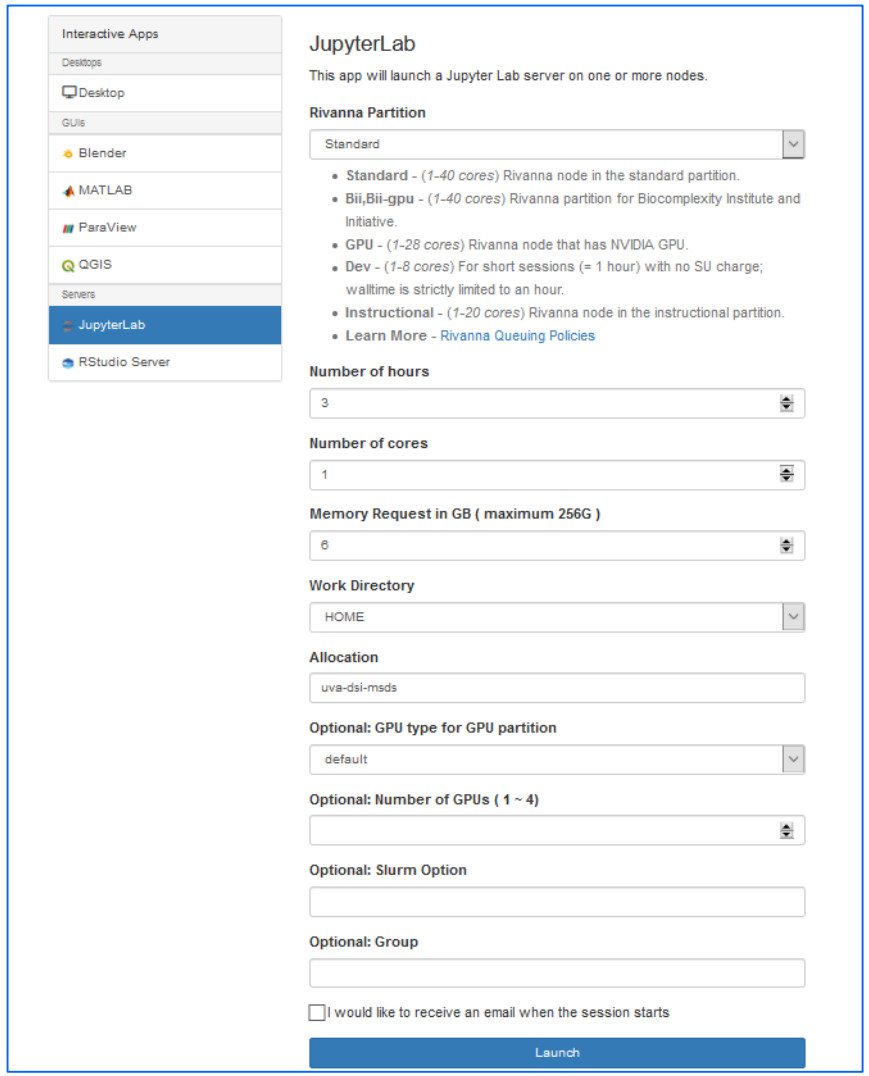
Today, we are going to look at JupyterLab and RStudio Server.

JupyterLab



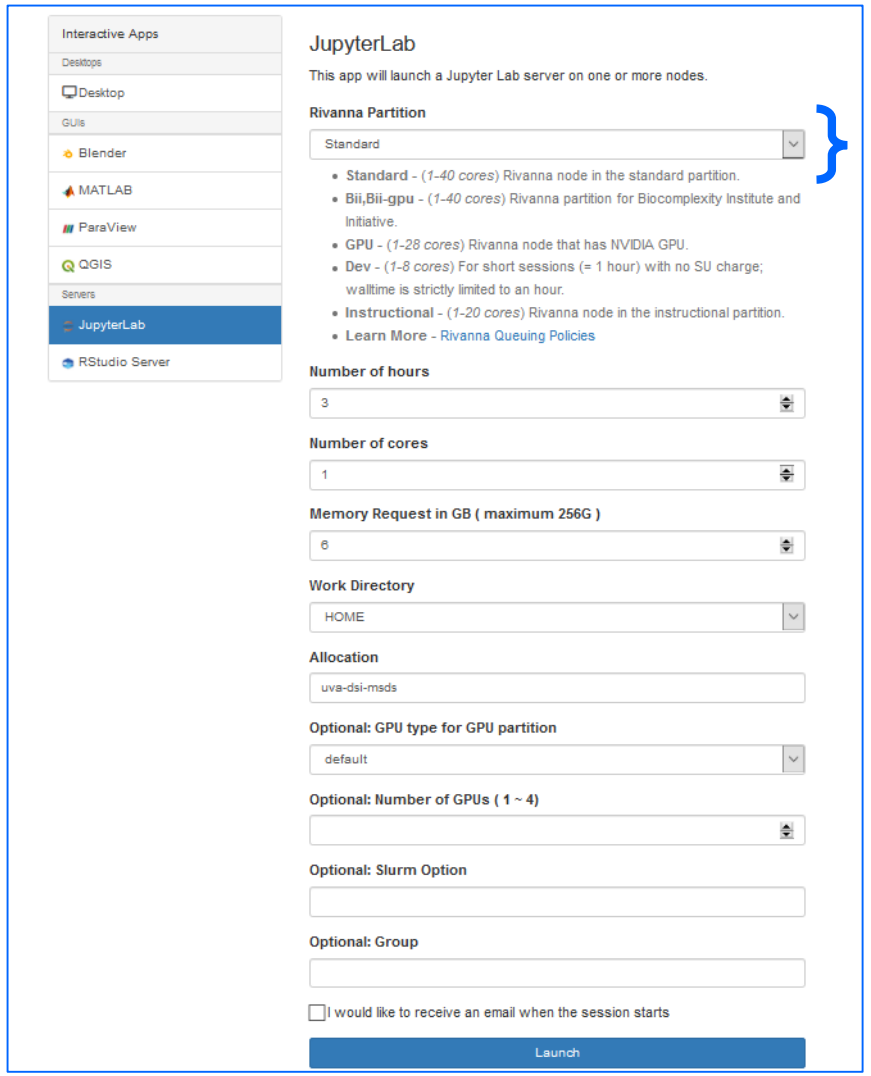
Click on JupyterLab.

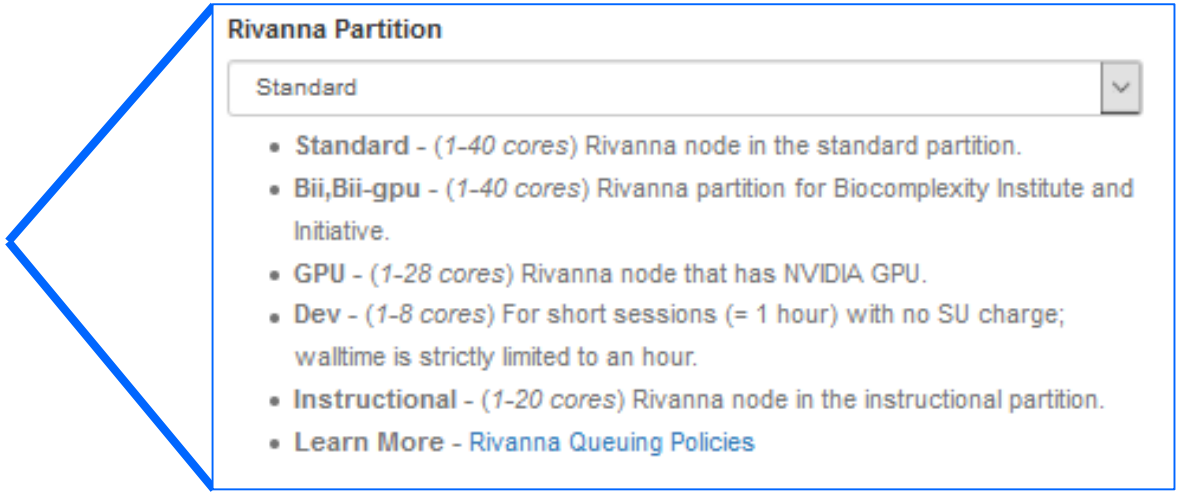
JupyterLab Web Form



- The Jupyter Web Form gathers information about the computing resources that you need for your Jupyter Notebook.
- After you fill in the form, it will re-populate with the same settings the next time that you connect to it.
- Let's look at how you would fill it in!

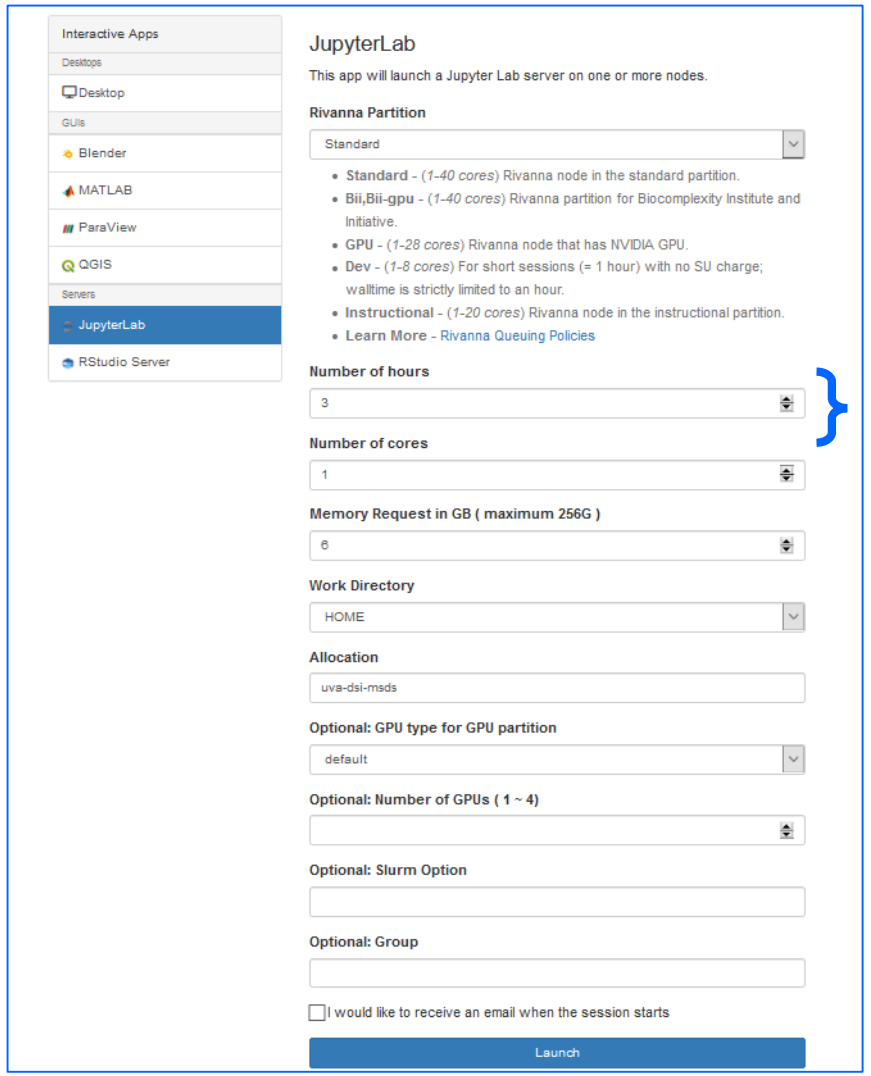
Rivanna Partition

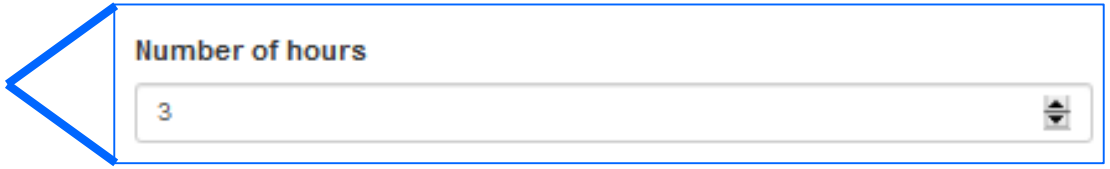




- Recall that Rivanna has lots of Compute Nodes.
- The nodes are partitioned (i.e., organized) by the type of processing that they can do.
- Most of the time, you will select the **Standard** partition.
- If you are running a deep learning model, you will want to choose a **GPU** Partition.

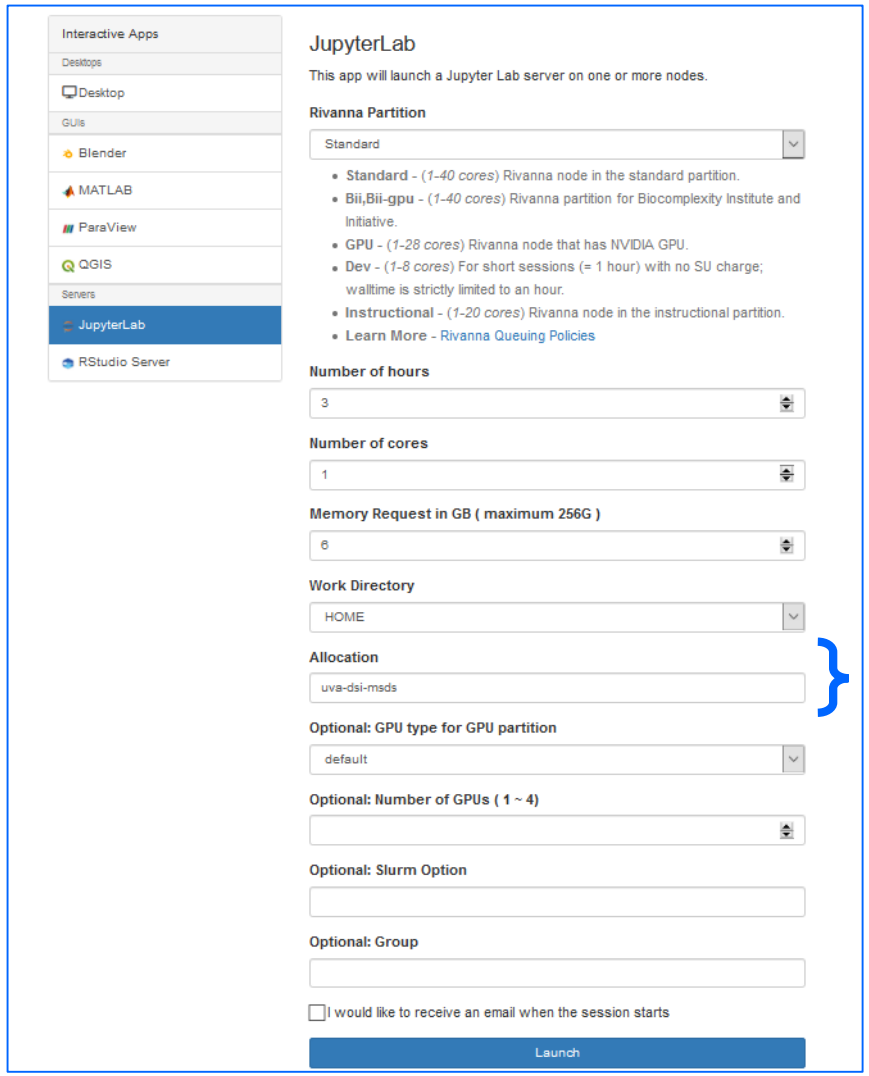
Number of Hours



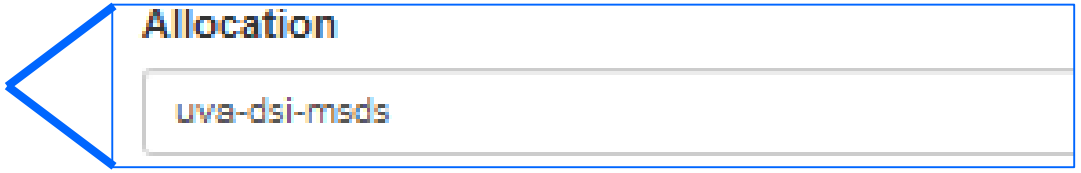


- The “Number of hours” is the amount of time that your session will be active.
- Beware! When time runs out the session will end without warning!

Allocation

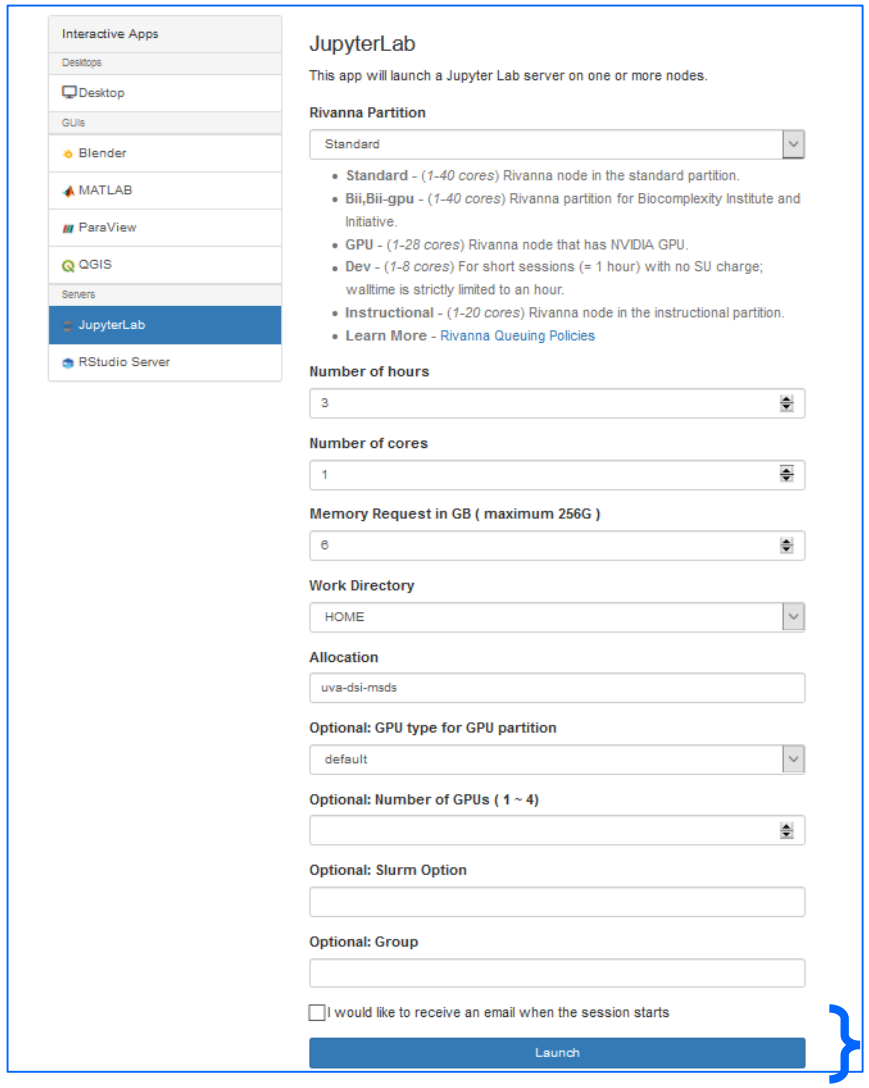


- The allocation is a special MyGroups group that allows you to have access to Rivanna.
- You must be a member of a Rivanna-enabled MyGroup to have an active account.
- In general, your professor or research advisor will add you as a member to an allocation.

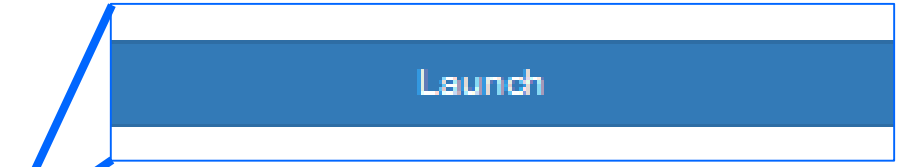


- You can have membership in more than one allocation.

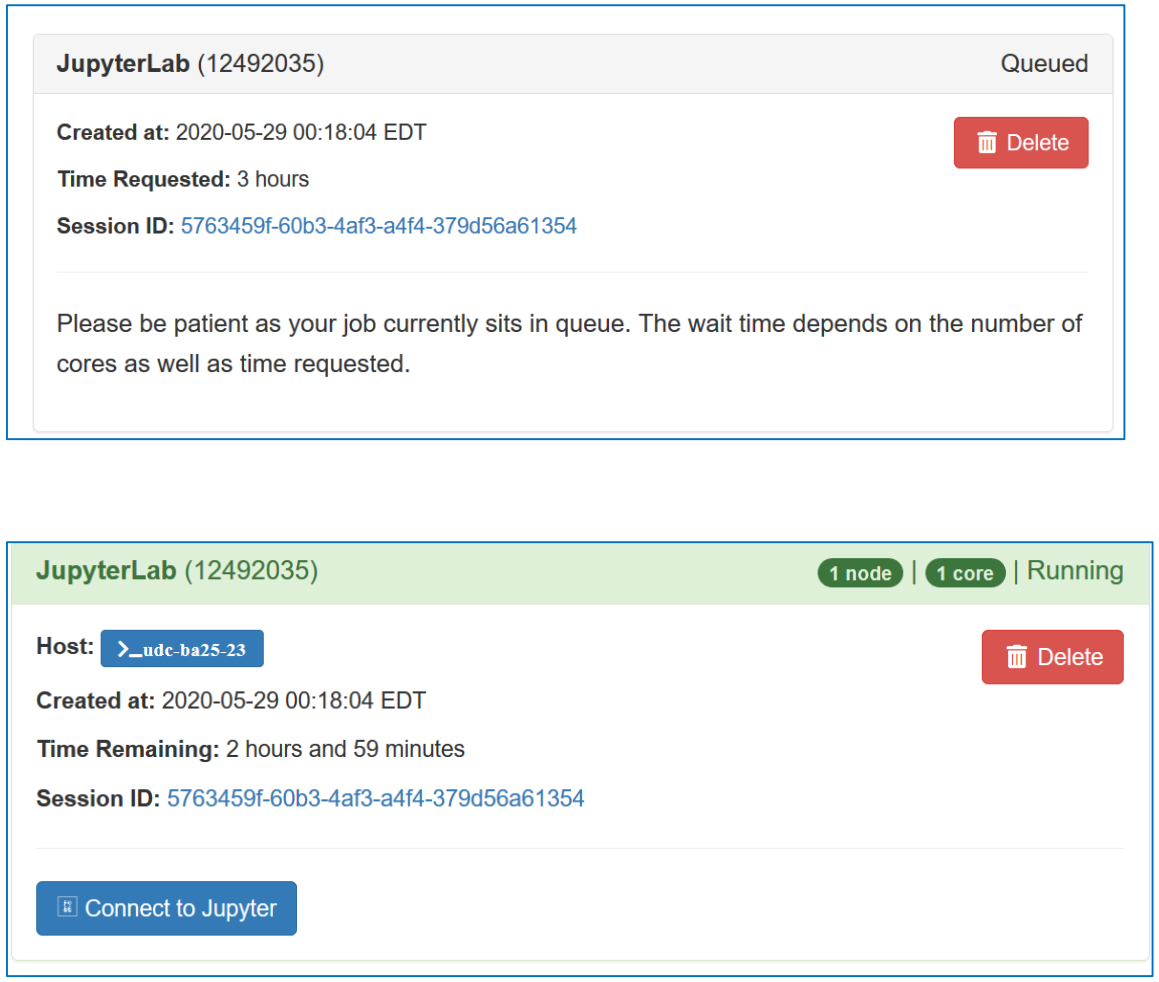
Launch



- Clicking on the “Launch” button will submit a request for the resources that you want.
- There will be a slight delay before the resources are available.



Waiting for the Session to Start

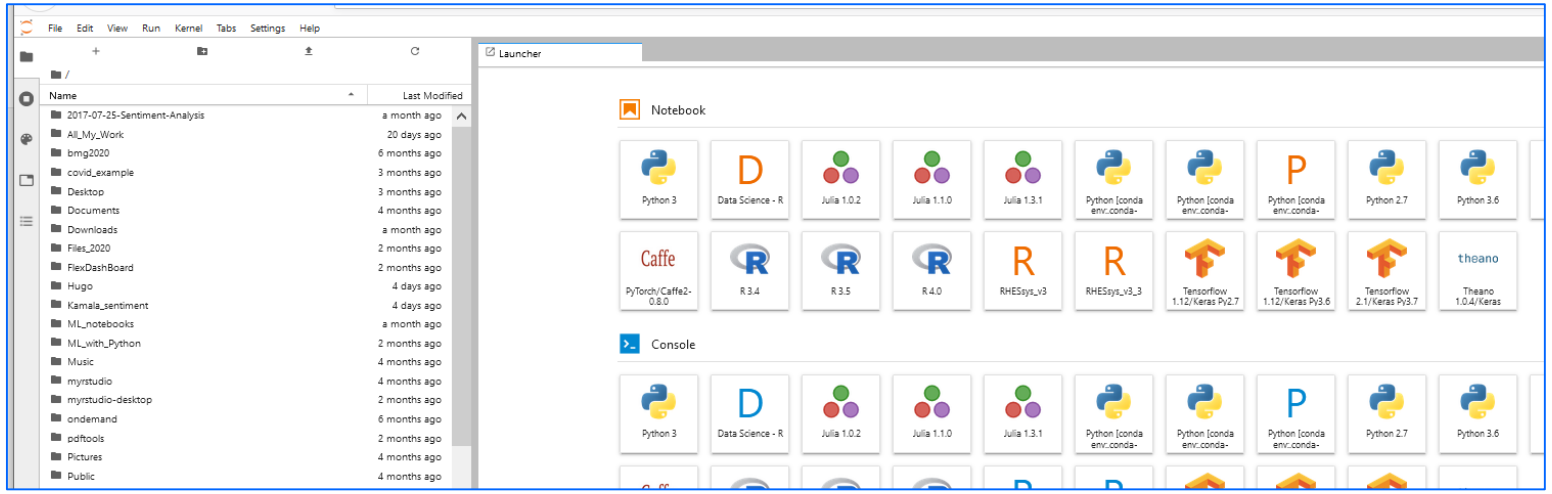


The screen will transition from a "Please be patient" statement to a "Connect to Jupyter" button.

Click on the "Connect to Jupyter" button.

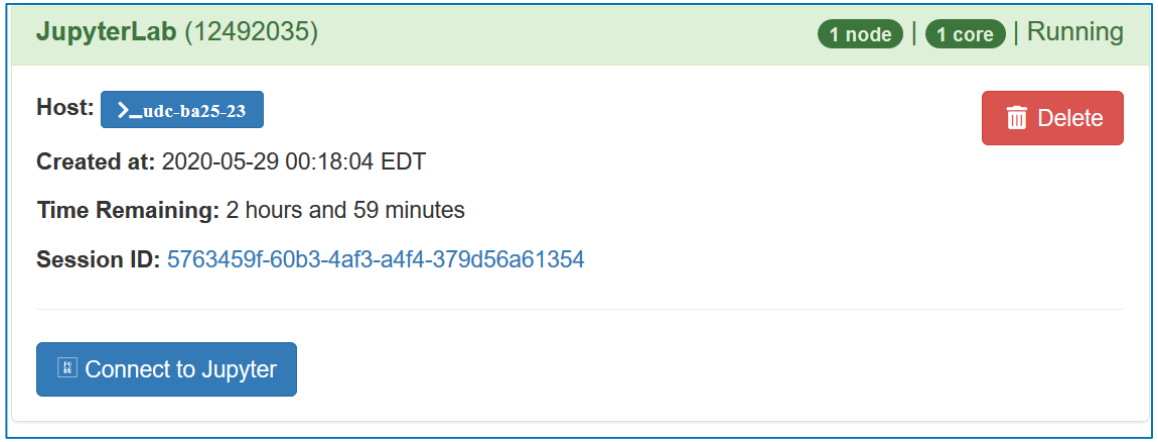
Hands-on Activity

- Connect to Open onDemand and start a JupyterLab session.
  - When it comes up, you will see a list of files (if any) in your home directory and a set of tiles for underlying applications (e.g., Python, R, Tensorflow).
  - You may see a slightly different set of tiles in your account – there are some customized tiles in this account.

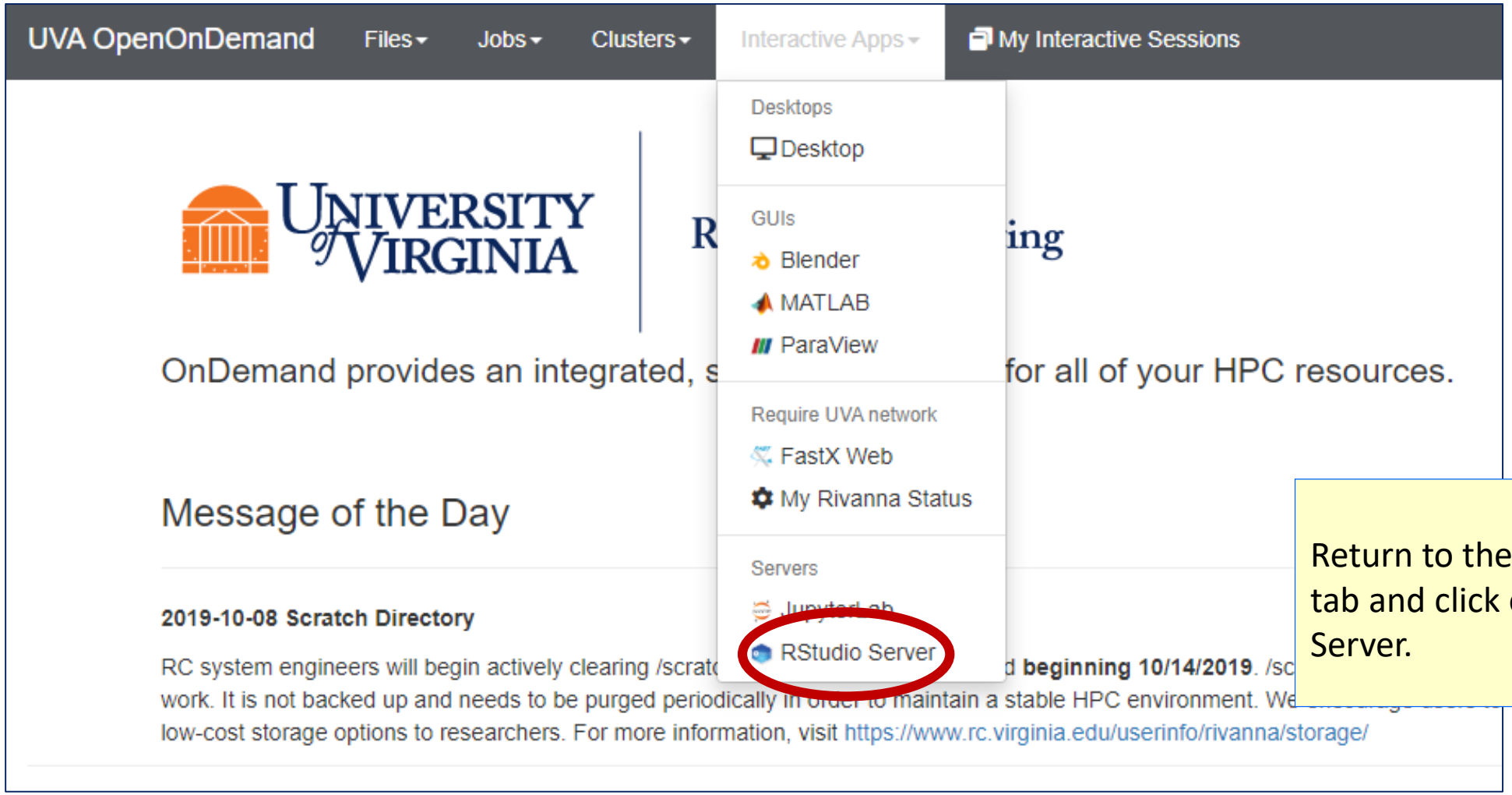


Deleting Your Session

- When you are done with your Jupyter Notebook, it is very important to delete the session.
- Go back to the browser tab labeled “Interactive Sessions” and click on the red “Delete” button.

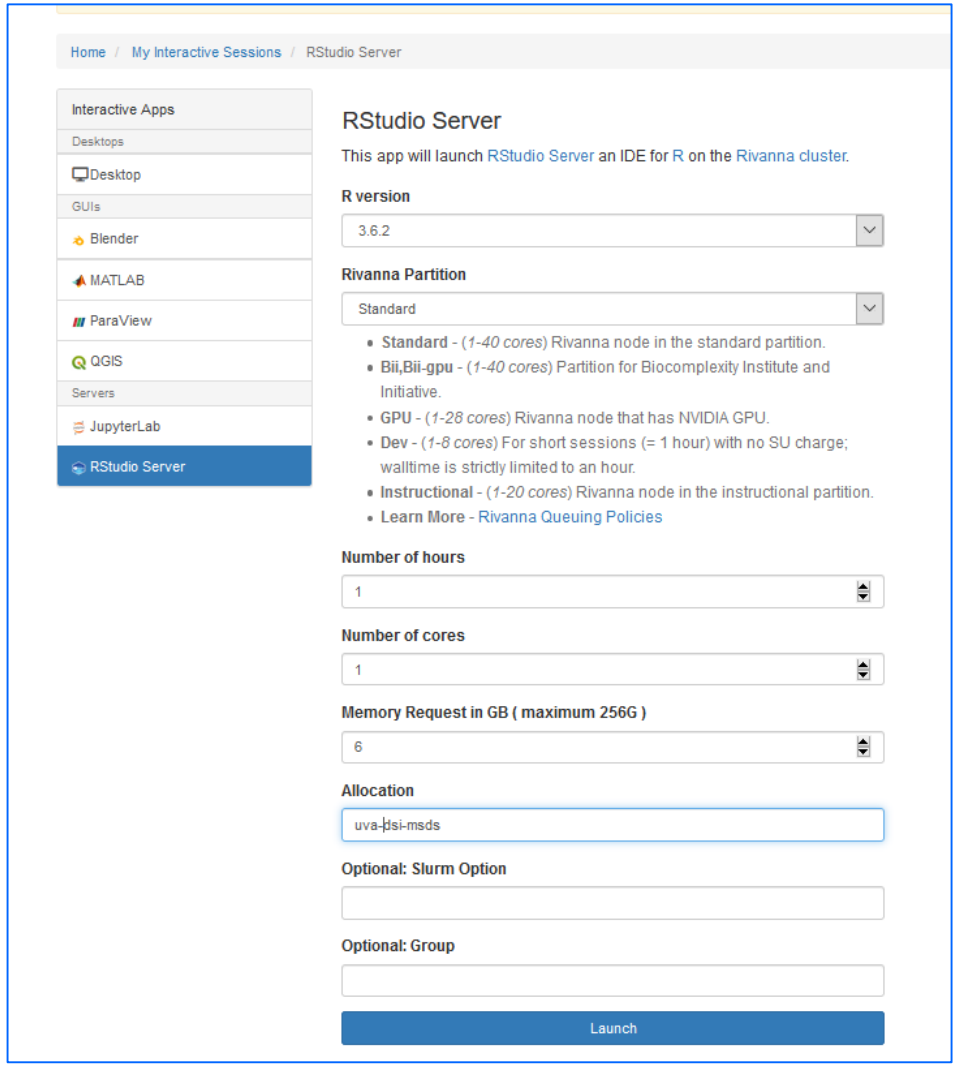


RStudio Server



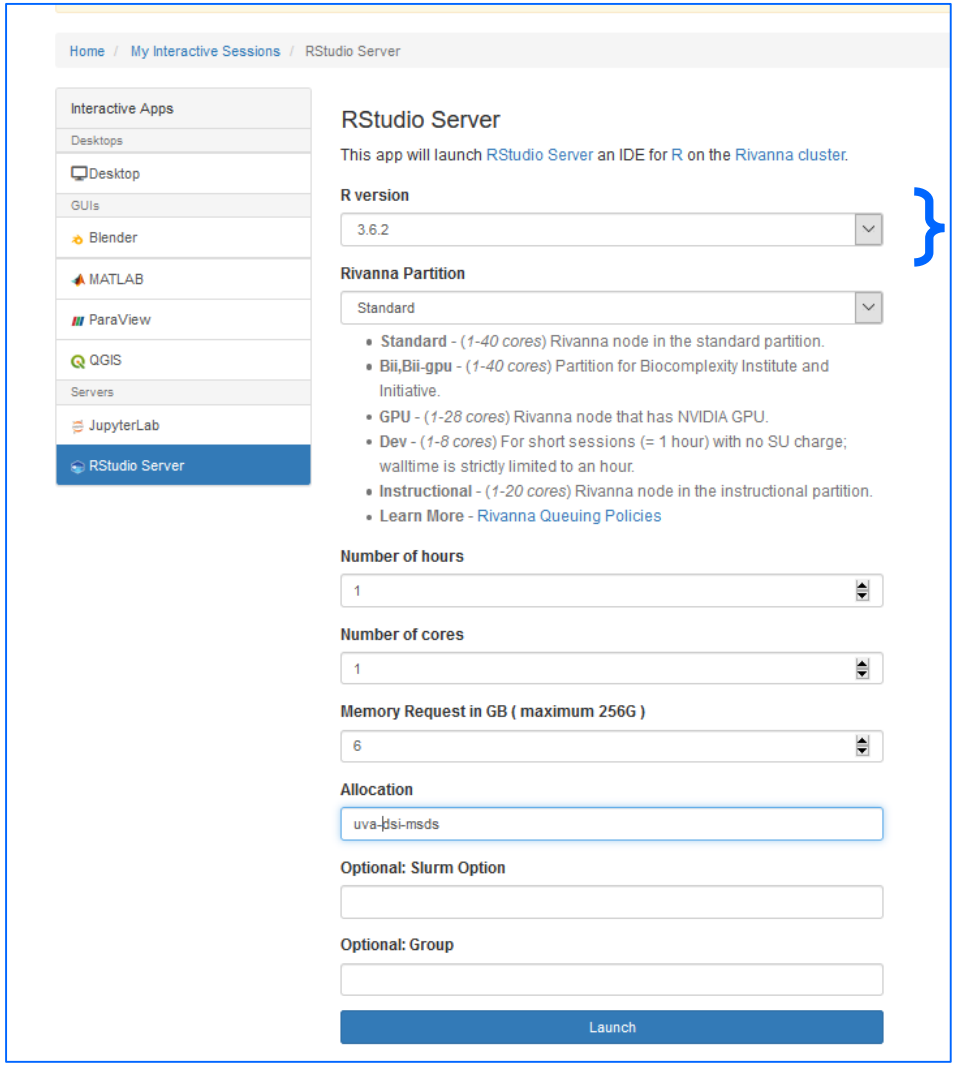
Return to the Dashboard tab and click on RStudio Server.

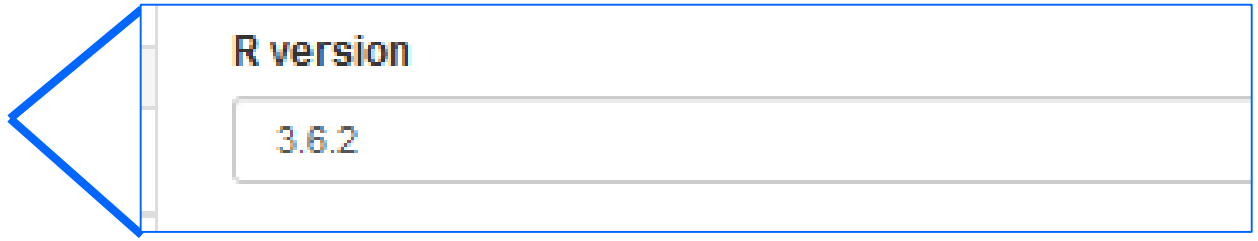
RStudio Web Form



- The RStudio Web Form is similar to what you saw with the Jupyter Web Form.
- Let's take a quick look at some of the fields.

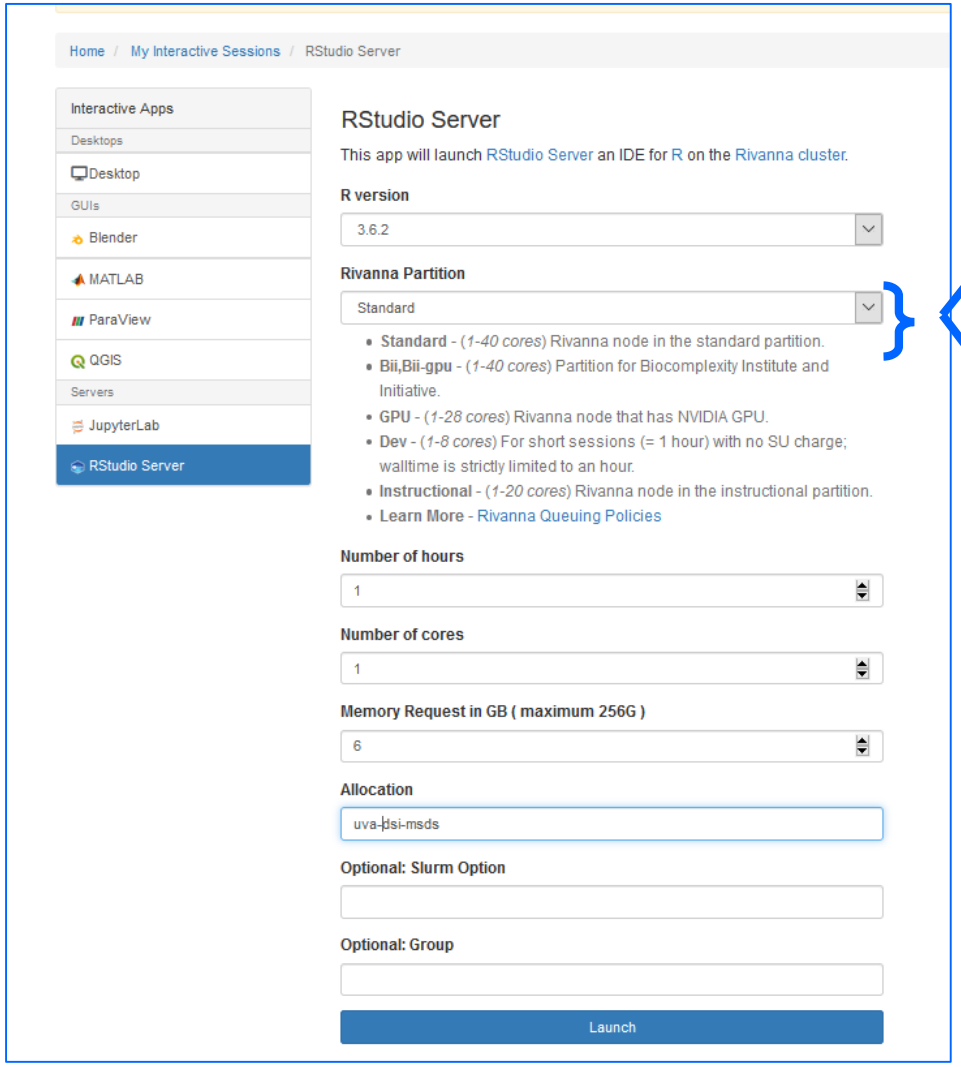
R Version

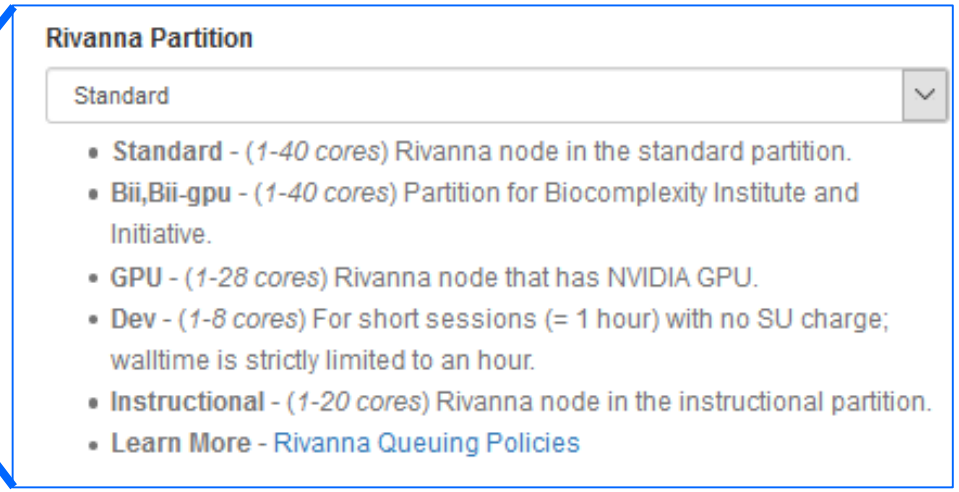




- We have three versions of R available.
- R/3.6.2 is the most stable at this time.

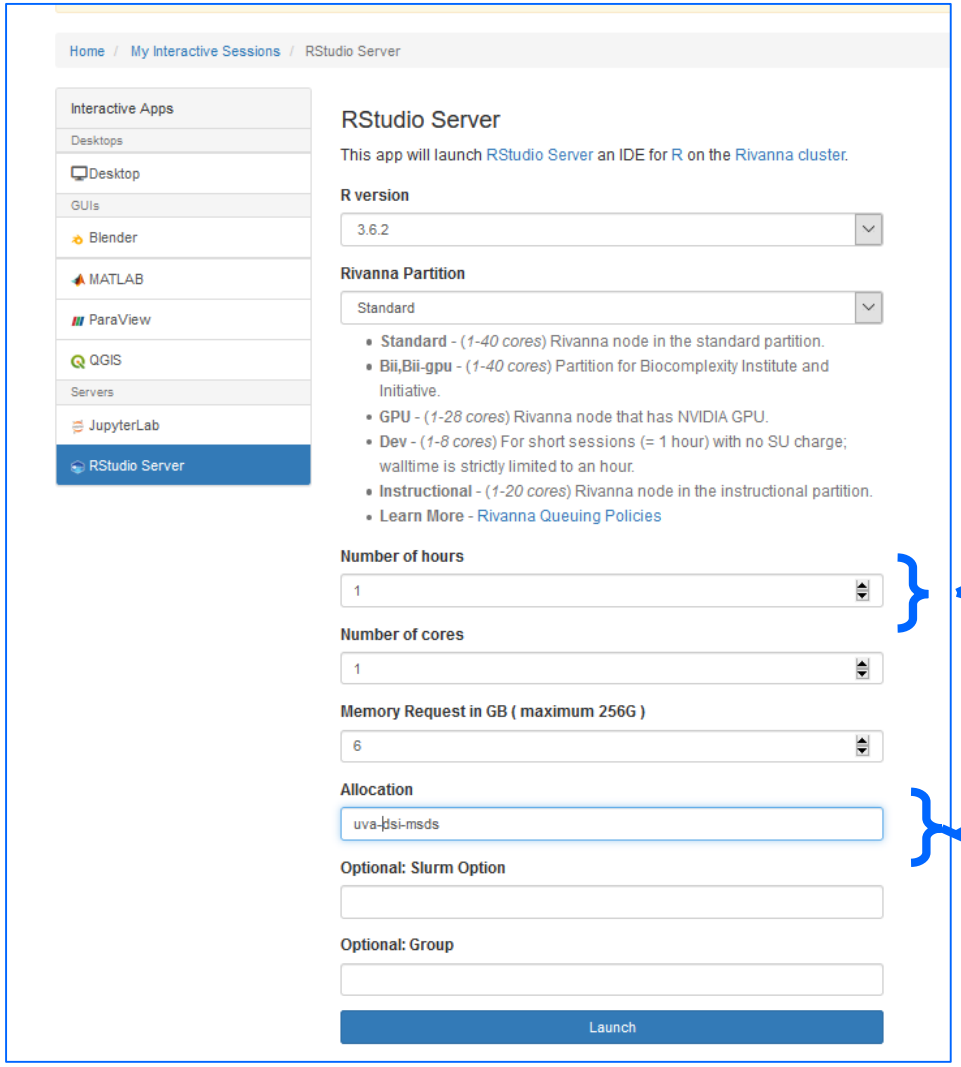
Rivanna Partition



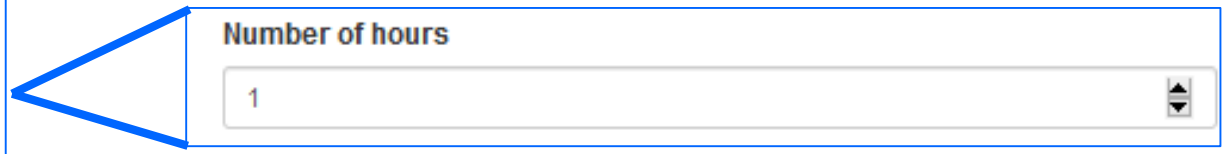


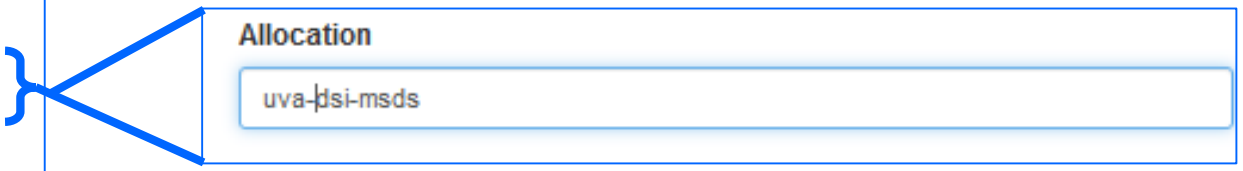
- The partitions are the same as in the JupyterLab Web Form, but we recommend mostly Standard, Dev, or Instructional for RStudio Server.

Number of hours & Allocation

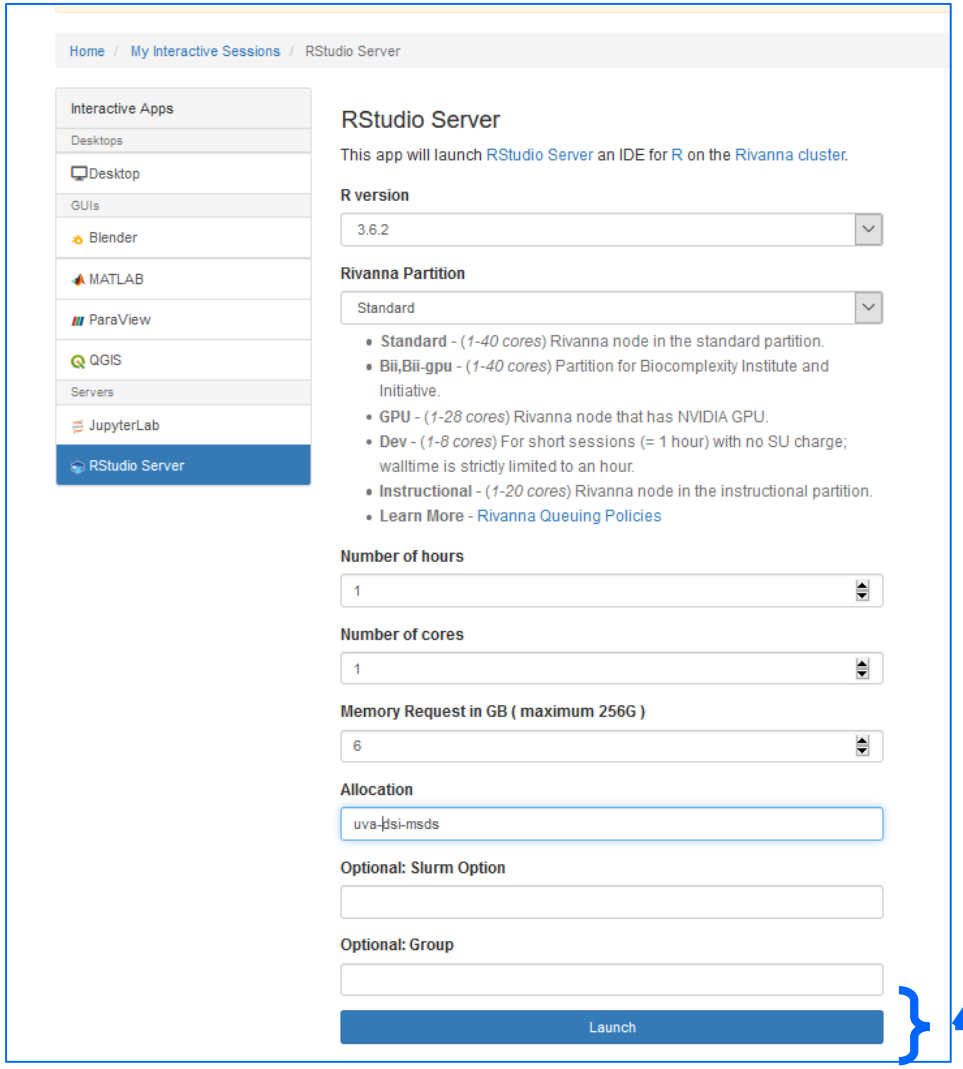


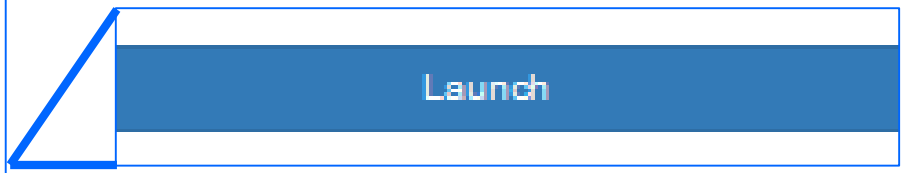
- The Number of hours and the Allocation are the same as in the JupyterLab Web Form.





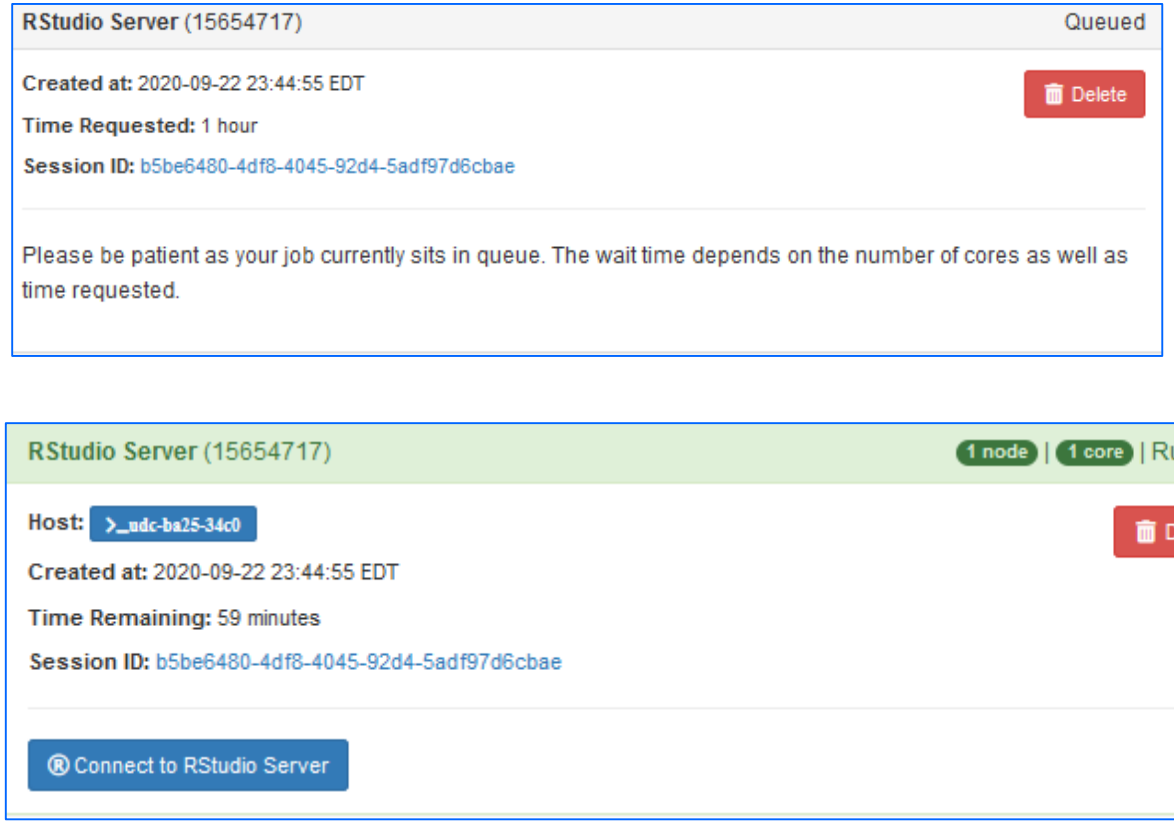
Launch





- Clicking on the “Launch” button will submit a request for the resources that you want.
- Again, there will be a slight delay before the resources are available.

Waiting for the Session to Start

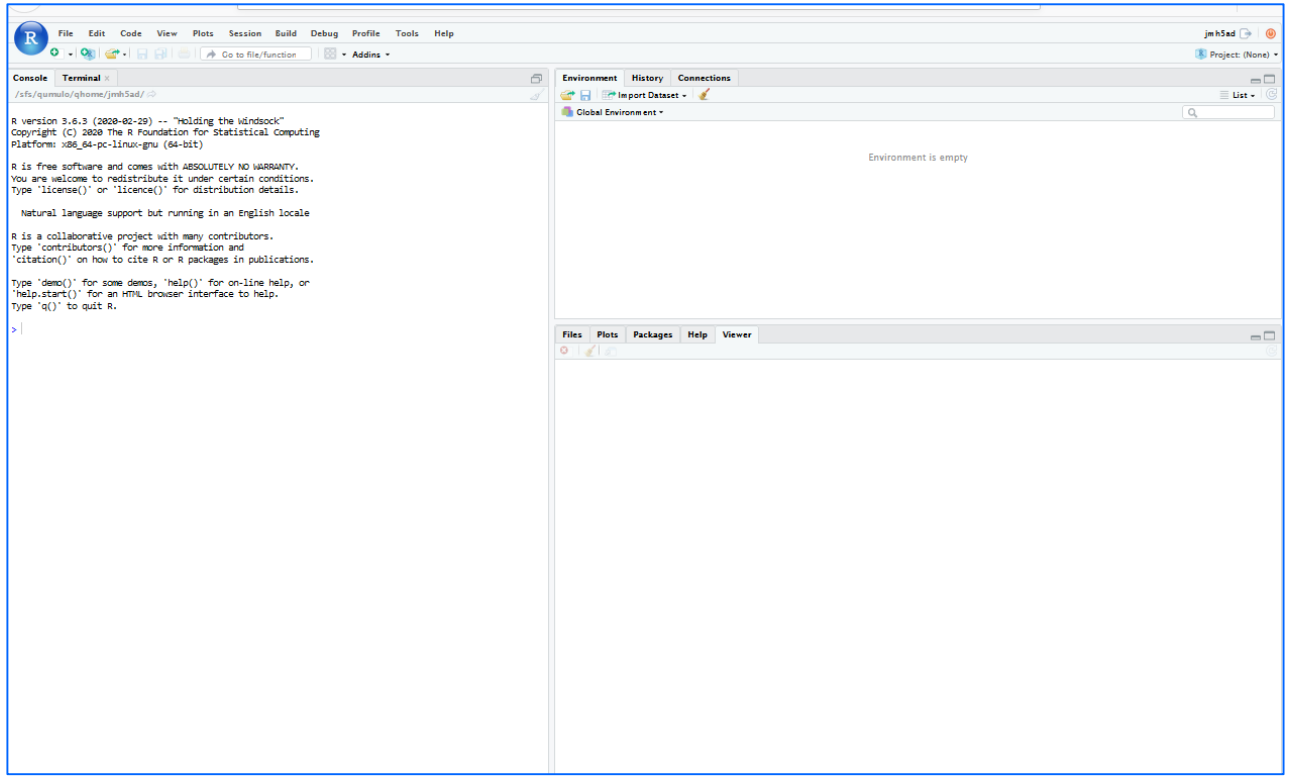


The screen will transition from a "Please be patient" statement to a "Connect to RStudio Server" button.

Click on the "Connect to RStudio Server" button.

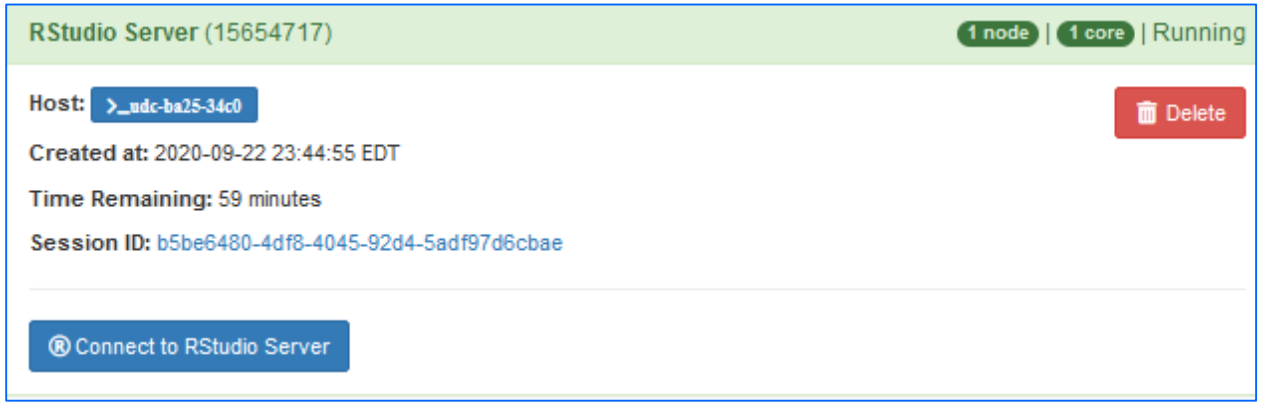
Hands-on Activity

- Connect to Open onDemand and start an RStudio Server session.
  - When it comes up, you will see an RStudio interface – just like you would see on your laptop.



Deleting Your Session

- Again, when you are done with RStudio Server, it is very important to delete the session.
- Go back to the browser tab labeled “Interactive Sessions” and click on the red “Delete” button.



What about the other fields?

Field	Description
Number of cores	Used in parallel processing. Your code must be modified to take advantage of using multiple cores.
Memory Request in GB	When dealing with Big Data, you will need to increase the amount of memory. My rule-of-thumb: request 2 to 3 times the size of data that you are reading in or generating.
Work Directory	Allows you to change the working directory of a Jupyter Notebook to your /scratch folder.
Optional: Slurm Option	Allows you to provide advanced features, like requesting specific nodes or providing a reservation
Optional Group	Only needed in you are in more than 16 allocations. You may need to force Rivanna to see your allocation.
Optional: GPU type for GPU partition & Optional: Number of GPUs	Only needed in you are running on a GPU node. The "default" for GPU type will put you on the first available GPU node. For now, the number of GPUS should be 1.

- You may have noticed fields on the Web Forms that we left blank or stayed with the default values.
- The most important one will be the Memory Request

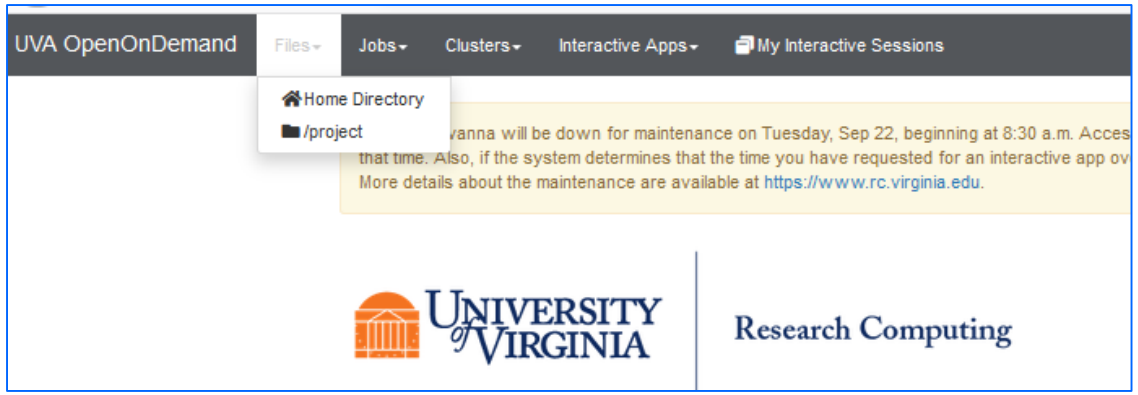
MOVING DATA (OR FILES) ONTO RIVANNA

Open onDemand File Explorer

Globus File Transfer

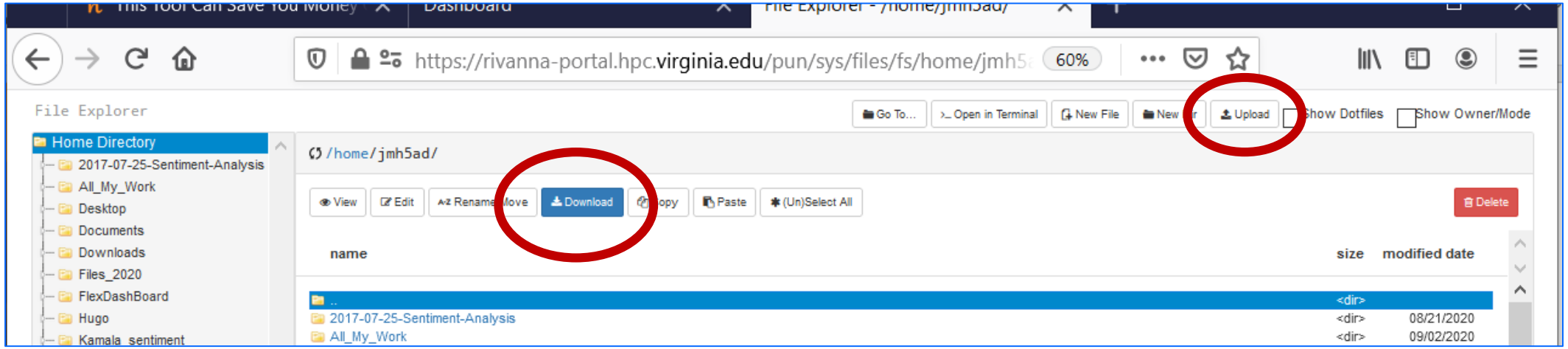
Moving Relatively Small Files: Open onDemand

- If your files are less than 10 GB, you can use the Open onDemand File Explorer.
  - Go to the Open onDemand Dashboard (https://rivanna-portal.hpc.virginia.edu)
  - Click on Files
  - Select Home Directory (or /project if your advisor has purchased additional storage)



Open onDemand File Explorer

- In the File Explorer, you can choose the button to Upload or Download Files.
- In each case a browser prompt will appear to select the files to upload or download.



Moving Large Files: Globus File Transfer

- If your files are 10 GB or more, we recommend that you use the Globus File Transfer service.
  - You will need to set up an endpoint (or collection) on your laptop.
  - Then, you can use the Globus web interface to transfer files.

- Details for these steps (and more) are available at

  https://www.rc.Virginia.edu/userinfo/globus/

Wrap Up

- These slides provide just an introduction on how to connect to Rivanna.
- There are many more details that we should cover.
- To learn more about Rivanna, we recommend:
  https://workshops.rc.virginia.edu/lesson/introduction-to-rivanna/
  https://workshops.rc.virginia.edu/lesson/rivanna-in-command-line/

Questions about Rivanna?

You can

- check out our frequently-asked questions at https://www.rc.virginia.edu/userinfo/faq/rivanna-faq/
- go to https://www.rc.virginia.edu/support/ and click on “Open a Support Ticket” to ask a specific question
- email us directly at hpc-support@virginia.edu
- visit us during our Zoom Office Hours (see https://www.rc.virginia.edu/support/#office-hours for days/times and links)