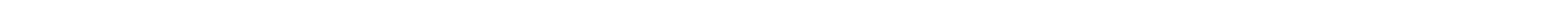


Using RStudio on Hoffman2

Hoffman2 Happy Hour

Charles Peterson



Welcome to Hoffman2 Happy Hours



Welcome to the Hoffman2 Happy Hours

- Short presentations on topics related to HPC and Hoffman2
- Thoughts for future “Happy Hour” topics?



cpeterson@oarc.ucla.edu



Files For This Presentation

This presentation can be found on our github page

https://github.com/ucla-oarc-hpc/H2HH_rstudio

The html slides can be found at

https://ucla-oarc-hpc.github.io/H2HH_rstudio

More information and scripts on using RStudio on Hoffman2

<https://github.com/ucla-oarc-hpc/H2-RStudio>

RStudio Information

What Is RStudio

RStudio is a great IDE for R and visualize files.



But why do you want to use
RStudio on Hoffman2 when
you can use your own
computer???

RStudio on Hoffman2 provides access:

- higher memory
- multi-core
- GPUs
- Your data on Hoffman2

RStudio Formats

There are two main (free) RStudio formats that researchers can use

RStudio Desktop

- Standalone desktop application
- Installed locally on your machine

RStudio Server

- Run RStudio as a server process
- Open on a web browser

RStudio on Hoffman2

RStudio Desktop can be inefficient on Hoffman2

- require X11 forwarding
- slow
- sluggish interaction

RStudio Server is the best way to use RStudio on Hoffman2

- Docker container with R/Rstudio
 - Ran with Apptainer
 - [Workshop on using containers on Hoffman2](#)
- Isolate container OS image
 - Separate version of R that was built on Hoffman2

Running RStudio

Running RStudio (1)

Get An Interactive Job

Containers cannot run on login nodes.

```
1 qrsh -l h_data=10G
```

- You **MUST** use a compute node

Modify the qrsh to meet your RStudio computing needs

- More memory and/or job time
- More cores
- Using GPUs

```
1 qrsh -l h_data=50G,h_rt=5:00:00
```

```
1 qrsh -l h_data=10G -pe shared 10
```

```
1 qrsh -l h_data=10G,gpu,V100
```

Running RStudio (2)

Create Temp Directories

- Create writable temp directories

- RStudio writes small files

- Anywhere you have write access

```
1 mkdir -pv $SCRATCH/rstudiotmp/var/lib  
2 mkdir -pv $SCRATCH/rstudiotmp/var/run  
3 mkdir -pv $SCRATCH/rstudiotmp/tmp
```

Running RStudio (3)

Load the Apptainer Module

- Apptainer is software that will run the Rstudio container

```
1 module load apptainer
```

RStudio Server on Hoffman2 created from Docker

- Separate R from modules on Hoffman2
 - DO NOT load R modules
 - R packages may need to be reinstalled

Running RStudio (4)

Start Up RStudio

```
1 apptainer run \
2 -B $SCRATCH/rstudiotmp/var/lib:/var/lib/rstudio-server \
3 -B $SCRATCH/rstudiotmp/var/run:/var/run/rstudio-server \
4 -B $SCRATCH/rstudiotmp/tmp:/tmp \
5 $H2_CONTAINER_LOC/h2-rstudio_4.1.0.sif
```

- `apptainer run`
 - Starts the RStudio container
- `-B $SCRATCH/rstudiotmp/[dir]:[/dir]`
 - Mounts tmp directories to the container
- `$H2_CONTAINER_LOC/h2-rstudio_4.1.0.sif`
 - Location of RStudio container
 - Can be change to different RStudio versions
- Information will display about RStudio session
 - Note the **compute node name** and **port number**.
 - Displays `ssh -N -L ...` info to be ran



Note

KEEP THIS TERMINAL OPEN UNTIL YOU JOB IS DONE

Running RStudio (5)

- Open another terminal on your local computer
- Run the port forward command
 - Creates a connection from local computer to compute node

```
1 ssh -N -L 8787:nXXX:8787 username@hoffman2.idre.ucla.edu
```

- Change port **8787** if needed
- **nXXX** is the compute node name
- **username** is your Hoffman2 username

Running RStudio (6)

- Finally, open a web browser
 - Type URL of RStudio Server
 - Will ALWAYS be localhost
 - Change port **8787** if needed

```
1 http://localhost:8787
```

Running Rstudio - The Easy Way

- [h2_rstudio.sh](#)
 - Script that runs everything from the previous slide
 - Starts Rstudio and opens a web browser for you
 - Runs on your **local computer** (not Hoffman2)
- Download script
- To display how to use this script
- Run script
 - Replace [username](#) with Hoffman2 username



h2-studio.sh Information

Look at our [Github page](#)

```
1 wget https://raw.githubusercontent.com/ucla-oarc-hpc/H2-RStudio/main/h2_
2 chmod +x h2_rstudio.sh
```

```
1 ./h2_rstudio.sh -h
```

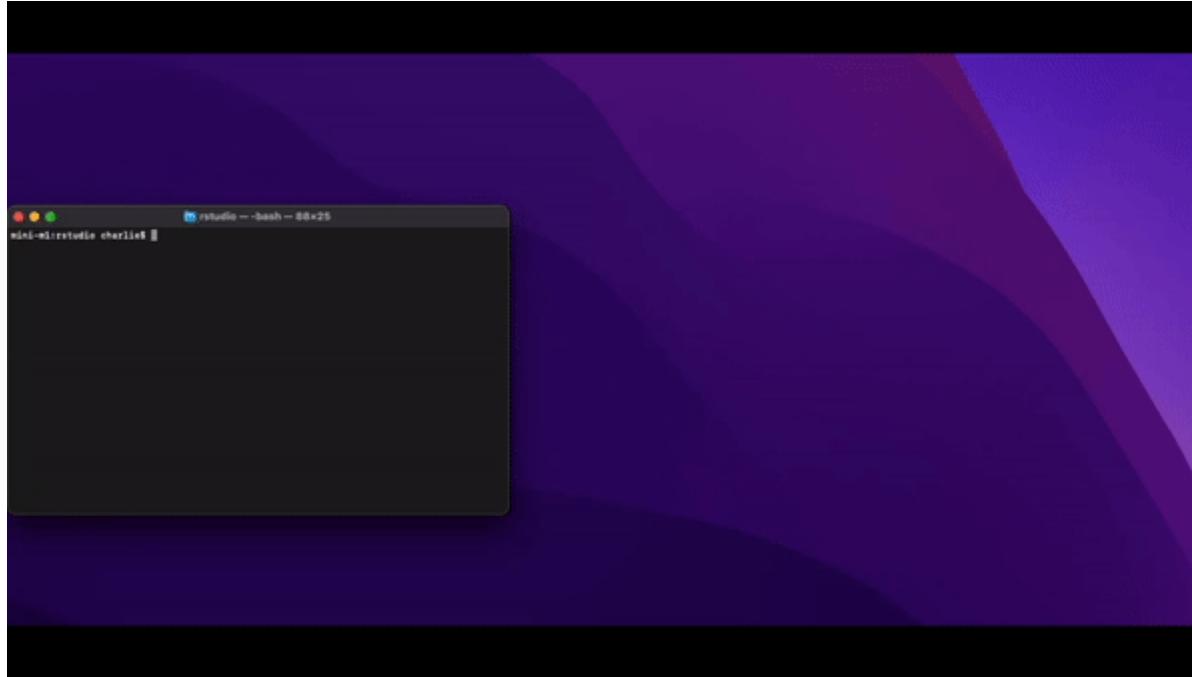
```
1 ./h2_rstudio.sh -u username
```



Tested Platforms

- ✓ Mac's terminal app
- ✓ Window's WSL2
- ✓ MoboXterm
- ✗ GitBash

RStudio Script



This RStudio Script is currently on our [GitHub page](#)

Info on this RStudio Container (1)

- Rstudio container was built using Docker
 - Based on RStudio images from the [Rocker Project](#)
 - Hoffman2 containers located at `$H2_CONTAINER_LOC`
 - RStudio containers are named:
 - `h2-rstudio_X.Y.Z.sif`
 - Where `X.Y.Z` is the R version
- View all available RStudio containers by running

```
1 module load apptainer
2 ls $H2_CONTAINER_LOC/h2-rstudio*sif
```

Info on this RStudio Container (2)

- Separate build of R and
- R packages installed in unique directory
 - `~/R/APPTAINER/h2-rstudio_4.1.0` (for `h2_rstudio-4.1.0.sif`)
- [HPC Container files](#)
 - Docker and definition files for Hoffman2 containers
 - RStudio Dockerfiles have all you need to build RStudio

R Package Installs

- Some R packages require extra libraries or software in the container
- Contact us to update this container
 - OR you can modify the Dockerfile for your own container

Tips for Running RStudio (1)

- If Rstudio does not at start up
 - Possibly due to previous RStudio not shutdown correctly
- Clear out any tmp directories config files

```
1 rm -rf $SCRATCH/rstudiotmp
```

- Clear out RStudio config files

```
1 rm -rf ~/.config/rstudio
```

Tips for Running RStudio (2)

Using Batch R

- Instead of interactive RStudio, you can run R as a non-interactive batch job
 - Use R from inside RStudio container as a qsub job
- Create a job script
 - Load Apptainer
 - use RStudio container with a .R script
 - `apptainer exec`

```
1 #!/bin/bash
2 #\$ -cwd
```

```
3 #$ -o rstudio_batch.out.$JOB_ID
4 #$ -j y
5 #$ -l h_rt=3:00:00,h_data=10G
6 #$ -pe shared 1
7
8 # Load Apptainer module
9 . /u/local/Modules/default/init/modules.sh
10 module load apptainer
11
12 # Run R with a R script, named myRtest.R
13 apptainer exec $H2_CONTAINER_LOC/h2-rstudio_4.1.0.sif R CMD BATCH myRtest.R
```

- Then run this job script

```
1 qsub rstudio_batch.job
```

Summary

- Utilize RStudio Server on Hoffman2
 - Access through on your web browser
 - Applicable to other HPC resource as well
- RStudio can be used interactively or as a non-interactive batch job
- Use the `h2_rstudio.sh` script for easy setup

Thanks and Happy Computing!

Questions? Comments?

- [!\[\]\(5c7d278cafc95e469a28c5619b31ce08_img.jpg\) cpeterson@oarc.ucla.edu](mailto:cpeterson@oarc.ucla.edu)
- Look at for more [Hoffman2 workshops](#) and other OARC workshops



