Running Jupyter Notebook and RStudio on Minerva

Minerva Scientific Computing Environment

https://labs.icahn.mssm.edu/minervalab

Patricia Kovatch
Lili Gai, PhD

Hyung Min Cho, PhD
Eugene Fluder, PhD
Jielin Yu, PhD
Wei Guo, PhD
Kali McLennan

April 05, 2023

Icahn
School of
Medicine at
Mount
Sinai

Outline

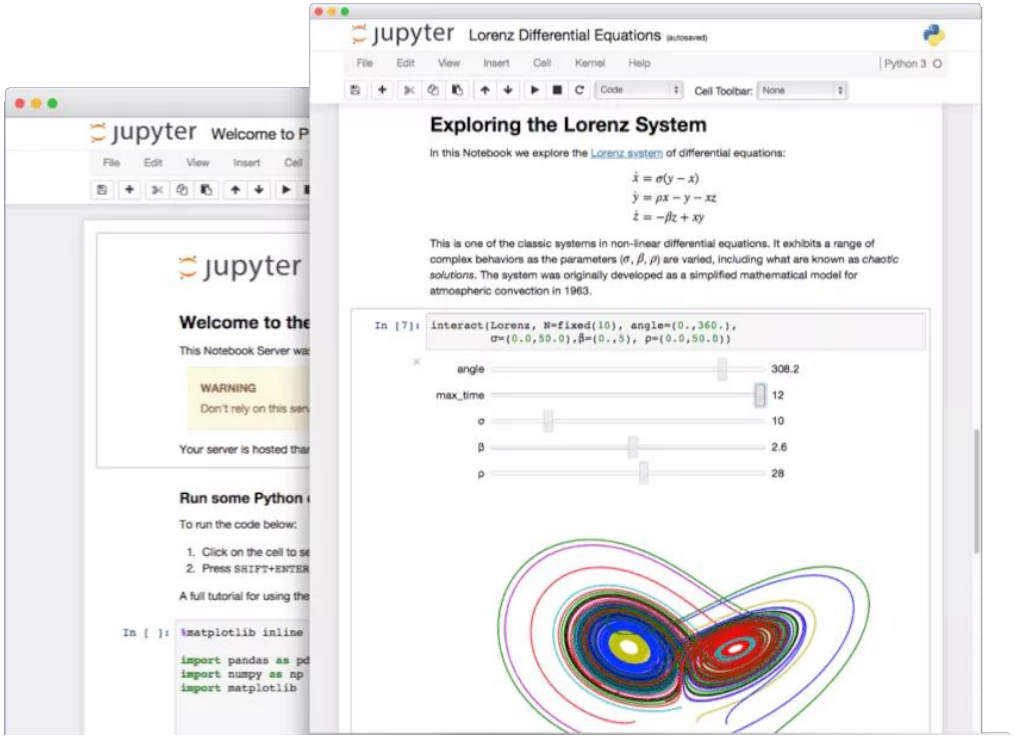
- What is a container
- How to run Jupyter notebook on Minerva
  - Option 1: On-the-fly Jupyter Notebook
  - Option 2: Command line
- How to run Rstudio on Minerva
  - Option 1: On-the-fly Rstudio over Web in a Minerva job
  - Option 2: Run RStudio over GUI (graphical user interface)
  - How to install R packages with On-the-fly RStudio over Web on Minerva
  - RStudio connect server

What is a container?

- **Simply put, a container is a sandboxed process on your machine that is isolated from all other processes on the host machine. A container:**
  - Is a runnable instance of an image.
  - Can be run on local machines, virtual machines or deployed to the cloud.
  - Is portable (can be run on any OS).
  - Is isolated from other containers and runs its own software, binaries, and configurations.
- **Why use containers?**
  - Software developers can build their stack onto whatever operating system base fits their needs best, and create distributable runtime environments so that users never have to worry about dependencies and requirements, that they might not be able to satisfy on their systems.
- **More information:**
  - Docker: https://docs.docker.com/get-started/
  - **Singularity**: https://docs.sylabs.io/guides/3.11/user-guide/introduction.html
  - Minerva training sessions: https://labs.icahn.mssm.edu/minervalab/resources/the-minerva-user-group-and-training-classes/

Jupyter Notebook

The Jupyter Notebook is the original web application for creating and sharing computational documents. It offers a simple, streamlined, document-centric experience. (https://jupyter.org/)



How to Run Jupyter Notebook on Minerva

Option 1 : On-the-fly Jupyter Notebook in a Minerva job

https://labs.icahn.mssm.edu/minervalab/documentation/python-and-jupyter-notebook/#jupyter

- Scripts in /usr/local/bin on head nodes only:
  - **`minerva-jupyter-web.sh`**
    - Runs jupyter installed inside a singularity container
    - No module system setup
  - **`minerva-jupyter-module-web.sh`**
    - Runs jupyter installed in module python/3.7.3
    - Can access modules on minerva
    - Can load users' conda environments (jupyter must be installed in conda env)
- Add --help to the script to get help message/usage
- How to create conda environment on Minerva: https://labs.icahn.mssm.edu/minervalab/documentation/conda/

How to Run Jupyter Notebook on Minerva

Option 2 : Jupyter Notebook in a batch job

- Primarily, the nbconvert tool allows you to convert a Jupyter .ipynb notebook document file into another static format including HTML, LaTeX, PDF, Markdown, reStructuredText, and more. nbconvert can also add productivity to your workflow when used to execute notebooks programmatically.
  https://nbconvert.readthedocs.io/en/latest/usage.html
- Use the commands below in your job script:

```
ml python
jupyter nbconvert --to notebook --execute myfile.ipynb
```

- The result is saved in a file named **myfile.nbconvert.ipynb**

How to Run RStudio on Minerva

Option 1 : On-the-fly Rstudio over Web in a Minerva job

https://labs.icahn.mssm.edu/minervalab/documentation/r/#rstudio

- Scripts in /usr/local/bin on head nodes only:
  - **`minerva-rstudio-web-r4.sh`**
  - ~~**`minerva-rstudio-web-r4.2.1.sh`**~~
  - **`minerva-rstudio-web-r4.2.3.sh`**
- All scripts run RStudio Server installed inside singularity containers
- No module system setup
- Add --help to the script to get help message/usage

Option 2 : Run rstudio over GUI (graphical user interface)

- Enable X11 forwarding
- Run "`ml rstudio proxies; rstudio`"

Install R packages with On-the-fly RStudio over Web

minerva-rstudio-web-r4.sh -h

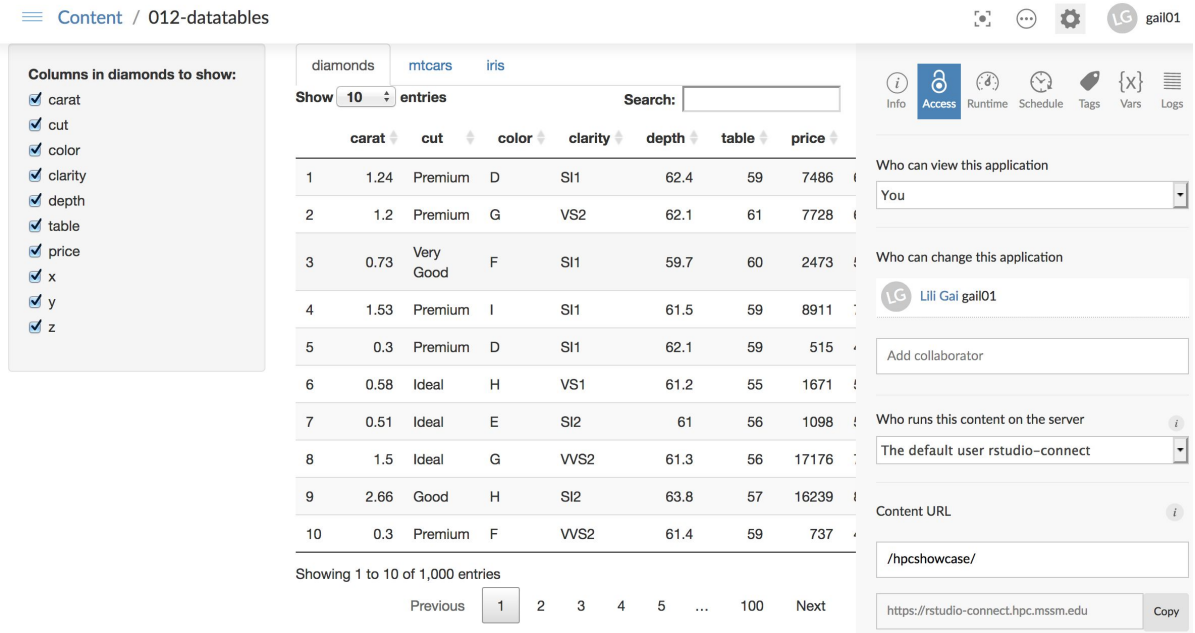
...

```
[INFO] === Package Installations ===
[INFO] To install R packages, do the following in the RStudio web *Shell terminal console*
[INFO] $ export http_proxy=http://172.28.7.1:3128
[INFO] $ export https_proxy=http://172.28.7.1:3128
[INFO] $ export all_proxy=http://172.28.7.1:3128
[INFO] $ export no_proxy=localhost,*.hpc.mssm.edu,*.chimera.hpc.mssm.edu,172.28.0.0/16
[INFO] $ R
[INFO] >>> install.packages(name_of_package)
[INFO] The packages will be installed in your
/hpc/users/yuj25/R/x86_64-pc-linux-gnu-library/R_VERSION
[INFO] If the package is not available in your RStudio Web interface by R
library('name_of_package')
[INFO] You can restart the RStudio job
```

…

RStudio Connect server https://rstudio-connect.hpc.mssm.edu

- You can publish Shiny, R Markdown for collaborators or others
- If interested in publishing on Rstudio-connect, please check instruction at https://labs.icahn.mssm.edu/minervalab/documentation/r/#rstudioconnect
  - Run "`ml rstudio proxies firefox; rstudio`" on a compute node
  - Shiny app example at https://github.com/rstudio/shiny-examples



Acknowledgements

- Supported by the Clinical and Translational Science Awards (CTSA) grant UL1TR004419 from the National Center for Advancing Translational Sciences, National Institutes of Health.

Last but not Least

- Got a problem? Need a program installed? Send an email to:

hpchelp@hpc.mssm.edu