JupyterHub at NCAR
and related things

Davide Del Vento

Wednesday, September 23 2020
10am in Colorado (or somewhere in the world....)
About the speaker

- Davide Del Vento, Ph.D.
  helping users of NCAR’s HPC infrastructure since 2008

- It’s fine to interrupt with questions in context of what I am discussing

https://staff.ucar.edu/users/ddvento
https://www.linkedin.com/in/delvento/
What is JupyterHub

- “The” way to serve Jupyter Notebooks to multiple users

What is a Jupyter Notebook

- A single-user web app to create documents containing **arbitrary, user-provided live code**, equations, visualizations and narrative text.

- Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

Source: [https://jupyter.org/](https://jupyter.org/)
How this differs from old one

- The Jupyter (formerly ipython) Notebook has been available at NCAR since at least early 2015… and it still is...

- But requires a clumsy setup with double-login, ssh-tunneling, etc.

- Yet, that is more robust and so still supported

Doc: https://www2.cisl.ucar.edu/resources/jupyter-and-ipython
Creating your own clone of the NCAR Package Library

Creating a personal clone of the package library in your GLADE space is useful if you want to add or update packages, or develop Python code while using NCAR-provided packages. Simply specify the clone option (e.g., to gear up) when selecting a

Conda environments on JupyterHub

Conda is not installed on the Cheyenne or Casper clusters, but you can install Miniconda in your own user space by following the instructions here.

Miniconda is a lightweight version of the Anaconda distribution that contains the Conda package manager. After installing it, you can create new virtual Conda environments and install packages in them. See this Conda documentation for information on how to create environments and install packages.

Install the ipykernel package in Conda environments that you create. The NCAR JupyterHub will then detect and display your environments in the JupyterLab Launcher with names like "Python [condaenv myenv]."

To activate an environment and install the package, follow this example, in which "myenv" is the name of your Conda environment:

```
conda activate myenv
```

Creating Jupiter kernels from NPL versions

If you wish to use the NCAR Package Library inside of other Jupiter instances on Cheyenne (for example, inside a personal NPL clone, conda environment, or a JupyterHub instance), you can run the following command to produce a user-space kernel from an available NPL version:

```
conda activate NPL
conda env create --name NPL --file ./NPL.yaml
```

Useful documentation

JupyterHub at NCAR

The JupyterHub deployment at CISL manages allows "push-button" access to NCAR's Cheyenne supercomputing resource and the Casper cluster of nodes used for data analysis and visualization, machine learning, and deep learning.

It gives users the ability to create, save, and share Jupyter Notebooks through the JupyterLab interface and to run interactive, web-based analysis, visualization and compute jobs on Cheyenne and Casper. JupyterHub is an alternative to X11 access for interacting with those resources to run jobs as well as for using web-based interactive shell functionality without the need to install or use software such as SSH or PUTTY.

NOTE
This JupyterHub deployment is a test system that is subject to more planned and unplanned downtimes than other resources in the NCAR/CISL high-performance computing ecosystem, such as login nodes. Responding to outages and providing updates to JupyterHub may be delayed while CISL staff support our full production HPC systems.

Getting started
Use your web browser to go to jupyterhub.ucar.edu. Chrome and Firefox are recommended.
Useful documentation

- [https://www2.cisl.ucar.edu/resources/jupyterhub-ncar](https://www2.cisl.ucar.edu/resources/jupyterhub-ncar)
- [https://www2.cisl.ucar.edu/resources/python-%E2%80%93-ncar-package-library#clone](https://www2.cisl.ucar.edu/resources/python-%E2%80%93-ncar-package-library#clone)
- [https://www2.cisl.ucar.edu/resources/conda-environments](https://www2.cisl.ucar.edu/resources/conda-environments)
- [https://www2.cisl.ucar.edu/resources/jupyter-and-ipython](https://www2.cisl.ucar.edu/resources/jupyter-and-ipython)
- Stay tuned: more to come...
JupyterHub @ NCAR

Available NCAR Resources

- Cheyenne Supercomputer
- Casper DAV

This material is based upon work supported by the National Center for Atmospheric Research, a major facility sponsored by the National Science Foundation and managed by the University Corporation for Atmospheric Research. Any opinions, findings and conclusions or recommendations expressed in this material do not necessarily reflect the views of the National Science Foundation.
Where to go from here

- Documentation provides additional details
- Pangeo project provides specific information e.g. for parallel analysis
- Future talk will cover kernel customization

CISL/NCAR documentation: see previous slides
Pangeo: [http://pangeo.io/](http://pangeo.io/)
Future work

● Ability to load modules before kernel startup

● Improve user-friendliness of the login process (e.g. kernel logos overlays)

● More documentation and training (let us know what you need!!)
The end

- Thank you for your attention!
- Please fill a quick survey which will be sent to you
- Do not hesitate to contact user support
- Yes, slides and recording will be provided
- Any more questions?
Backup slides
Logging in JupyterHub

Chrome & Firefox are recommended
Note the highlighted: this is beta
Starting the Jupyter Server
What is a Jupyter Server?

- A program responding to HTTP requests
- For a single user!!!
- The same program you can start yourself “the old way” or on your laptop
- Usually “inaccessible” from the outside
- Needs “something” to become accessible (e.g. ssh-tunnel)
Submitting a PBS (or Slurm) job
Why submitting a job?

- Your node, your resources…
- i.e. others will not step on your toes (& vice versa)
- The memory is yours, if you run out, you have exhausted hw
- CPU is yours, if you run out…
- GPU (if you requested) is yours…
Launch a kernel
What is a kernel?

- The kernel is a program which interprets what you type, execute it in the language of choice, and keep the “session” in scope (so variables are not forgotten at each RETURN).
What kernels can you use?

- Three varieties (that I know of):
  - Provided by NCAR/CISL
    (actually two kinds of these, next slide)
  - Provided by you as mods of CISL’s ones
  - Provided by you as conda environments
How to choose a kernel?

● What programming language would you like to use?

● Often Python, but many more are available (C, julia, matlab..)

● What libraries do you need?
Working with a notebook
Working with the console
Working with the shell
Back to presentation