Using Jupyter at NERSC

New User Training
June 16, 2020

Rollin Thomas
Data and Analytics Services Group
What Is Jupyter?

Interactive open-source web application

Allows you to create and share documents, “notebooks,” containing:

- Live code
- Equations
- Visualizations
- Narrative text
- Interactive widgets

Things you can use Jupyter notebooks for:

- Data cleaning and data transformation
- Numerical simulation
- Statistical modeling
- Data visualization
- Machine learning
- Workflows and analytics frameworks
Why Does NERSC Care About Jupyter Usage?

**Integral part of Big (Data) Science & Superfacility:**
LSST-DESC, DESI, ALS, LCLS, Materials Project, NCEM, LUX, LZ, KBase

**Generational shift in data science:**
UCB’s Data 8 course, entirely in Jupyter
“I’ll send you a copy of my notebook”
Training events adopting notebooks (DL)

**Reproducibility and science outreach:**
Open source code and open science
Jupyter notebooks alongside publications

---

**2017 ACM Software System Award:**
“... a de facto standard for data analysis in research, education, journalism and industry. Jupyter has broad impact across domains and use cases. Today more than 2,000,000 Jupyter notebooks are on GitHub, each a distinct instance of a Jupyter application—covering a range of uses from technical documentation to course materials, books and academic publications.”
Jupyter at NERSC Timeline

- **2013**: Users running IPython via login nodes
- **2014**: Jupyter as a NERSC “science gateway” app
- **2015**: Access to Cori via 1 login node enabled
- **2016**: Transition to Docker-based JupyterHub Deployment
- **2017**: JupyterLab Beta enabled at NERSC
- **2018**: NBViewer, more Cori login nodes, expand compute access

**NERSC Talks, Papers, Posters, and/or Demos:**
- SC16 • CUG17 • JupyterCon17
- IDEAS/ECP • ISC18
- JupyterCon18 • ECP2019
- BlueWaters Webinar • Community Workshop • NUG2019
- NUG VC • SciPy2020

**2020**: Added 2 more login nodes, CPU and GPU compute jobs for Jupyter

**2021**: First JupyterCon

**2019**: Jupyter team receives ACM Software Systems Award
Number of Jupyter Users per Month

OK, How Do I Use Jupyter at NERSC?

Jupyter at NERSC is provided through a JupyterHub deployment we manage:
- Authenticates you (username, password, and OTP)
- Spawns a notebook server for you somewhere at NERSC
- Manages communication between you and your notebook
- Keeps track of and manages your notebook process
- Can provide helpful additional services

https://jupyter.nersc.gov

Authenticate → Choose → Go!
How Do I Choose a Notebook Server to Spawn?

**Cori Shared CPU Node:**
Notebook on cori{13,14,19}
Can see /cfs, $HOME, etc
Can see Cori $SCRATCH
Same Python env as ssh login
Can submit jobs via `%sbatch`

**Spin Shared CPU Node:**
External to Cori, in Spin
Can’t see $SCRATCH
Can’t run jobs
But can see /cfs, $HOME

**Cori Shared GPU Node:**
Notebook on cgpu{01-18}
Like Cori Shared CPU
Runs in a 4h job
Enabled if you have GPU QOS

Shared ⇒ Other users are on the same node as you
NERSC uses JupyterHub’s Services feature
A process that interacts with the Hub's REST API
May perform a specific action or task:
  - Shutting down idle notebook servers (16 hours)
  - Posting announcements on the hub
  - Rendering or sharing notebooks

**Announcement**
Notices about upcoming maintenances
Communication about known issues
(Not a replacement for NERSC MOTD)

**NBViewer (Coming Soon)**
Render a notebook as static HTML
Copy a notebook to your server and start it up
Can copy the kernel used with the notebook
JupyterLab Interface
JupyterLab Interface: NERSC Goodies

Favorites (NERSC/jupyterlab-favorites)

Bookmark your favorite places on the file system

Prepopulate with $HOME and $SCRATCH

Add the current directory by clicking the ★ icon
JupyterLab Interface: NERSC Goodies

Open from Path...
Jump to where you want to go on the file system

Recents (NERSC/jupyterlab-recents)
Recent locations you’ve visited on the file system
Kernels: How You Compute with Jupyter

Your Own Jupyter Kernel

Most common Jupyter question:
“How do I take a conda environment and use it from Jupyter?”

Several ways to accomplish this, here’s the easy one.

$ module load python
$ conda create -n myenv python=3.7
$ source activate myenv
(myenv) $ conda install ipykernel <other-packages>...
(myenv) $ python -m ipykernel install --user --name myenv-jupyter

Point your browser to jupyter.nersc.gov.
(You may need to restart your notebook server via control panel).
Kernel “myenv-jupyter” should be present in the kernel list.

This creates a “kernelspec” file.
The kernelspec File

```
(myenv) rthomas@cori01:~> cat \
   $HOME/.local/share/jupyter/kernels/myenv-jupyter/kernel.json
{
  "argv": [
    "/global/homes/r/rthomas/.conda/envs/myenv/bin/python",
    "-m",
    "ipykernel_launcher",
    "-f",
    "{connection_file}"
  ],
  "display_name": "myenv-jupyter",
  "language": "python"
}
```
```
{
    "argv": [
        "/global/homes/r/rthomas/.conda/envs/myenv/bin/python",
        "-m",
        "ipykernel_launcher",
        "-f",
        "{connection_file}"
    ],
    "display_name": "myenv-jupyter",
    "language": "python",
    "env": {
        "PATH": "...",
        "LD_LIBRARY_PATH": "...",
    }
}
```
Additional Customization

```json
{
    "argv": [
        "/global/homes/r/rthomas/jupyter-helper.sh",
        "-f",
        "{connection_file}\n"
    ],
    "display_name": "myenv-jupyter2",
    "language": "python",
}
```

Meanwhile, in jupyter-helper.sh:
```
#!/bin/bash
export SOMETHING=123
module load texlive
exec python -m ipykernel "$@
```

The helper script is the most flexible approach for NERSC users since it easily enables modules.
A Shifter Kernelspec

```json
{
    "argv": [
        "shifter",
        "--image=continuumio/anaconda3:latest",
        "/opt/conda/bin/python",
        "-m",
        "ipykernel_launcher",
        "-f",
        "{"connection_file}""n],
    "display_name": "my-shifter-kernel",
    "language": "python"
}
```
Debugging Jupyter Stuff

(myenv) rthomas@cori01:~> cat ~/.jupyter.log

[I 2018-03-19 16:00:08.175 SingleUserNotebookApp manager:40] [nb_conda_kernels] enabled, 5 kernels found
[I 2018-03-19 16:00:08.248 SingleUserNotebookApp extension:53] JupyterLab beta preview extension loaded from
/usr/common/software/python/3.6-anaconda-4.4/lib/python3.6/site-packages/jupyterlab
[I 2018-03-19 16:00:08.248 SingleUserNotebookApp extension:54] JupyterLab application directory is
/global/common/cori/software/python/3.6-anaconda-4.4/share/jupyter/lab
[I 2018-03-19 16:00:09.123 SingleUserNotebookApp handlers:73] [nb_anacondacloud] enabled
[I 2018-03-19 16:00:09.129 SingleUserNotebookApp handlers:292] [nb_conda] enabled
[I 2018-03-19 16:00:09.181 SingleUserNotebookApp __init__:35] ✓ nbpresent HTML export ENABLED
[I 2018-03-19 16:00:09.181 SingleUserNotebookApp __init__:43] ✗ nbpresent PDF export DISABLED: No module
named 'nbrowserpdf'
version 0.8.0.rc1
[I 2018-03-19 16:00:09.190 SingleUserNotebookApp log:122] 302 GET /user/rthomas/ →
/user/rthomas/tree/global/homes/r/rthomas? (@128.55.206.24) 0.62ms
[I 2018-03-19 16:00:09.194 SingleUserNotebookApp notebookapp:1445] 0 active kernels
[I 2018-03-19 16:00:09.194 SingleUserNotebookApp notebookapp:1445] The Jupyter Notebook is running at:
[I 2018-03-19 16:00:09.194 SingleUserNotebookApp notebookapp:1445] http://0.0.0.0:56901/user/rthomas/
[I 2018-03-19 16:00:09.194 SingleUserNotebookApp notebookapp:1446] Use Control-C to stop this server and shut
down all kernels (twice to skip confirmation).
/user/rthomas/tree/global/homes/r/rthomas? (@::ffff:10.42.245.15) 0.39ms
Near Future Jupyter Support

**Working on:**
Expanding resources to support Jupyter
Preparing the strategy for Jupyter with Perlmutter
*Streamlining Dask, IPyParallel cluster launch and management*
Custom JupyterLab notebook servers
Expanding JupyterLab interface to:
  - Track and monitor batch jobs
  - New viewers
Jupyter at NERSC

- Go to [https://jupyter.nersc.gov](https://jupyter.nersc.gov) to use Jupyter at NERSC
- Use a kernel-spec to use a conda environment in your notebook
- You can customize those kernelspec files in many ways
- We work on making Jupyter work and work better for you

Always looking for:
New ways to empower Jupyter users
Feedback, advice, and even help:
[https://help.nersc.gov/](https://help.nersc.gov/)
rcthomass@lbl.gov
Thank You and Welcome to NERSC!