

Jupyter in HPC

Andrea Zonca @ SDSC

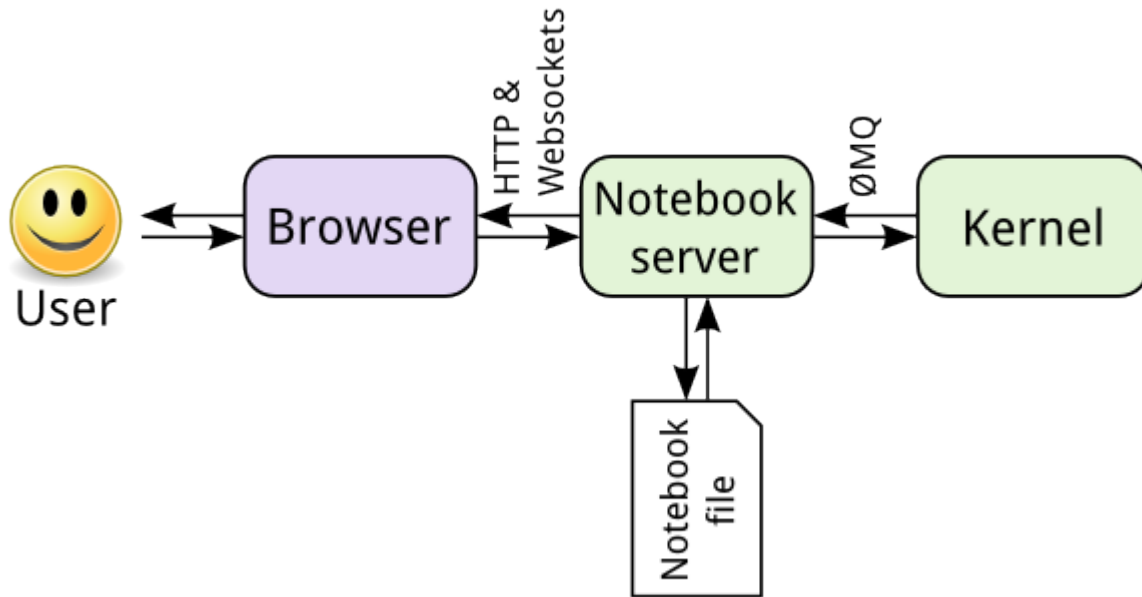
Why the Notebook?

- Code, text, equations, plots in the same document
- Reproducible computation
- Easy to share

Notebook demo

- Post on Nature: <http://www.nature.com/news/interactive-notebooks-sharing-the-code-1.16261>
- or <http://tmpnb.org>

How the Notebook works



ipynb documents

- JSON format
- includes plots in binary format
- easy to convert to .html/.pdf for sharing
- <http://nbviewer.ipython.org>
- run non-interactively with [runipy](#)

Workflows and Gateways

- Self-documenting script
- Node in a workflow
- As a script to be uploaded to a Gateway for more flexibility

Notebook: interactive work

- Best use of Notebook for data exploration
- On HPC to access:
 - data
 - large memory
 - centrally managed environment

Notebook: interactive work

- Notebook distributed architecture suitable for remote computing
- Notebook running on computing node
- SSH-tunnel port to local browser

Jupyterhub

- Multi-user server for Jupyter Notebooks
- Web application:
 - authenticates users
 - spawns Jupyter Notebooks
 - proxies the Jupyter Notebooks port for each user back to the Jupyterhub web application
 - supports plugins for custom Authenticator and Spawner

Jupyterhub / Gordon

- Implemented proof-of-concept plugin RemoteSpawner
- Released on github.com/zonca/remotespawner
- Interfaces with Torque to spawn Jupyter Notebooks on computing nodes as jobs

Demo of Jupyterhub-HPC

See screenshot at: zonca.github.io/2015/04/jupyterhub-hpc.html

How to operate Jupyterhub

- XSEDE or CILogon Authentication - working on this now
- High priority queue?

Thanks

- Feedback very welcome: zonca@sdsc.edu