

Python for HPC

Andrea Zonca - SDSC

Jupyter Notebook

Data exploration in your browser

What is the notebook?

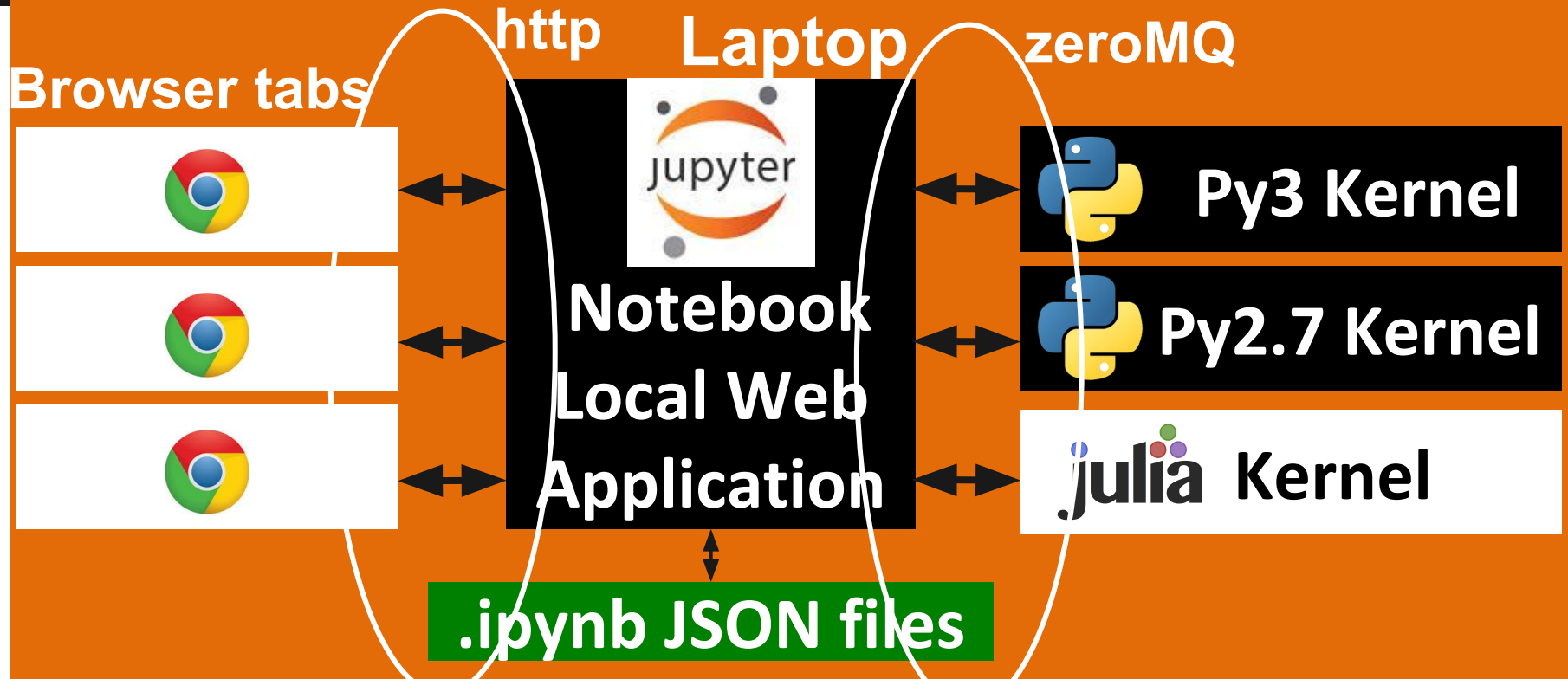
- Browser based interactive console
- Supports multiple sessions in browser tabs
- Each session has a Kernel executing computation
- Saved in JSON format

Notebooks for LIGO

Interactive data analysis of gravitational waves
from black holes merging:

[http://beta.mybinder.org/repo/losc-tutorial/LOSC
Event tutorial](http://beta.mybinder.org/repo/losc-tutorial/LOSC_Event_tutorial)

Jupyter notebook local



Jupyter notebook remote

Laptop



https +
password

Server

Jupyter
Notebook
Web
Application



Py3 Kernel



Py2.7 Kernel



Kernel

.ipynb JSON files

Clone workshop repository

ssh into comet with training account

git clone URL

URL is

<https://github.com/sdsc-scicomp/2018-11-02-comet-workshop-ucr>

Launch notebook job

- `cd python_hpc/`
- `sbatch notebook_singularity.cmd`

connect with browser

Open browser on your laptop and connect to
`cat jupyter-notebook*.out`

New -> Notebook

!hostname

IPython notebook demo

- Python code
- Formatted text
- Equations
- Plots
- Cells execution, cells order
- Clear output

Why the notebook?

- Literate programming: code and explanation together
- Reproducible science: document easily every step
- Easy to share computations: send one single notebook instead of scripts/plots/.doc

ipynb documents

- JSON format
- includes plots in binary format
- easy to convert to .html/.pdf for sharing
- <http://nbviewer.ipython.org>
- Recently rendered automatically on Github

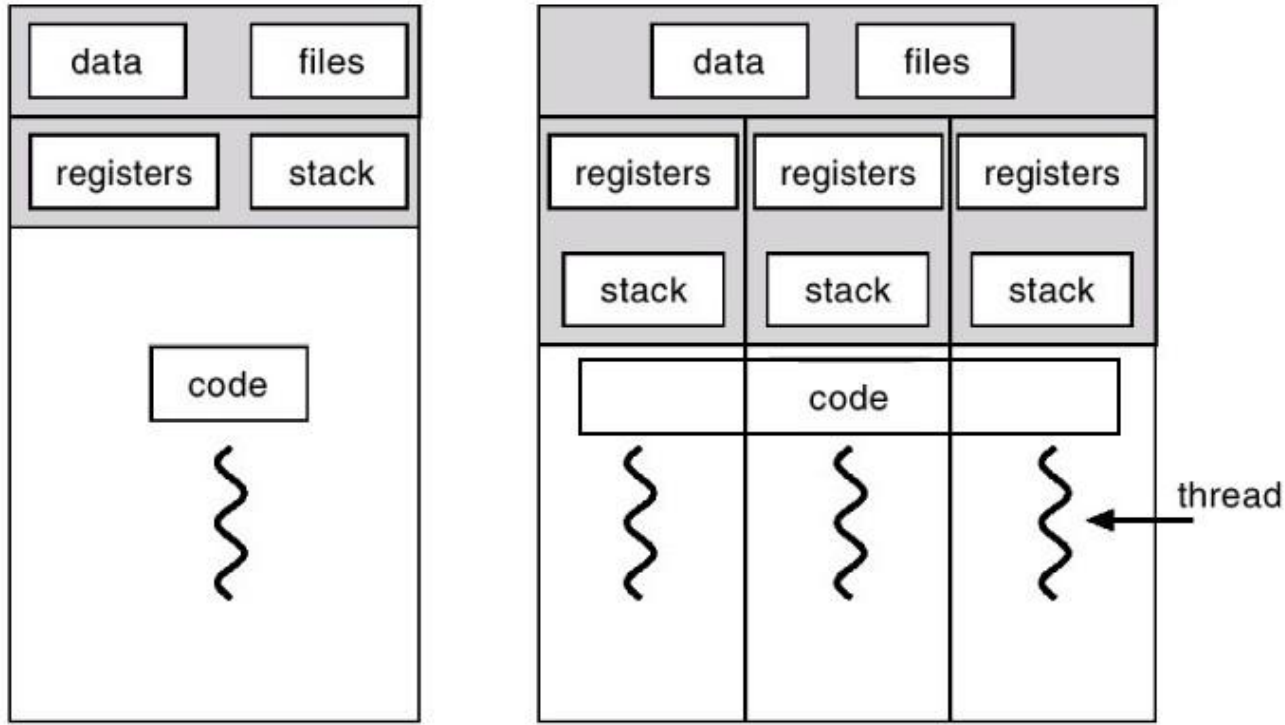
HPC: interactive notebooks

- Analyze large amount of data
- In-situ visualization
- Centralized Python stack
- Check long-running computations
- Prepare and submit batch jobs

Threads and processes



Threads vs processes



threaded

Out of order execution

Some people, when confronted with a problem, think
"I know, I'll use regular expressions." Now they have two problems.



Davidlohr Bueso

@davidlohr

Follow



A programmer had a problem. He thought to himself, "I know, I'll solve it with threads!".
has Now problems. two he

2:16 PM - 8 Jan 2013

4,637 Retweets 1,396 Likes



Numba

Run code on GPU with Python

JIT compiler for Python

- based on LLVM (compiler infrastructure behind clang, Apple's C++ compiler)
- turns Python code into machine code
- on-the-fly