A quick word on the Jupyter project

Jean-Baptiste Poline

UC Berkeley Brain Imaging Center

INCF Nidash TF
The Lifecycle of a Scientific Idea (schematically)

1. Individual exploratory work
2. Collaborative development
3. Parallel production runs (HPC, cloud, ...)
4. Publication & communication (reproducibly!)
5. Education
6. Goto 1

We treat this as a single, coherent problem
A simple and generic architecture

- IPython Kernel
- Terminal console
- Qt Console
- Web Notebook
- ZeroMQ
- Client: monitor, email, publish, ..
Not just about Python: Kernels in any language

- IPython "Official", we ship it.
- IJulia
- IRKernel
- IHaskell
- IFSharp
- Ruby
- IScala
- IErlang
- Lots more! \(\sim 37\) and counting
Not just about Python: Kernels in any language

- IPython "Official", we ship it.
- IJulia
- IRKernel
- IHaskel
- IFSharp
- Ruby
- IScala
- IErlang
- Lots more! ~37 and counting
IPython … Jupyter

- Interactive Python shell at the terminal
- Kernel for this protocol in Python
- Tools for Interactive Parallel computing

- Network protocol for interactive computing
- Clients for protocol
  - Console
  - Qt Console
  - Notebook
- Notebook file format & tools (nbconvert...)
- Nbviewer

Language Agnostic
nbviewer: seamless notebook sharing

- Zero-install reading of notebooks
- Just share a URL
- nbviewer.ipython.org
Changing the scientific culture

http://www.nature.com/news/interactive-notebooks-sharing-the-code-1.16261
Notebook Workflows: The Big Picture

Image credit: Joshua Barratt
**Case study - Question Class Analysis Using IPython Notebook**

**Effort:**

<table>
<thead>
<tr>
<th>Feature</th>
<th>LOC</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPython analysis notebook</td>
<td>20</td>
<td>Python</td>
</tr>
<tr>
<td>Extract relevant experiment results (module “result_proc”)</td>
<td>200</td>
<td>Python</td>
</tr>
<tr>
<td>Total:</td>
<td>220</td>
<td></td>
</tr>
</tbody>
</table>

**Setup:**

- IPython Notebook installed in Python virtual environment
  - Packages: Pandas, NumPy, SymPy, Matplotlib, scikit-learn
- Deployed nbserver for each team member in cluster environment
- Shared file system for direct access to experiment results
- Workers deployed across 10 machines for parallel processing