COLLEGE OF ENGINEERING
NUCLEAR ENGINEERING & RADIOLOGICAL SCIENCES
UNIVERSITY OF MICHIGAN

Lecture 25
HPC + Jupyter Notebooks + Workflows

Prof. Brendan Kochunas

12/4/2019

NERS 590-004

Outline

- Introduction to Jupyter Notebooks
- Running Jupyter on Great Lakes
- Tips & Tricks in Jupyter
- Hands on Demos

Today's Learning Objectives

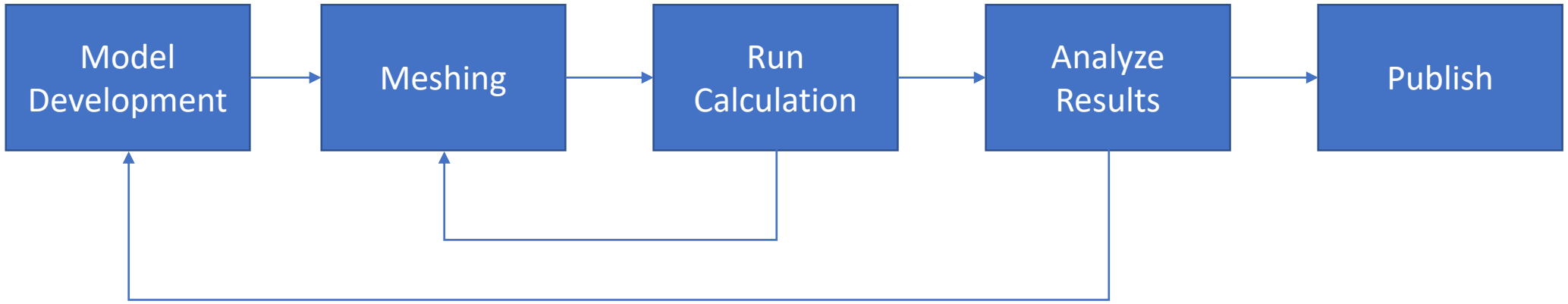
- Know how to run a jupyter notebook on Great Lakes
- Know how to get other code into Jupyter
- Know how to interact with Great Lakes from inside Jupyter

Why should I care about Jupyter Notebooks and HPC?

- As we've learned throughout the semester HPC is challenging to use
  - This is somewhat inherent in the design
- But—does it have to be difficult?
- Other languages (e.g. Python or MATLAB) are much better for *productivity*
- From Lecture 2
  - "***Simulation*** *in the climate community* ***is dominated by the nexus of the Fortran/C/C++ languages*** *and the OpenMP/OpenACC programming models. While both* ***Fortran and C/C++ are interoperable*** *on almost all computing platforms available today, there is a* ***strong desire for other productivity oriented scripting languages*** *in a distributed environment (e.g., Python/pyMPI), as well as the desire to move away from the flat MPI model, which may be functional on the next generation of systems but will lose a factor of 10 to 50 times on* ***GPU-accelerated systems***."

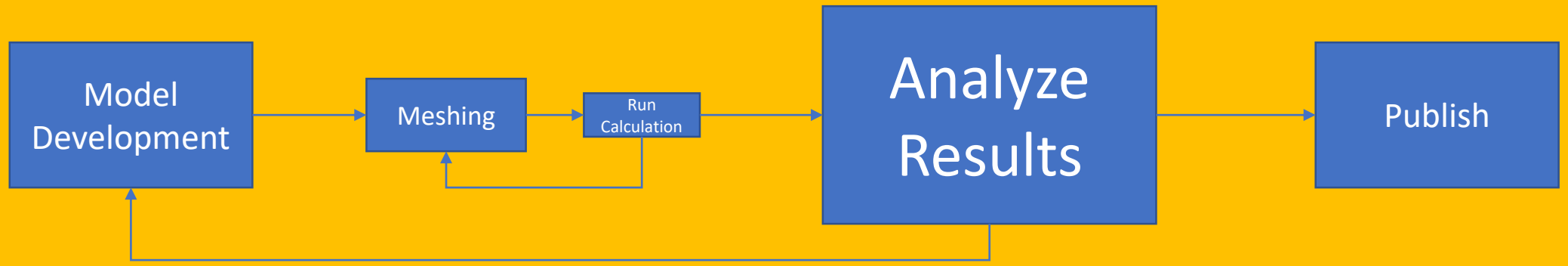
Gerber, R., et. al, "*Crosscut report: Exascale Requirements Reviews,*" *U.S. DOE Office of Science,* Tysons Corner, Virginia, March 9–10, (2017). doi:10.2172/1417653

Typical Modeling and Analysis Workflows



What is the distribution of effort for these tasks?

Typical Modeling and Analysis Workflows



Notebooks can be used to integrate and automate these workflows.

Jupyter Notebooks

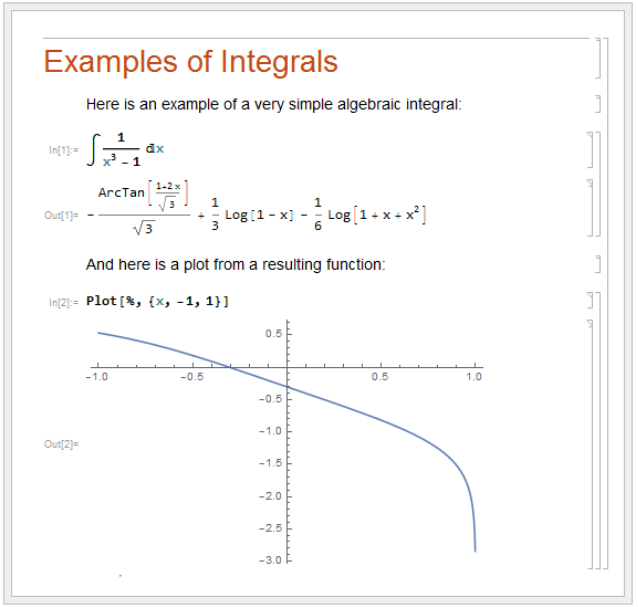
Background

- Jupyter – **JUL**ia, **PYT**hon, and **R**
  - Open Source and developed by Project Jupyter—a nonprofit
- Grew out of the IPython (Interactive Python)
  - Evolved and replaced the IPython notebook
- IPython was designed to utilize parallel computing
  - e.g. task and data parallelism, SIMD, and MIMD
- Key concept is the Notebook Interface

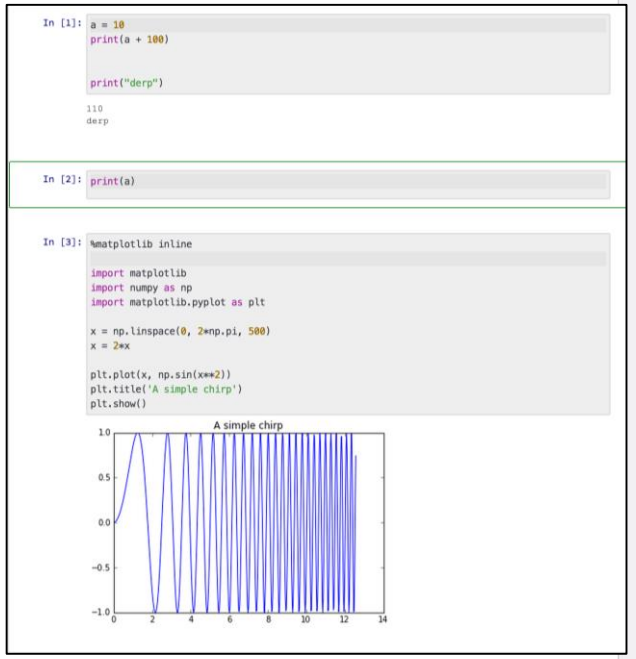
Concept of a notebook

- Notebook is composed of cells
- Cells can be of different types
  - Functional/Evaluatable
  - Text or Documentation (with nice formatting)
- Can be exported to other formats
  - HTML, PDF, etc.
- Not a new concept
  - Existed in Mathematica long before Python
- Excellent for expositions that involve computations

Mathematica



Jupyter Notebook

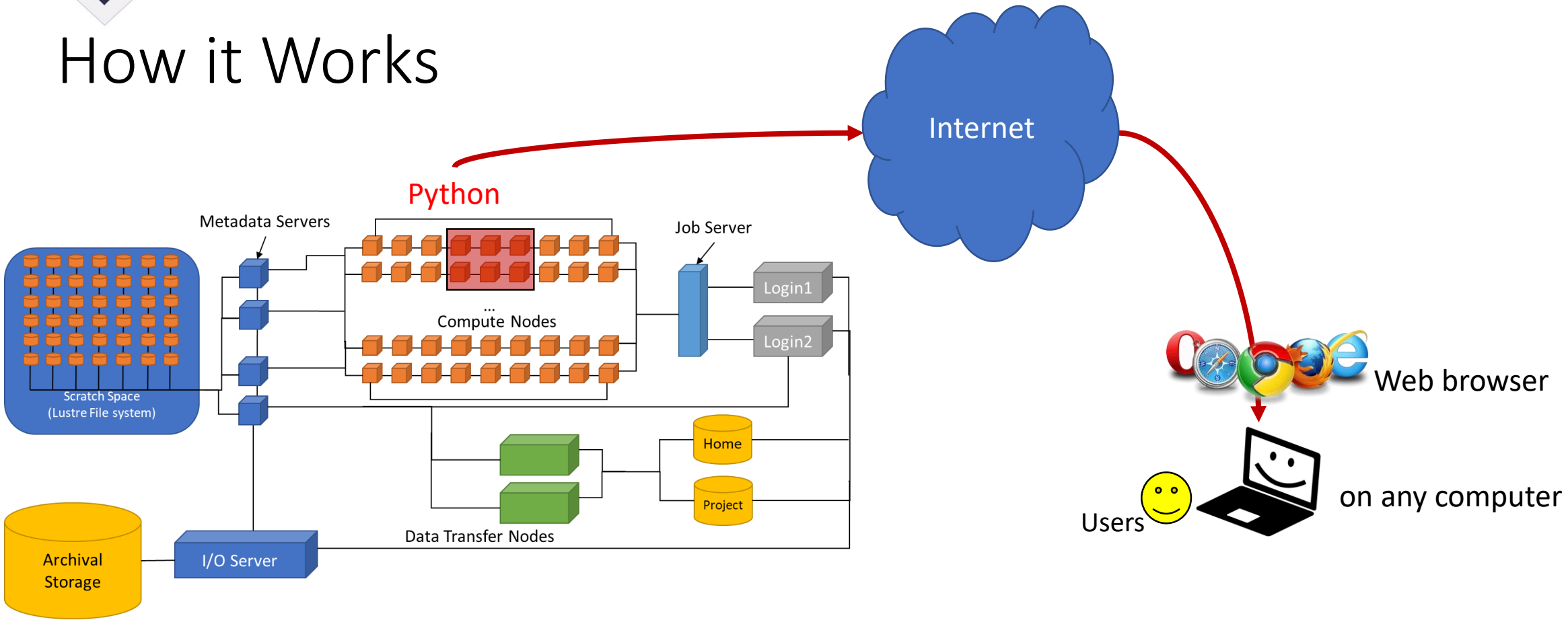


Under the hood

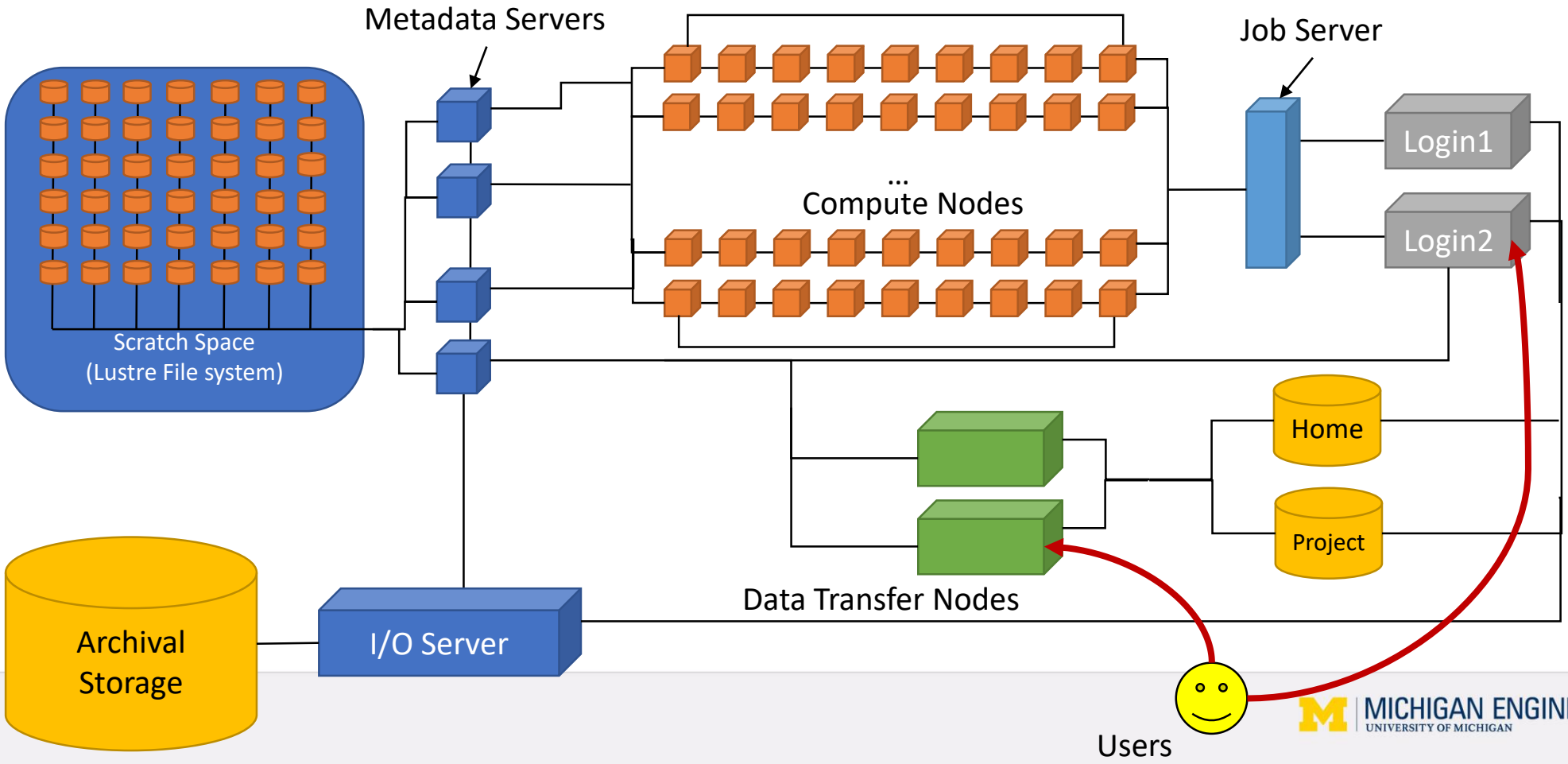
- The IPython Notebook (ipynb) is just a JSON text file
- Suitable for version control with git
- Enhances interoperability with Web applications
  - e.g. Github

```
...
 "metadata": {
  "kernelspec": {
   "display_name": "Python 3",
   "language": "python",
   "name": "python3"
  },
  "language_info": {
   "codemirror_mode": {
    "name": "ipython",
    "version": 3
   },
   "file_extension": ".py",
   "mimetype": "text/x-python",
   "name": "python",
   "nbconvert_exporter": "python",
   "pygments_lexer": "ipython3",
   "version": "3.7.3"
  }
 },
 "nbformat": 4,
 "nbformat_minor": 2
}
```

How it Works


Internet
Python
Metadata Servers
Job Server
Login1
Login2
...
Compute Nodes
Scratch Space
(Lustre File system)
Home
Project
Data Transfer Nodes
Archival
Storage
I/O Server
Web browser
Users
on any computer

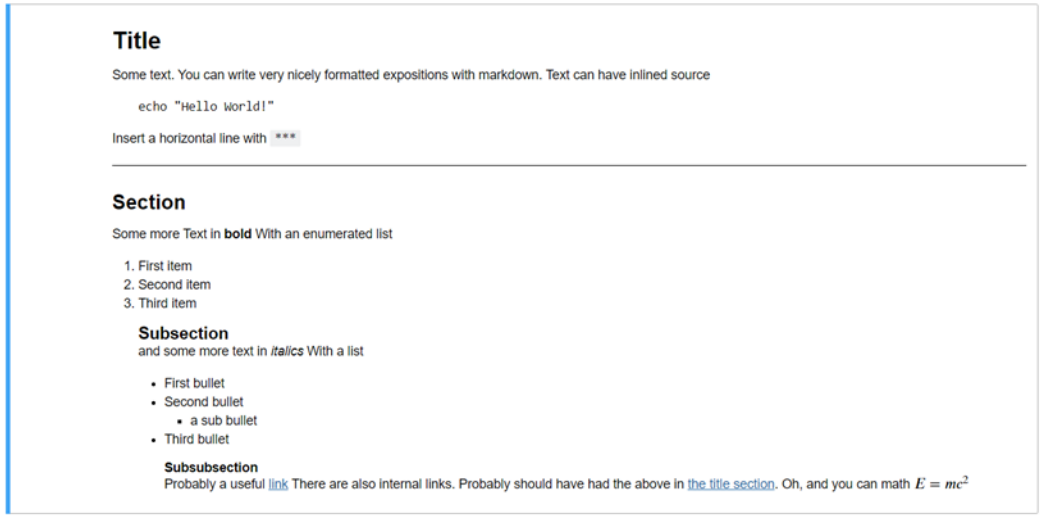
Normal HPC Interface



In-line Documentation with Markdown

- Toggle cell type to markdown with `m`
- Toggle cell back to evaluatable with `y`
- https://markdown-it.github.io/

```
# Title
Some text.
You can write very nicely formatted expositions with markdown.
Text can have inlined source
```
echo "Hello World!"
```
Insert a horizontal line with `***`
***
## Section
Some more Text in __bold__
With an enumerated list
1. First item
2. Second item
3. Third item
### Subsection
  and some more text in _italics_
  With a list
  - First bullet
  - Second bullet
    - a sub bullet
  - Third bullet
#### Subsubsection
Probably a useful [link](https://medium.com/ibm-data-science-experience/markdown-for-jupyter-notebooks-cheatsheet-386c05aeebed)
There are also internal links. Probably should have had the above in [the title section](#Title).
Oh, and you can math $E = mc^2$
```

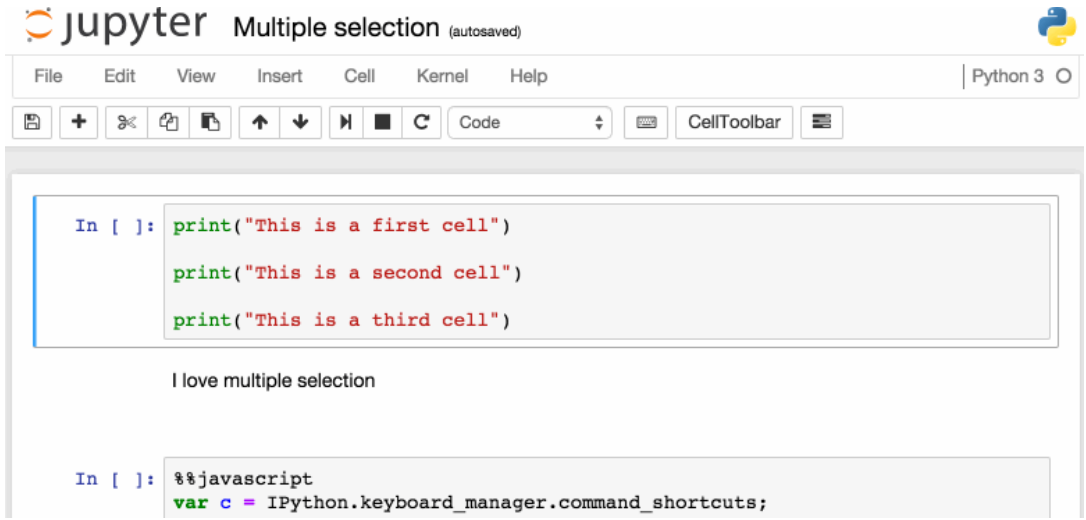


Tips and Tricks

Keyboard Shortcuts

- Command mode and insert mode
  - Just like VIM!
- Adding/removing cells
  - a – insert above
  - b – insert below
  - dd – delete cell (just like VIM!)
- Merging/Splitting Cells
  - Shift + m to merge cells
  - Ctrl + Shift + - to split cell at cursor
- Toggle Output: Esc + o
- Evaluate Current Cell: Ctrl + Enter

Ctrl+Shift+p for Command Prompt



Interacting with the Shell

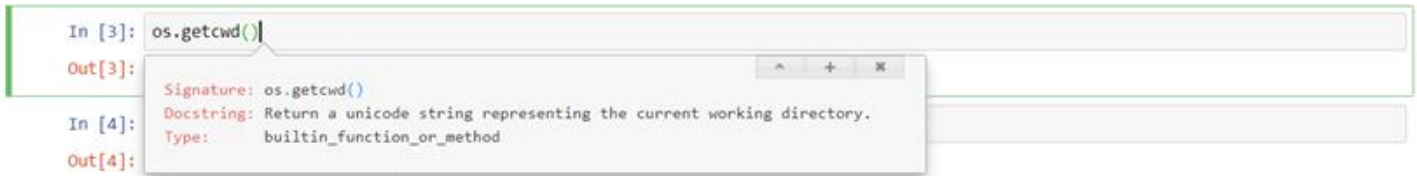
- Shell commands are simply preceded by a "!"

```
In [5]: !ls -l
```

```
total 5941
drwxrwxr-x 12 bkochuna bkochuna     681 Sep 20 10:47 BUILD
drwxr-xr-x  2 bkochuna bkochuna       0 Dec  2 12:34 Desktop
drwxr-xr-x  3 bkochuna bkochuna      37 Dec  2 12:38 Documents
drwxr-xr-x  2 bkochuna bkochuna       0 Dec  2 12:34 Downloads
-rw-rw-r--  1 bkochuna bkochuna     277 Sep 20 10:24 example_job.stdout
drwxrwxr-x  3 bkochuna bkochuna      26 Oct 23 10:22 flux-xfer
drwxr-xr-x  2 bkochuna bkochuna       0 Dec  2 12:34 Music
drwxrwxr-x 14 bkochuna bkochuna     292 Nov 21 15:32 NERS590
-rw-r--r--  1 bkochuna bkochuna   20390 Dec  3 22:40 NERS-590_demo.ipynb
drwxr-xr-x  3 bkochuna bkochuna      22 Dec  2 12:33 ondemand
drwxr-xr-x  2 bkochuna bkochuna       0 Dec  2 12:34 Pictures
drwxr-xr-x  2 bkochuna bkochuna       0 Dec  2 12:34 Public
-rw-rw-r--  1 bkochuna bkochuna    3137 Oct 18 10:35 qr.f90
-rw-r--r--  1 bkochuna bkochuna   12840 Dec  2 12:47 results.h5
drwxr-xr-x 10 bkochuna bkochuna     497 Sep 20 10:46 scalapack-2.0.2
-rw-rw-r--  1 bkochuna bkochuna 4779534 May  2  2012 scalapack-2.0.2.tgz
-rw-rw-r--  1 bkochuna bkochuna      24 Sep 20 10:07 slurm-609955.out
-rw-rw-r--  1 bkochuna bkochuna      24 Sep 20 10:12 slurm-610042.out
-rw-rw-r--  1 bkochuna bkochuna      60 Sep 20 10:17 slurm-610155.out
-rw-rw-r--  1 bkochuna bkochuna      60 Sep 20 10:19 slurm-610159.out
-rw-rw-r--  1 bkochuna bkochuna     277 Sep 20 10:21 slurm-610209.out
drwxrwxr-x  3 bkochuna bkochuna      30 Sep 26 21:29 sw
drwxr-xr-x  2 bkochuna bkochuna       0 Dec  2 12:34 Templates
-rw-rw-r--  1 bkochuna bkochuna     464 Sep 20 10:22 test.sh
drwxr-xr-x  2 bkochuna bkochuna       0 Dec  2 12:34 Videos
```

In-line Help

- Preced command with a “?”
- Keyboard shortcut for showing doc string of previous word: Shift+Tab



Python Magic!

- %load – Load a python file
- %timeit – Time the following command
- %%writefile <filename> – Write cell contents to a file
- %pycat – cat's a file's contents to a popup

Column Selection

- Similar to Sublime Text
- Hold down Alt key and select with mouse

```
In [ ]: x = [
            'one'
            'two'
            'three'
            'four'
            'five'
        ]
```

Hands On Demos

Detailed ARC-TS Instructions: https://arc-ts.umich.edu/greatlakes/user-guide/#document-3

Basically open a browser to: https://greatlakes.arc-ts.umich.edu/

Detailed ARC-TS Instructions

- https://arc-ts.umich.edu/greatlakes/user-guide/#document-3

- Basically open a browser to: https://greatlakes.arc-ts.umich.edu/