



jupyterhub

Jupyterhub as a Playground for Everyone

Loïc Hausammann, ETHZ

loic.hausammann@id.ethz.ch

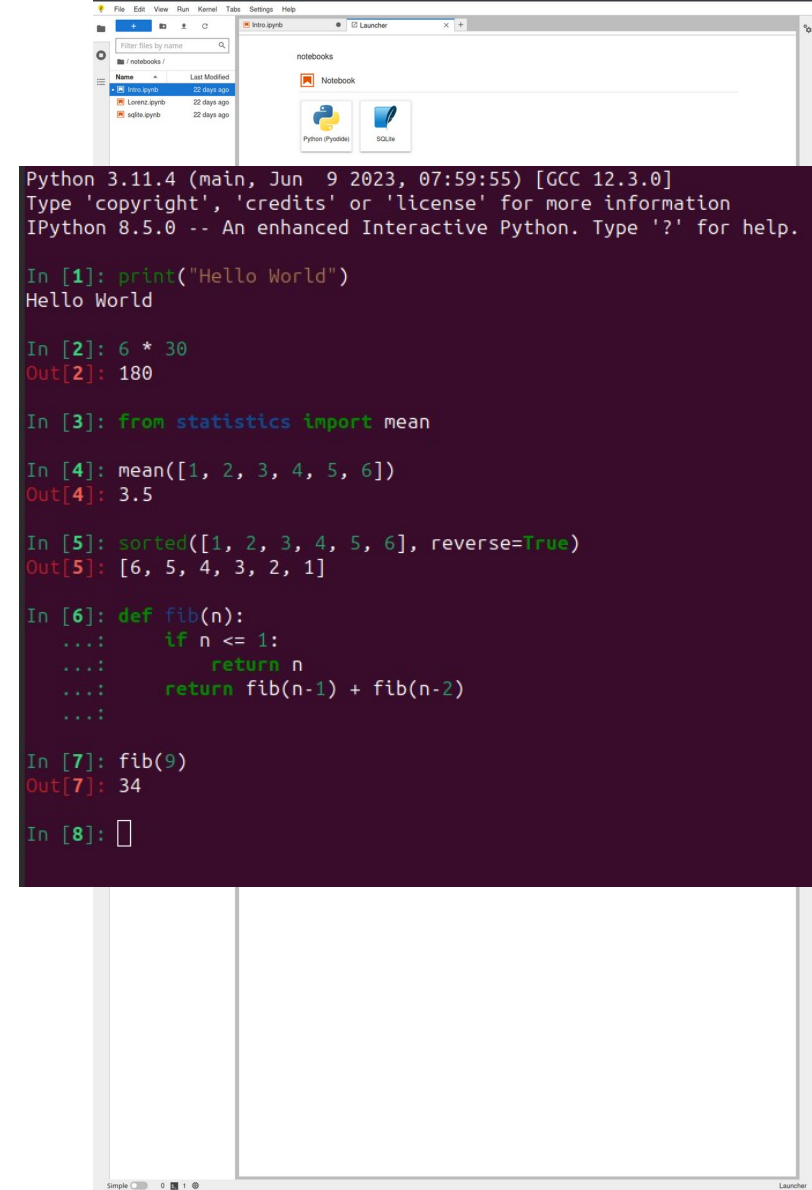
Introduction

- What is Jupyter?
- What is the typical user experience?
- How is it deployed at the ETHZ?
- How to use Jupyterhub as a service provider?



History of Jupyter

- IPython (2011)
- Jupyter notebook (2014)
- First version on Euler@ETHZ (2016)
- Jupyterlab (2018)
- Jupyterhub (2019)



The screenshot displays the Jupyter Notebook environment. At the top, a file browser shows the 'notebooks' directory containing three files: 'Intro.ipynb', 'Lorenz.ipynb', and 'sgtk.ipynb'. Below the file browser, a terminal window shows the Python version (3.11.4) and the IPython version (8.5.0). The terminal contains the following code and output:

```
Python 3.11.4 (main, Jun 9 2023, 07:59:55) [GCC 12.3.0]
Type 'copyright', 'credits' or 'license' for more information
IPython 8.5.0 -- An enhanced Interactive Python. Type '?' for help.

In [1]: print("Hello World")
Hello World

In [2]: 6 * 30
Out[2]: 180

In [3]: from statistics import mean

In [4]: mean([1, 2, 3, 4, 5, 6])
Out[4]: 3.5

In [5]: sorted([1, 2, 3, 4, 5, 6], reverse=True)
Out[5]: [6, 5, 4, 3, 2, 1]

In [6]: def fib(n):
...:     if n <= 1:
...:         return n
...:     return fib(n-1) + fib(n-2)
...:

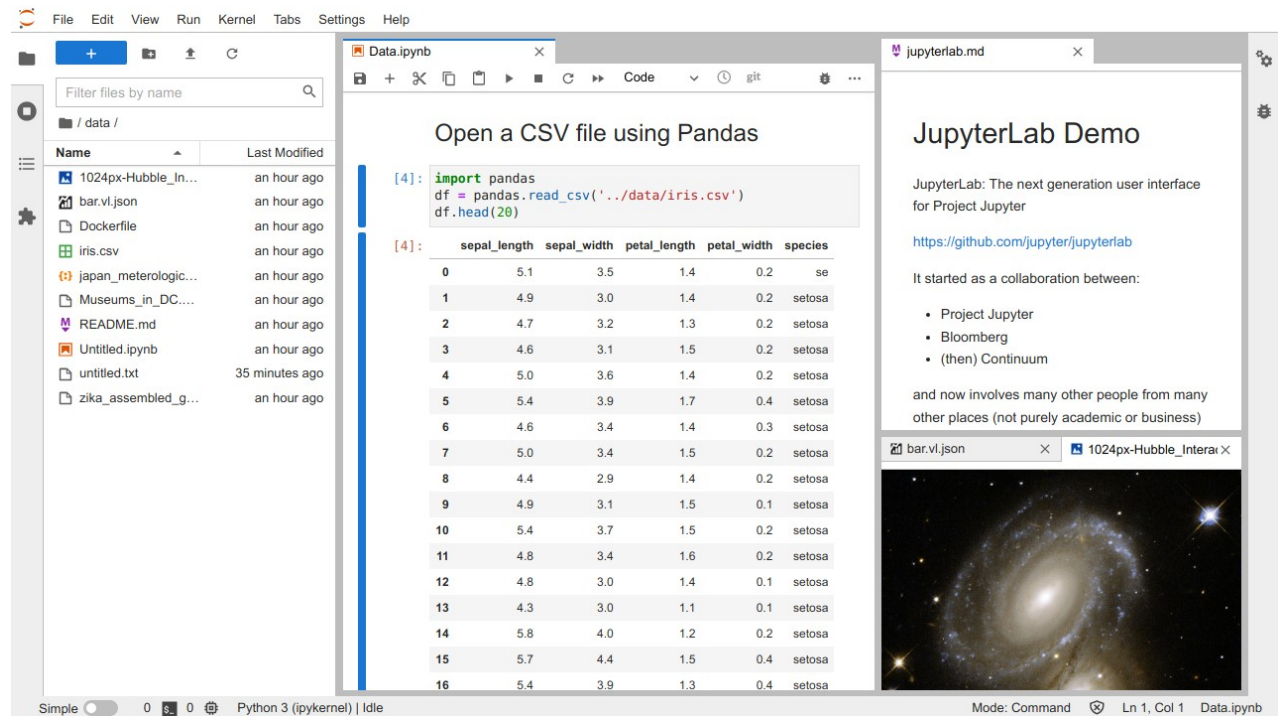
In [7]: fib(9)
Out[7]: 34

In [8]:
```

Below the terminal, a blank Jupyter Notebook interface is visible, showing a large empty area for writing code and a toolbar at the bottom.

What is Jupyterlab?

- Web base user interface:
 - Terminal
 - Python
 - Markdown
 - Text Editor
 - CSV Viewer
 - ...

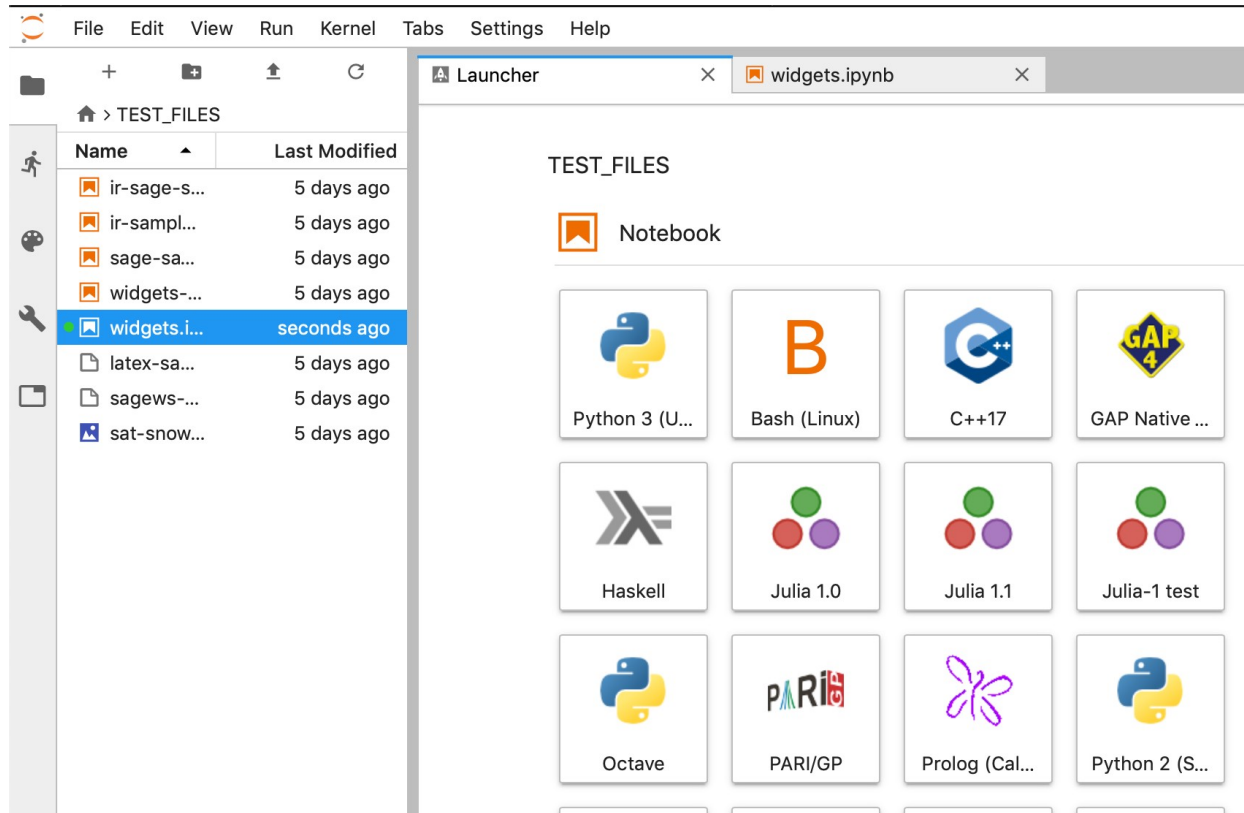


The screenshot displays the JupyterLab web interface. On the left is a file browser showing a directory named '/ data /' with various files including '1024px-Hubble_In...', 'bar.vl.json', 'Dockerfile', 'iris.csv', 'japan_meterologic...', 'Museums_in_DC...', 'README.md', 'Untitled.ipynb', 'untitled.txt', and 'zika_assembled_g...'. The main area is a code editor titled 'Data.ipynb' with the text 'Open a CSV file using Pandas'. It contains two code cells: the first imports pandas and reads a CSV file, and the second displays the first 20 rows of the data as a table. On the right is a markdown viewer titled 'jupyterlab.md' with the text 'JupyterLab Demo' and a list of contributors.

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	se
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
5	5.4	3.9	1.7	0.4	setosa
6	4.6	3.4	1.4	0.3	setosa
7	5.0	3.4	1.5	0.2	setosa
8	4.4	2.9	1.4	0.2	setosa
9	4.9	3.1	1.5	0.1	setosa
10	5.4	3.7	1.5	0.2	setosa
11	4.8	3.4	1.6	0.2	setosa
12	4.8	3.0	1.4	0.1	setosa
13	4.3	3.0	1.1	0.1	setosa
14	5.8	4.0	1.2	0.2	setosa
15	5.7	4.4	1.5	0.4	setosa
16	5.4	3.9	1.3	0.4	setosa

Kernels

- On Euler:
 - Python
 - R
 - Bash
 - Julia
- Users can install their own kernels



The screenshot displays the JupyterLab interface. On the left, a file browser shows the directory structure for 'TEST_FILES'. The file 'widgets.i...' is selected, indicating it is the active notebook. The main area on the right shows the 'TEST_FILES' directory with a 'Notebook' icon. Below this, a grid of kernel options is visible, including Python 3 (U...), Bash (Linux), C++17, GAP Native..., Haskell, Julia 1.0, Julia 1.1, Julia-1 test, Octave, PARI/GP, Prolog (Cal...), and Python 2 (S...).

Name	Last Modified
ir-sage-s...	5 days ago
ir-sampl...	5 days ago
sage-sa...	5 days ago
widgets-...	5 days ago
widgets.i...	seconds ago
latex-sa...	5 days ago
sagews-...	5 days ago
sat-snow...	5 days ago

TEST_FILES

Notebook

Python 3 (U...)

Bash (Linux)

C++17

GAP Native ...

Haskell

Julia 1.0

Julia 1.1

Julia-1 test

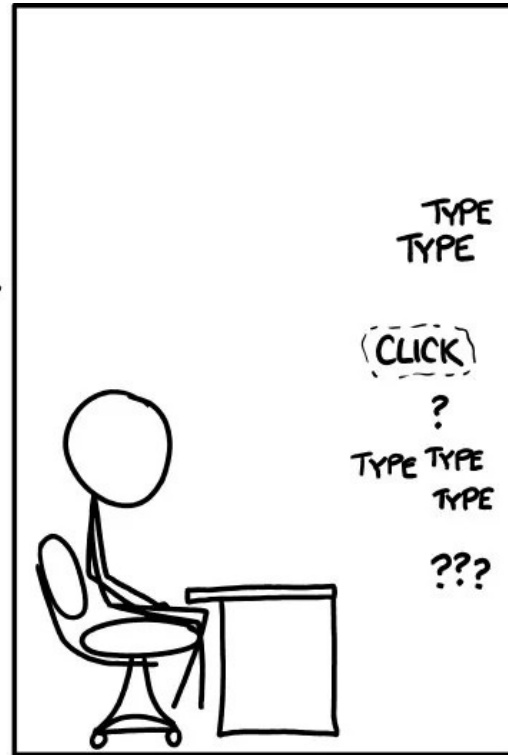
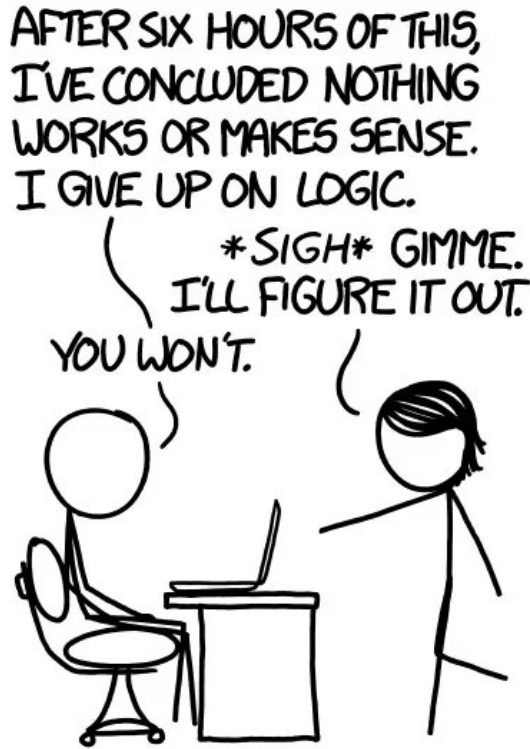
Octave

PARI/GP

Prolog (Cal...)

Python 2 (S...)

User Experience



Euler will have a maintenance between the 24th and 27th October so JupyterHub will stop working during it. Please check our [wiki](#) for more information.

Before being able to login to Jupyterhub, you need to login at least once through SSH.

Please don't forget to [stop your servers](#)

News

Added notebooks for Segmentation (U-Net) Training (beta)

Migrated to Jupyterhub 3.0 (still the same Jupyterlab)

Fix issue in caching the 'Other Software' field

Sign in

Username:

Password:

Sign in

Euler will have a maintenance between the 24th and 27th October so JupyterHub will stop working during it. Please check our [wiki](#) for more information.

Please don't forget to [stop your servers](#)

News

- Added notebooks for Segmentation (U-Net) Training (beta)
- Migrated to Jupyterhub 3.0 (still the same Jupyterlab)
- Fix issue in caching the 'Other Software' field

Server Options

Other Software
JupyterLab (None) ▾

Number of Cores
●
1 ▾

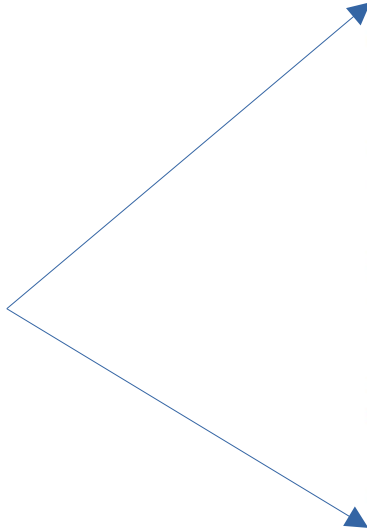
Number of GPUs ⓘ
●
0 ▾

Total Amount of RAM (GB)
●
1 ▾

Run Time In Hours
●
1 ▾

Custom Script to Load ⓘ
Default (depends on the software) (~/.jupyterlabrc) ▾

▶ **Advanced Options**



Euler will have a maintenance between the 24th and 27th October so JupyterHub will stop working during it. Please check our [wiki](#) for more information.

Please don't forget to [stop your servers](#)

News

Added notebooks for Segmentation (U-Net) Training (beta)

Migrated to Jupyterhub 3.0 (still the same Jupyterlab)

Fix issue in caching the 'Other Software' field

Your server is starting up.

You will be redirected automatically when it's ready for you.

Cluster job running... waiting to connect

Event log

Name	Last Modified
mpi-test	2 months ago
nginx	8 months ago
R	a month ago
root	a month ago
scratch	a year ago
slurm	7 months ago
slurm-api	a month ago
swiftsim	8 days ago
system_ser...	10 months ago
tensorboar...	8 months ago
test-multin...	22 days ago
work	3 months ago
abacus_v6...	6 months ago
code-server	5 months ago
fix_permis...	3 months ago
jupyterhub...	6 days ago
jupyterhub...	6 days ago
jupyterhub...	5 days ago
jupyterhub...	5 days ago
jupyterhub...	a minute ago
jupyterhub...	seconds ago
jupyterhub...	7 days ago
jupyterhub...	6 days ago
logs	2 days ago
logs2	2 days ago
loop.sh	2 days ago
myjobs	22 days ago
ref.ipynb	2 months ago
script.sh	9 months ago
slurm-8191...	8 days ago
slurm-8192...	8 days ago
slurm-8193...	8 days ago
slurm-8195...	8 days ago
test	3 months ago

```
Launcher x ref.ipynb x +
Code git Python 3 (ipykernel)

[2]: %matplotlib inline
import matplotlib.pyplot as plt

def mandel_pt(c, MAX_ITER=100):
    """Calculates the number of iterations for point <c> for the Mandelbrot set"""
    iter = 0
    z = 0
    while (iter < MAX_ITER):
        z = z**2 + c
        iter+= 1
        if abs(z) > 2.0:
            break
    return iter if iter!=MAX_ITER else 0

def mandel(width, height, zmin, zmax):
    """Calculates the mandelbrot set for a certain image size (width, height) over a
    # This function maps an (i,j) position of the screen to a complex number inside o
    def map_pos(i, j):
        return zmin + j*(zmax-zmin).real/width + i*(zmax-zmin).imag/height*1j

    # Now return the Mandelbrot set by mapping each point to the corresponding number
    return [[mandel_pt(map_pos(i, j)) for j in range(width)] for i in range(height)]

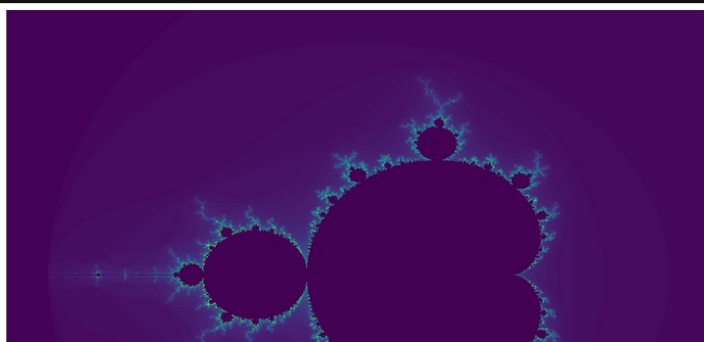
# Width and height of the image
(width, height) = (1024, 768)

# The complex plane to map
(zmin, zmax) = (-2.2-1.5j, +1.2+1.5j)

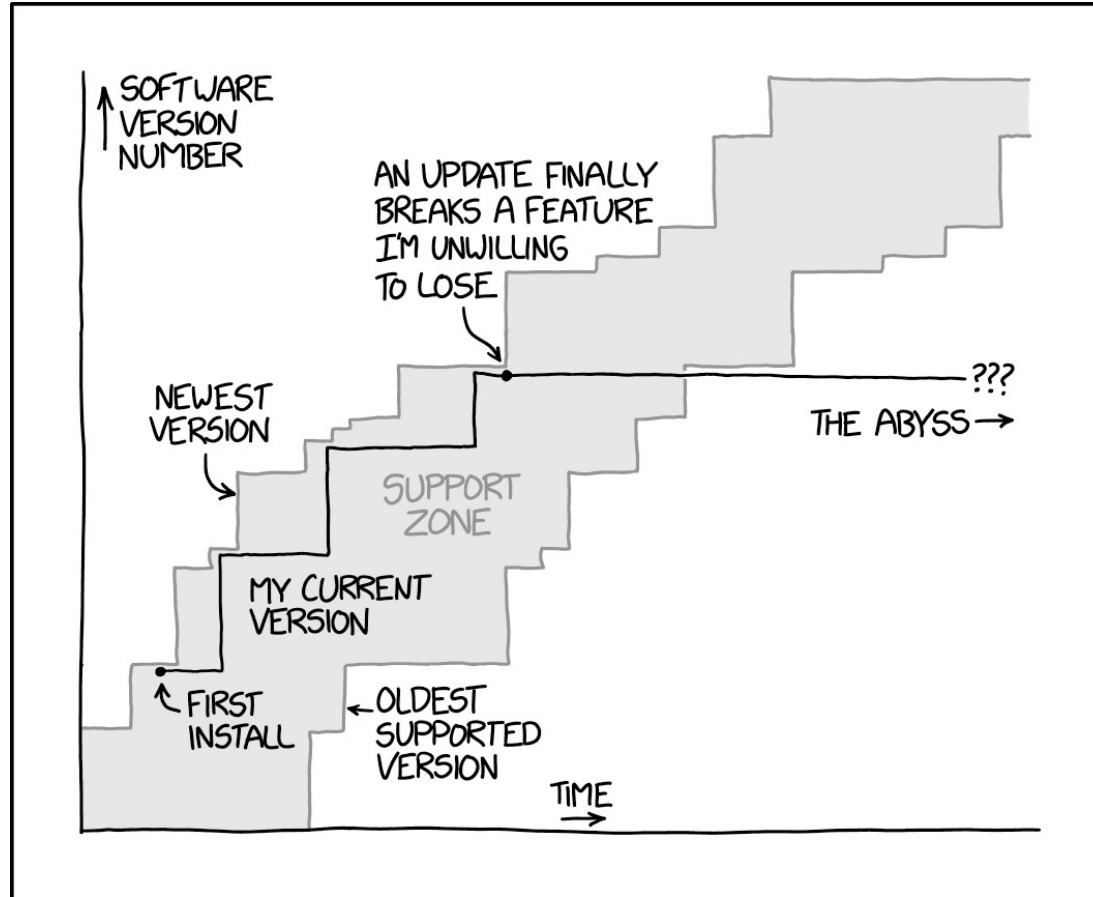
# The Mandelbrot set
mandel = mandel(width, height, zmin, zmax)

plt.figure(figsize=(16,10))
plt.axis('off')
plt.imshow(mandel)

[2]: <matplotlib.image.AxesImage at 0x2ab087c382b0>
```

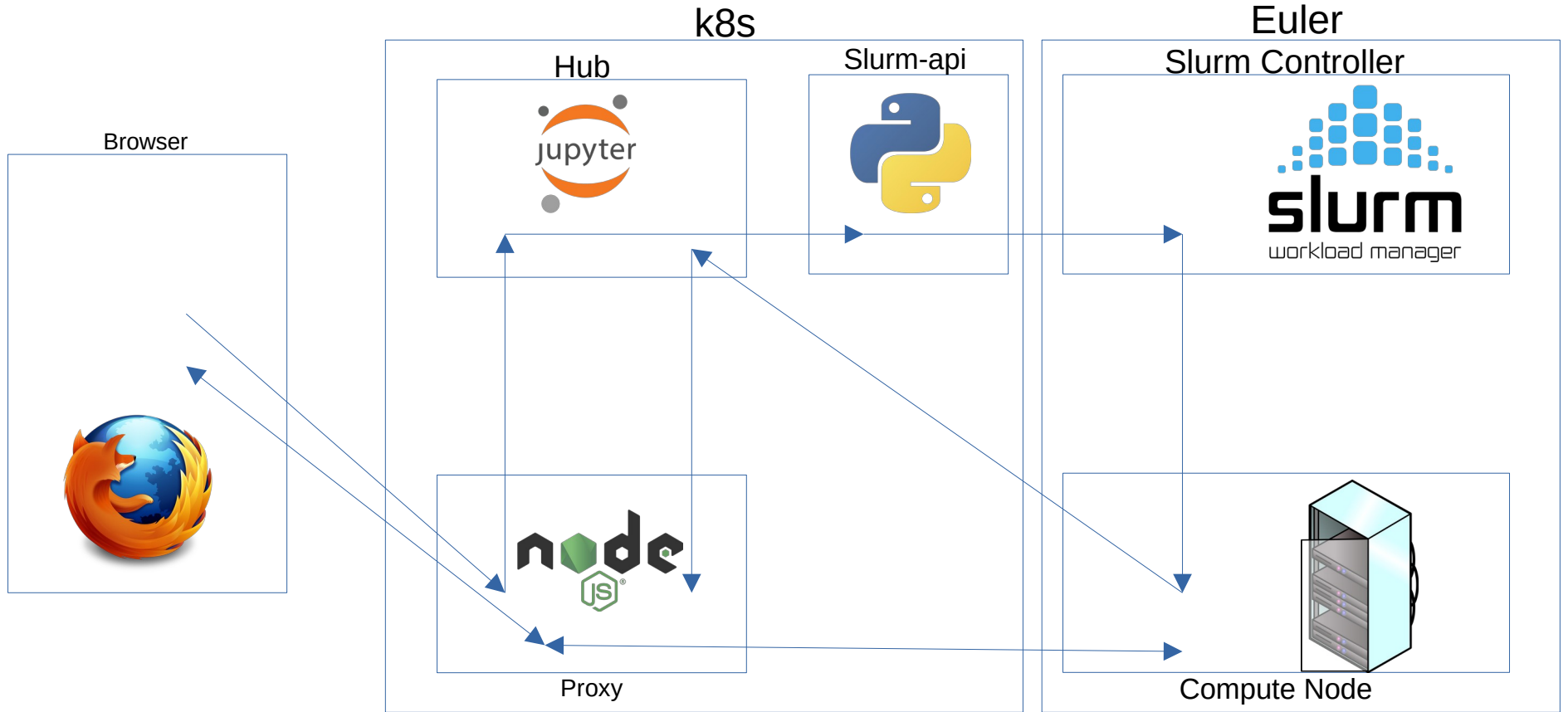


Deployment at ETHZ



ALL SOFTWARE IS SOFTWARE AS A SERVICE.

Architecture on Euler



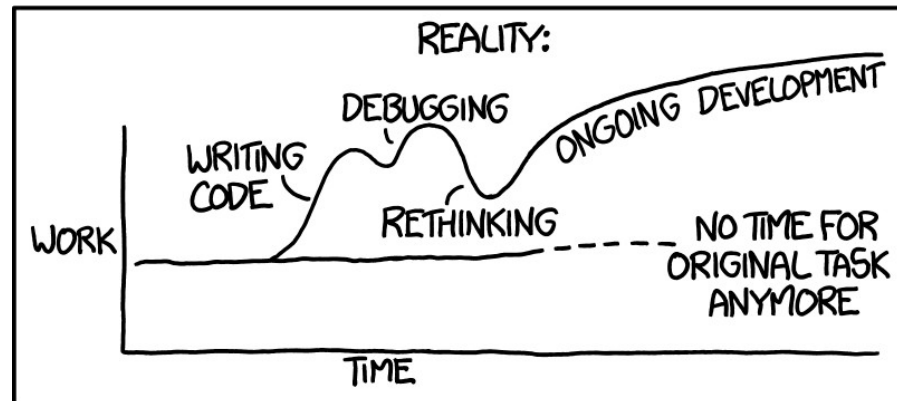
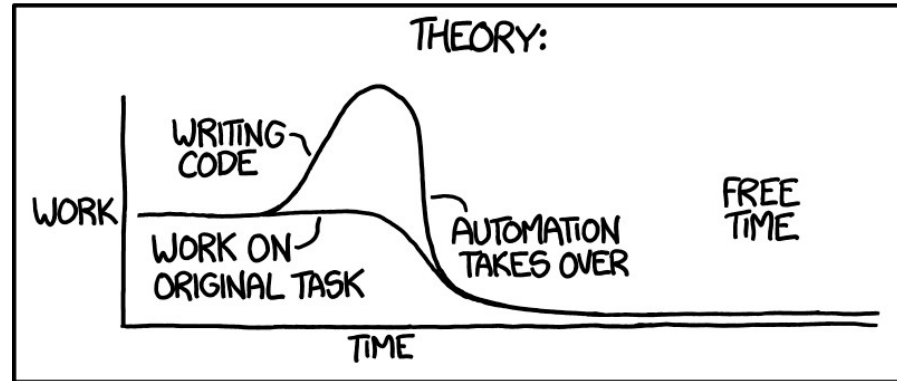
Why Using our own Slurm-api?

- Official REST API lacking endpoints
- Low coupling with Slurm
- Security designed for our need (defined by endpoints)
- Based on python so no hard requirements on the hardware
- Available online:
<https://gitlab.com/ethz-hpc/slurm-api/slurm-api>



Jupyter as a Service Provider

"I SPEND A LOT OF TIME ON THIS TASK.
I SHOULD WRITE A PROGRAM AUTOMATING IT!"



Initial Design

- Goal: Minimizing the user support for installing libraries in Jupyterlab
- Solution: Providing a way for the users to manage their own environment

```
if [ -f {{script}} ];  
then  
  echo Reading {{script}}  
  . {{script}}  
else  
  echo {{script}} does not exist \  
  skipping it  
fi
```



Euler will have a maintenance between the 24th and 27th October so JupyterHub will stop working during it. Please check our [wiki](#) for more information.

Please don't forget to [stop your servers](#)

News
Added notebooks for Segmentation (U-Net) Training (beta)
Migrated to Jupyterhub 3.0 (still the same Jupyterlab)
Fix issue in caching the 'Other Software' field

Server Options

Other Software
JupyterLab (None) ▾

Number of Cores
●
1 ▾

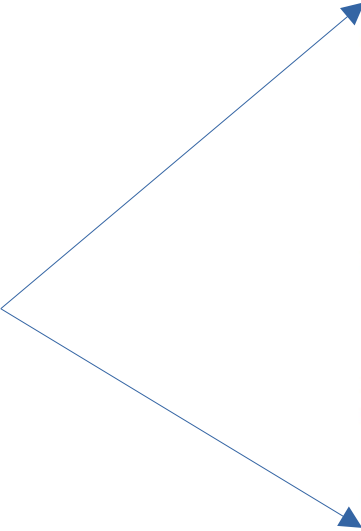
Number of GPUs ①
●
0 ▾

Total Amount of RAM (GB)
●
1 ▾

Run Time In Hours
●
1 ▾

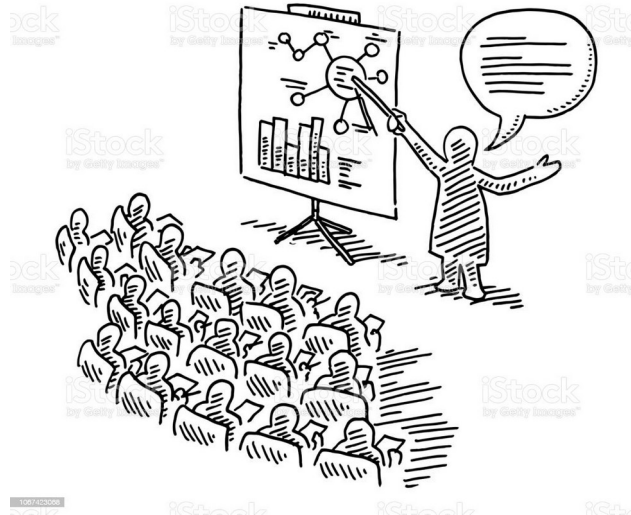
Custom Script to Load ①
Default (depends on the software) (~/.jupyterlabrc) ▾

▶ **Advanced Options**



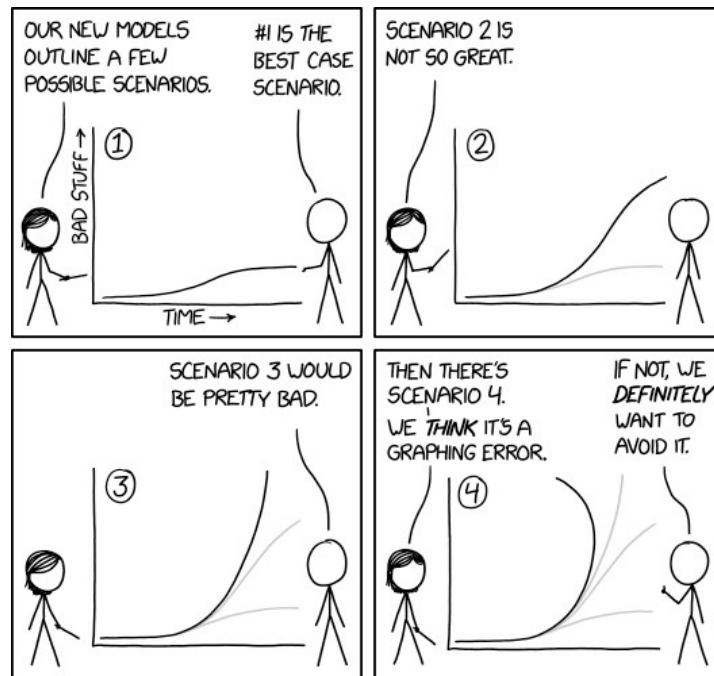
First Consequence

- Why not providing some standard notebooks?
 - Notebooks for the researcher (e.g. Alphafold)
 - Notebooks for workshops and lectures



Second Consequence

- Why not using the user customization for providing something else?
 - Tensorboard
 - Code-server
 - Rstudio
 - ???



Code-Server


- New Issues:
 - Cannot run with a different root path
- Solution:
 - Start a proxy at the same time than code-server



Tensorboard

- New Issues:
 - HTTPS is not provided
 - Need to get the right directory
- Solutions:
 - HTTPS solved with the proxy
 - Start a file explorer server before running tensorflow



 `/home/`

Use current directory

Show hidden directories

`ansys/`

`batchspawner/`

`benchmarks/`

`gpu-stress-test/`

`haproxy/`

`hpc/`

`hybrid/`

`jif3d/`

`julia/`

`mixcr/`

`mpi-test/`

`nginx/`

`R/`

`root/`

`scratch/`

`slurm/`

`slurm-api/`

`swiftsim/`

`system_services/`

`tensorboard_logs/`

`test-multinodes/`

`tickets/`

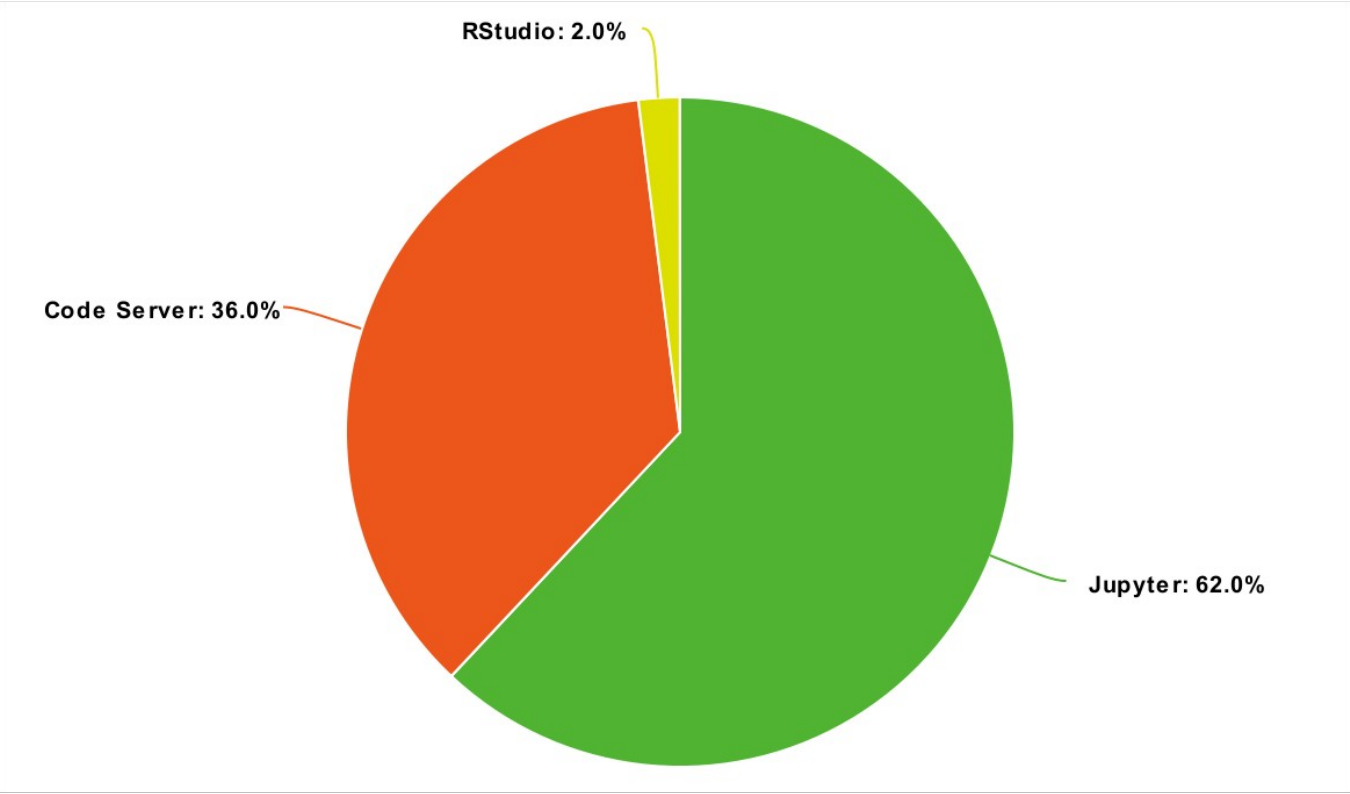
`work/`

RStudio

- New Issues:
 - Requires running with the root account
- Solution:
 - Use Singularity / Apptainer to run it in a container



Statistics over one Week



■ Jupyter ■ Code Server ■ RStudio

Conclusion

- Jupyterhub is a great tool to provide a simple access to a cluster
- It can be largely extended in order to provide many more services
- It can also provide standard tools to everyone

Questions?

- Most of the code is open source
- I can provide the remaining parts on request



loic.hausammann@id.ethz.ch