### JupyterHub A solution for running remote technical workshops

Ludmila MARIAN, 14 December 2022

### Outline

- Problem description and possible solutions
- What is Jupyter Notebook and JupyterHub
- JupyterHub setup for the ICTP Workshop

#### Joint ICTP-IAEA Workshop on Atomic Processes in Plasmas: Data-driven Research

13 - 17 December 2021 **An ICTP - IAEA Virtual Meeting** Trieste, Italy

This School will assist Ph.D. students and other earlystage career researchers to develop their skills and understanding in the application of modern data science techniques to the calculation and evaluation of data relevant to the physics and behaviour of plasmas.

#### **Description:**

The properties of laboratory, industrial, • Plasma emission modelling astrophysical and fusion energy plasmas are • Practical exercises using FLYCHK and the NIST determined by numerous atomic and molecular processes. Modelling such plasmas requires a • The principles of Molecular Dynamics large amount of data concerning its collisions, spectroscopic properties and plasma-material • Bayesian Inference for the determination of interactions, not all of which is directly accessible by experiment. Plasma physicists therefore turn • Demonstration and practical exercises using to calculations and approximations to complete their data needs.

The last 10 years have seen a huge increase • Managing atomic and molecular data in the amount of computing power available • Linux command line usage; Using Python to run the codes that calculate these data. Accompanying this expansion has been the development of so-called data science techniques which unify statistical methods, data analysis and machine learning.

This School will provide training for and knowledge-transfer between scientists working in the fields of computational plasma physics and data science. Participants are expected to have some experience with basic programming in a language such as Python.

How to apply:

Online application: http://indico.ictp.it/event/9657/

Female scientists are encouraged to apply.

### 

#### **Topics:**

- LIBS database
- modelling
- LAMMPS
  - data resources

#### **Registration:**

There is no registration fee.

**Further information:** http://indico.ictp.it/event/9657/ smr3638@ictp.it

plasma-material interaction parameters

Accessing, assessing and exploiting online

and the JupyterLab environment for research

#### **Directors:**

C. HILL, IAEA, Austria

- L. MARIAN, IAEA, Austria
- M. BAUTISTA, Western Michigan University, USA H.-K. CHUNG, Korea Institute of Fusion Energy (KFE),
- Republic of Korea
- Y. RALCHENKO, National Institute of Standards and Technology (NIST), USA
- U. VON TOUSSAINT, Max Planck Institute for Plasma Physics, Germany

#### Local Organiser:

G. THOMPSON, ICTP, Italy

#### **Speakers:**

- M. BAUTISTA, Western Michigan University, USA
- M.-L. DUBERNET, Paris Observatory PSL, France
- C. HILL, IAEA, Austria L. MARIAN, IAEA, Austria
- Y. RALCHENKO, National Institute of Standards and Technology (NIST), USA
- E. STAMBULCHIK, Weizmann Institute of Science, Israe
- **U. VON TOUSSAINT, Max Planck Institute for Plasma** Physics, Germany





**(CTF** 

#### 6 December 2021

e Abdus Salam International Centre ٢ for Theoretical Physics www.ictp.it Trieste, Italy

### **Problem** How to effectively run virtual technical workshops?



Each computer is the *same* (Operating System and Software Tools)

Installation problems can be checked on the spot and *fixed* 



Each computer is *different* (Operating System, Version, Dependencies, Software Tools)

Installation problems are *impossible* to troubleshoot

# **Options for running remote workshops**

- Postpone them until they can be run inperson
- 2. Give participants remote access to computers
  - A. Physical computers (the ones existing at the ICTP headquarters)
  - B. Virtual computers (in the public cloud AWS, Azure, Google)

#### 3. Use JupyterHub on the ICTP infrastructure

- not an option for the IAEA
- participants would need basic knowledge of using the chosen Operating System;
- access to Terminal or Remote Desktop;
- limited ways of transferring local files;
- running Windows instead of Linux;
- no admin rights for organizers;
- we needed several CPUs per participant, thus very expensive;
- laborious installation

# Using JupyterHub on the ICTP infrastructure

- Used via the web browser no need for participants to install anything extra • Free of charge - the ICTP offered us a cluster where they can install the
- JupyterHub
- No system administration tasks they would be handled by the ICTP IT administrators; we would only be responsible with the installation of the specific software for the workshop
- Modern tool for data science participants would become familiar with a modern tool used extensively in data science



# Jupyter Notebook / JupyterHub



Free software, open standards, and web services for interactive computing across all programming languages



## Jupyter: the Basics

- Open Source tool for collaborating.
- Write & Share from the Browser: code, text, equations, graphs.
- Makes your work more transparent, understandable, reproducible, and shareable.
- Support for: Julia, Python, R + ~50 languages.



### Jupyter: the Interface

- Sections
  - File Browser: folders or files available
    - Create new
    - Rename, Move, Download, Delete
    - Upload
  - Launcher
    - Create new notebooks, files
    - Start new terminals
    - Drag-and-drop tabs to arrange them as you wish
  - Running Kernels and Terminals



File Edit View Run Kernel Tabs Settings Help

 $\smile$ 

0

 $\equiv$ 

### Jupyter: the Interface

#### Sections

- File Browser: folders or files available
  - Create new
  - Rename, Move, Download, Delete
  - Upload
- Launcher
  - Create new notebooks, files
  - Start new terminals
  - Drag-and-drop tabs to rearrange them as you wish
- Running Kernels and Terminals



### Jupyter: the Interface

#### • Sections

- File Browser: folders or files available
  - Create new
  - Rename, Move, Download, Delete
  - Upload
- Launcher
  - Create new notebooks, files
  - Start new terminals
  - Drag-and-drop tabs to arrange them as you wish
- Running Kernels and Terminals



### Jupyter: Create a new notebook



for it, delete it.

## Jupyter: the Cell





Run Cell Run a cell content, stop the run, restart the kernel, restart and rerun all the notebook.

#### Type of Input

Select the type of input a specific cell will have: code, markdown, text.

#### Kernel

Shows the kernel chosen for this notebook.

## **Jupyter: the Kernel**

- Behind every notebook **runs** a kernel.
- Kernel state is persistent between cells and over time.
- The order of execution is stated in the label of each cell (be careful if you edit previous cells).
- Useful reset options in the menu:
  - **Restart Kernel** clear all variables
  - **Restart Kernel and Clear** clear also all the output
  - Restart Kernel and Run All clear everything and rerun all cells

### Jupyter: Markdown

G	≣ readme.md ×	♥ readme.md ×
OPEN TABS Close All	₩ C	
E readme.md readme.md KERNELS Shut Down All readme.md TERMINALS Shut Down All	<pre>### ICTP-IAEA Workshop ## Notebooks Let's **explore** some of the features of *JupyterLab Notebooks*. For more details check the [documentation](https://jupyterlab.readthedocs.io /en/stable/). Let's start by generationg some `random numbers`: ```python import random randlist = random.sample(range(1,100), 10) print(randlist) '`` Python 3.8.10 (default, Sep 28 2021, 16:10:42) Type 'copyright', 'credits' or 'license' for more information IPython 7.27.0 An enhanced Interactive Python. Type '?' for help. []: import random randlist = random.sample(range(1,100), 10) print(randlist) [93, 57, 99, 80, 12, 84, 65, 62, 8, 49] []:</pre>	<pre>ICTP-IAEA Workshop ICTP-IAEA Workshop Action and a statement and a statem</pre>

### Jupyter: some useful Features

• Insert LaTex in the code cells

 Run command line commands from inside the notebook  $sz = \arc{a^2 + b^2}$ 

$$z = \sqrt{a^2 + b^2}$$

#### [15]: !ls -lrth

total 220K

-rw-r--r- 1 smr3638 grpsmr3638j 21K Dec 10 19:21 output.png -rw-r--r- 1 smr3638 grpsmr3638j 584 Dec 10 19:41 text.csv drwxr-xr-x 1 smr3638 grpsmr3638j 36 Dec 10 19:44 src -rw-r--r- 1 smr3638 grpsmr3638j 25K Dec 10 23:36 a3Sg\_vf1\_plo t.png

### Jupyter: some useful Features

 Tab to complete methods or attributes available.

**Documentation on functions** 



#### help(datetime) [16]:

Help on module datetime:

#### NAME

datetime - Fast implementation of the datetime type.

#### MODULE REFERENCE

https://docs.python.org/3.8/library/datetime

The following documentation is automatically generated from the Python

# Jupyter: tool for Collaboration

- Easily shareable: *Export Notebook As.*.
- Add it to (automatically rendered).
- Use *Solution* to run the notebook in the cloud.



• Use to package the notebook + dependencies.

### Notebooks are automatically saved and stored on disk as *.ipynb* (JSON).

	File	Edit	View	Run	Kernel	Tabs	Settin	igs	Help			
	Ν	lew					•	ichei	r			×
	Ν	lew Lau	uncher			ۍ ا	ደ	8		Ċ	►	
	С	Dpen fro	om Path									
	Ν	lew Vie	w for No	otebool	k				IC	ΤР	W	/or
	Ν	lew Co	nsole fo	r Noteb	ook					••	••	
	C	lose Ta	ab			٢	c w		Pro	oble	em	de
	С	close ar	nd Shuto	lown N	otebook	^ <del>(</del>	βQ					
	C	lose Al	l Tabs			We would				would	like	to e
	S	ave No	tebook			:	жs	Asciidoc				
	S	ave No	tebook	As		<del></del>	жs		HTML			
	S	ave All							LaTeX			
	R	eload N	Voteboo	k from	Disk				Markdo	own		
	R	evert N	lotebool	k to Ch	eckpoint				PDF			
Rename Notebook							cture	d Tex	t			
	D	ownloa	ad						Execut	able S	Script	
									Reveal	.js Slio	des	
	E	xport N	loteboo	k As			•	-	Webpo	(† ••• <b>•</b> ••	<b>0D</b>	01
	S	ave Cu	rrent Wo	orkspa	ce As			1	30	iuti	on	ex
	S	ave Cu	rrent Wo	orkspa	ce				1 5			ما د
	Р	rint				:	¥Р		1. R	eao	an	aa





A multi-user version of the notebook designed for companies, classrooms and research labs



## **JupyterHub capabilities**

using the web browser, without any local installation.

- Provides a Jupyter Notebook environment for thousands of users;
- Flexible authentication for users;
- Can be deployed on any infrastructure (public cloud, virtual machines, local server);
- Can be accessed from any computer with internet;
- Ready to use environment, with all dependencies installed (no more waisted) time with installing and maintaining software);
- Administrator can login into the running instance of each user and troubleshoot:
- Jupyter and JupyterHub are open source.



## JupyterHub allows users to access a server from any computer with internet

Sign in	
Username:	
Password:	
Sign in	

# JupyterHub setup for the ICTP Workshop



## ICTP Workshop use case

- JupyterHub installed on an ICTP multi-CPU server and available at: <u>https://jupyter.ictp.it/smr3638/</u> Each participant had an account (with credentials communicated by the ICTP organisers in
- advance, via e-mail)
- All participants had access to exactly the same environment with all the software tools needed already installed
- All participants had access to a common folder (with data examples, tutorials, executables)
- On the first day we had an introductory presentation on Jupyter Notebooks and command-line tools
- One lecturer was presenting the tutorial, and another one was running it (just like a participant), to lacksquareanticipate any problems





# Thank you!

Questions: L.Marian@iaea.org

Try it: <u>https://calc4-nds.iaea.org</u>