



# Interactive HPC with Jupyter (1)

training course, 26+27.05.2021

Jens Henrik Göbbert, [j.goebbert@fz-juelich.de](mailto:j.goebbert@fz-juelich.de)

Christian Witzler, [c.witzler@fz-juelich.de](mailto:c.witzler@fz-juelich.de)

Jülich Supercomputing Centre (JSC)  
Forschungszentrum Jülich (FZJ)



European  
Commission

Horizon 2020  
European Union funding  
for Research & Innovation

*The CoEC project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952181.*

*The CoE RAISE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951733.*

# WELCOME

- Hello !
- Workshop information (live document)
  - <https://gitlab.version.fz-juelich.de/hedgedoc/oo2I4aZHSKO5eIJJOLPk3w?view>
- Workshop interaction
  - Zoom chat
- Workshop repository
  - <https://gitlab.version.fz-juelich.de/jupyter4jsc/CoE-2021.05-jupyter4hpc>

# WELCOME

## Agenda

- day 1: JupyterLab Introduction

- 9:00 - 11:00
  - Welcome and Login
  - Introducing JupyterLab
- 11:00 - 11:30
  - Break
- 11:30 - 13:00
  - JupyterLab extensions tour
  - Customizing your environment

- day 2: Jupyterlab for HPC

- 9:00 - 11:00
  - Welcome and Login
  - Build your own kernels
  - Using JupyterLab as Proxy
- 11:00 - 11:30
  - Break
- 11:30 - 13:00
  - Utilizing supercomputers with JupyterLab
  - Jupyter-JSC under the hood

## Resources

HDF Cloud (OpenStack)

JUSUF (HPC cluster)

# PRE-WORKSHOP TODOS

## Register & Login

<https://judoor.fz-juelich.de>

**Join** the project „training2109“

→ Wait to get joined by the project PI

## Sign usage agreement

→ Wait for creation of HPC accounts

## Check Connected Services:

jupyter-jsc

The screenshot shows the Jülich Supercomputing Centre user interface. At the top, there's a navigation bar with 'Your account', 'Mentoring', a search bar, and 'Detailed Statistics'. The main content area is divided into several sections:

- Account:** Fields for Salutation, E-mail address (with a checkmark), Telephone, and Address.
- Mentored projects:** A section for managing projects.
- Systems:** A table showing system status:
  - judac:** Managed by SH-keys, usage agreement confirmed on 18.04.2021 (green checkmark).
  - jusuf:** JUSUF\_CPU: training2109 (red X), with a message: 'You need to sign the usage agreement to access this system' (red X).
- Projects:** A section for managing projects, showing 'Interactive High-Performance Computing with Jupyter' (green checkmark).
- Software:** A section for managing software.
- Connected Services:** A section showing connected services: trac, lview, jards, gitlab, and jupyter-jsc (green checkmark).

For more details, please visit

https://gitlab.version.fz-juelich.de/hedgedoc/oo2I4aZHSKO5eIJJOLPk3w?view#Pre-Workshop-Todos  
or https://gitlab.version.fz-juelich.de/jupyter4jsc/CoE-2021.05-jupyter4hpc/-/tree/master/preparation

# MOTIVATION

your thinking, your reasoning, your insides, your ideas

“It is all about using and building a machinery **interface** **between** computational researchers and data, supercomputers, laptops, cloud **and** your thinking, your reasoning, your insides, your ideas about a problem.”

Fernando Perez, Berkely Institute for Data Science  
Founder of Project Jupyter



# MOTIVATION

## Rise of Jupyter's popularity

If popularity can be counted by

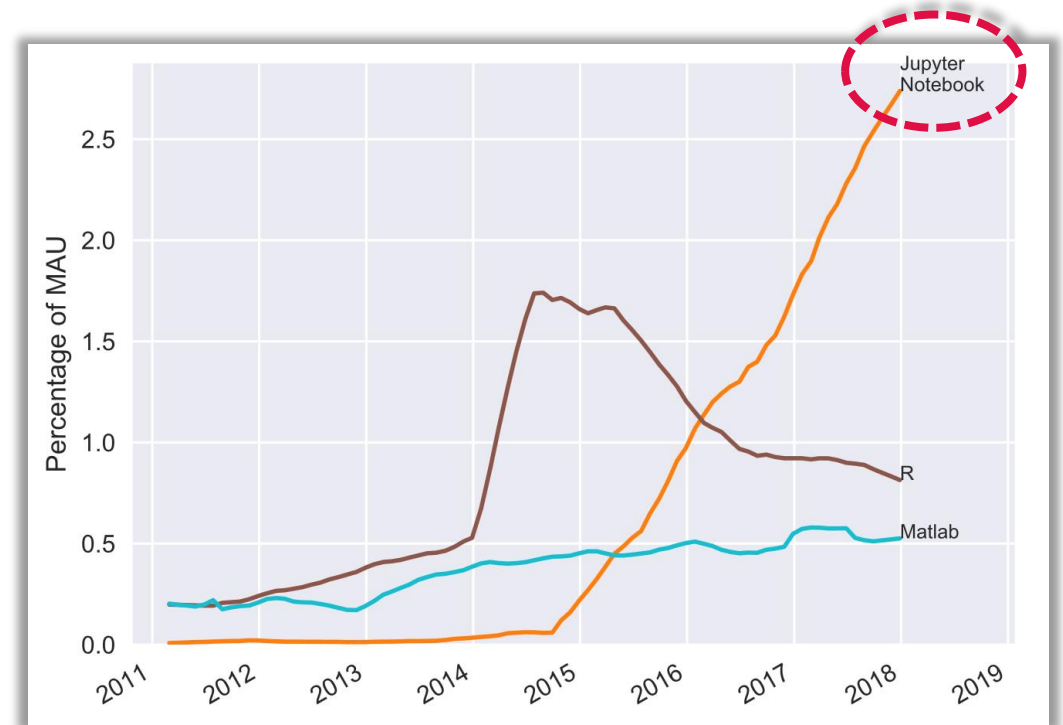
- Monthly aggregated number of user interactions with GitHub repos (= Monthly Active Users (MAU))

and

- Each repository is assigned to a single language (by looking at which language has the most bytes in the repo)

Jupyter Notebooks have seen significant and steady growth over the last years (still rising).

- Of course the popularity of Python in general is pushing this trend.



<https://www.benfrederickson.com/ranking-programming-languages-by-github-users/>  
<https://github.com/benfred/github-analysis>

# TERMINOLOGY

## What is JupyterLab

### JupyterLab

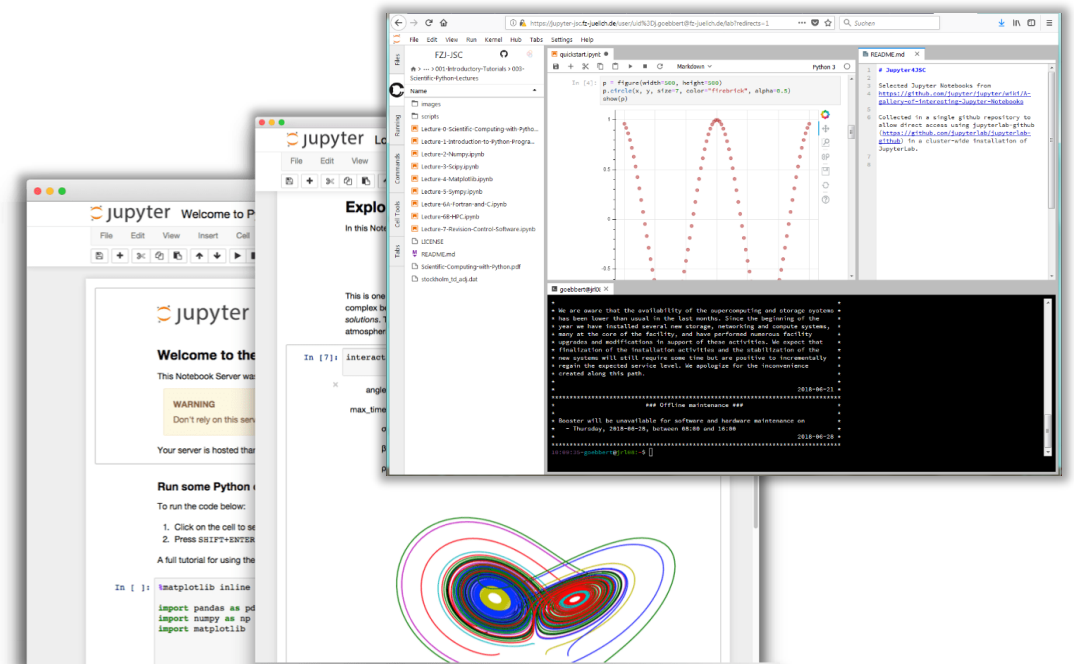
- **Interactive** working environment in the web browser
- For the creation of **reproducible** computer-aided narratives
- Very **popular** with researchers from all fields
- Jupyter = Julia + Python + R

### Multi-purpose working environment

- Language agnostic
- Supports execution environments (“*kernels*”)
  - For dozens of languages: Python, R, Julia, C++, ...
- Extensible software design („*extensions*”)
  - many server/client plug-ins available
  - Eg. in-browser-terminal and file-browsing

### Document-Centered Computing (“*notebooks*”)

- Combines code execution, rich text, math, plots and rich media.
- All-in-one document called Jupyter Notebook



<https://jupyterlab.readthedocs.io>

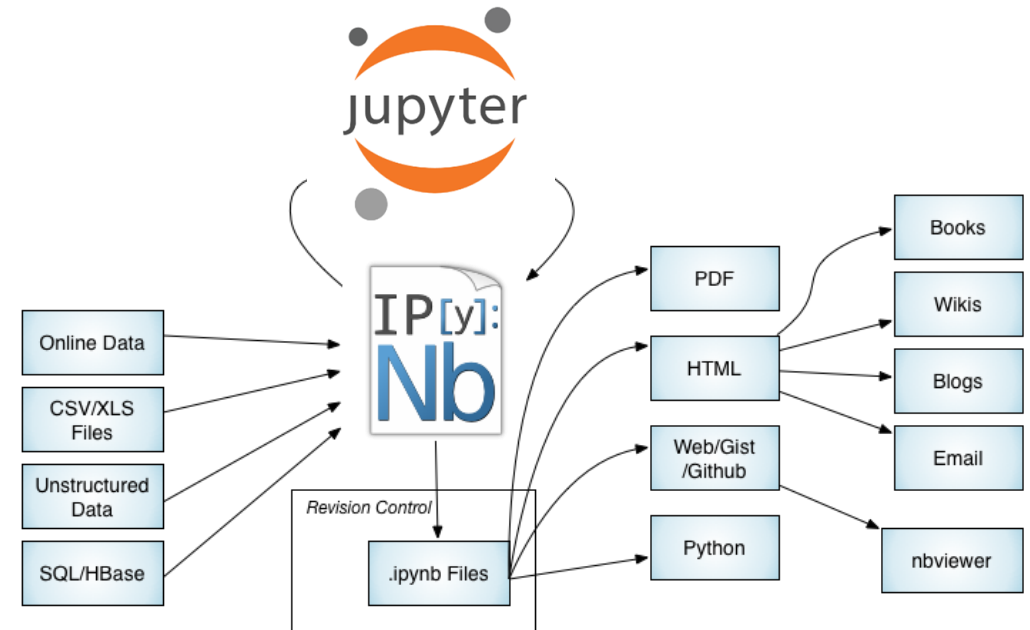
# TERMINOLOGY

## What is a Jupyter Notebook?

### Jupyter Notebook

A notebook document (file extension .ipynb) is a document that can be rendered in a web browser

- It is a file, which stores your work in JSON format
- Based on a set of open standards for interactive computing
- Allows development of custom applications with embedded interactive computing.
- Can be extended by third parties
- Directly convertible to PDF, HTML, LaTeX ...
- Supported by many applications such as GitHub, GitLab, etc..



<https://jupyter-notebook.readthedocs.io/>

<https://github.com/jupyter/jupyter/wiki/A-gallery-of-interesting-Jupyter-Notebooks>



# TERMINOLOGY

## What is a Jupyter Kernel?

### Jupyter Kernel

A “kernel” refers to the separate process which executes code cells within a Jupyter notebook.

### Jupyter Kernel

- **run code** in different programming languages and environments.
- can be **connected to** a notebook (one at a time).
- **communicates** via ZeroMQ with the JupyterLab.
- Multiple **preinstalled** Jupyter Kernels can be found on our clusters
  - Python, R, Julia, Bash, C++, Ruby, JavaScript
  - Specialized kernels for visualization, quantumcomputing
- You can easily **create your own kernel** which for example runs your specialized virtual Python environment.

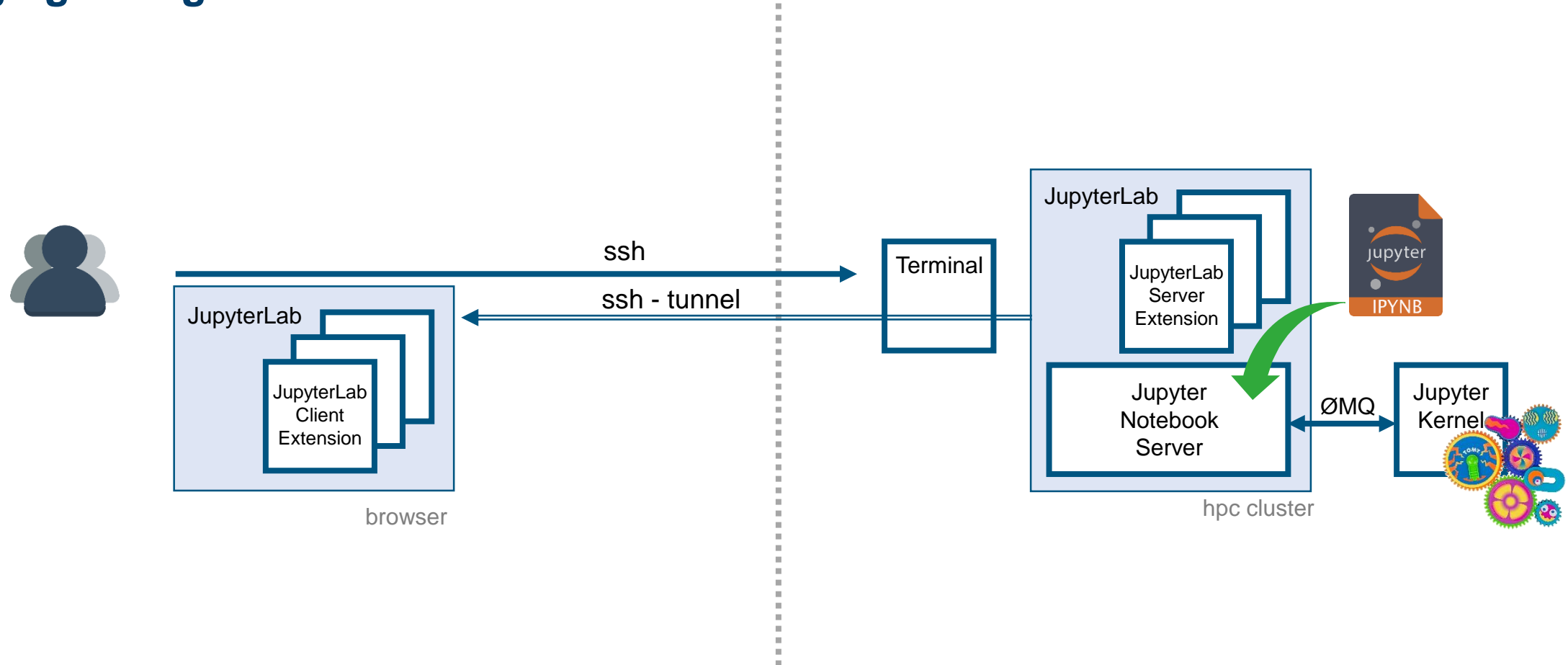


<https://jupyter-notebook.readthedocs.io/>  
<https://github.com/jupyter/jupyter/wiki/Jupyter-kernels>  
<https://zeromq.org>



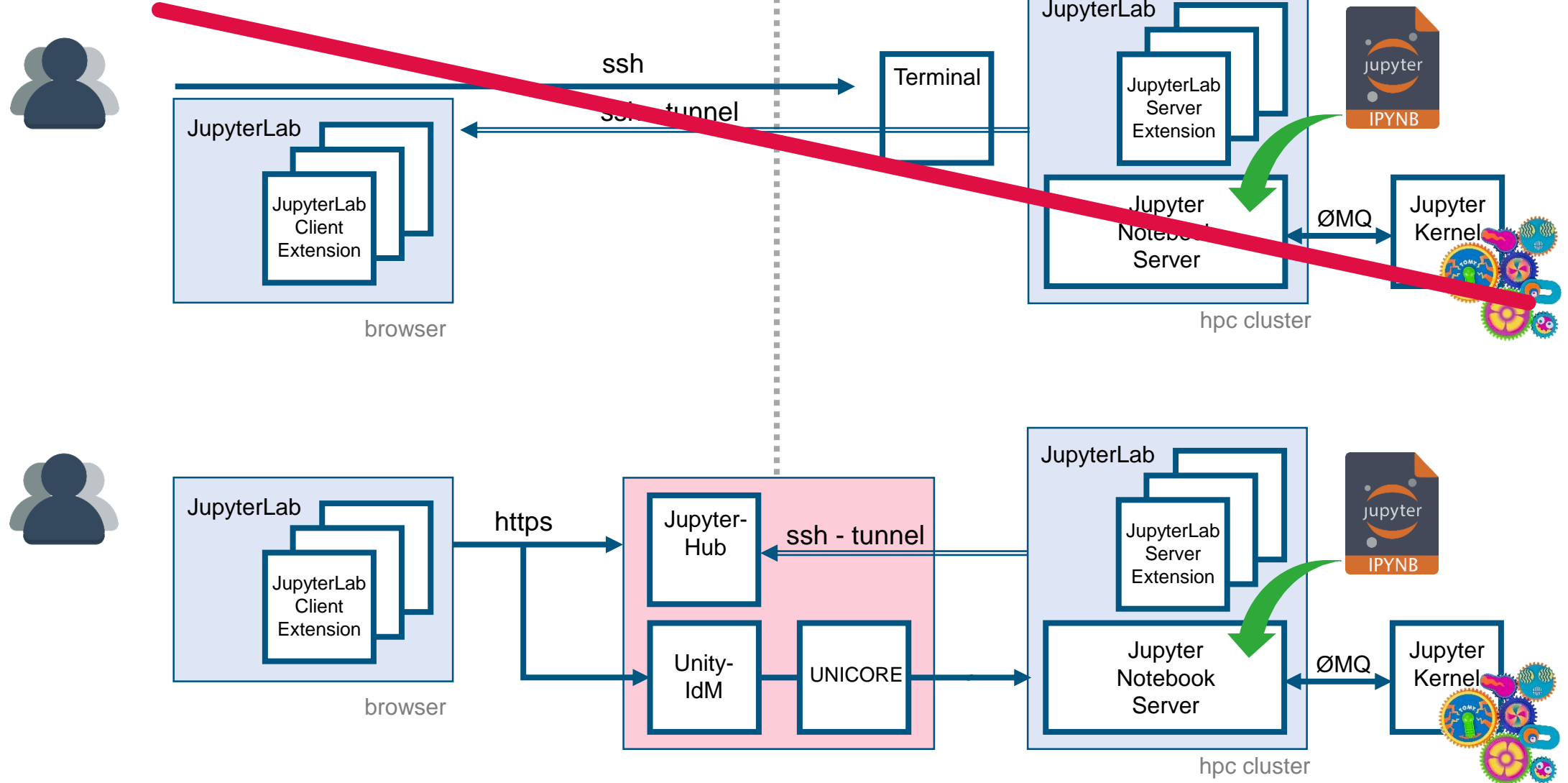
# TERMINOLOGY

Bringing all together



# JUPYTER-JSC WEBSERVICE

Start your JupyterLab (the easy way)



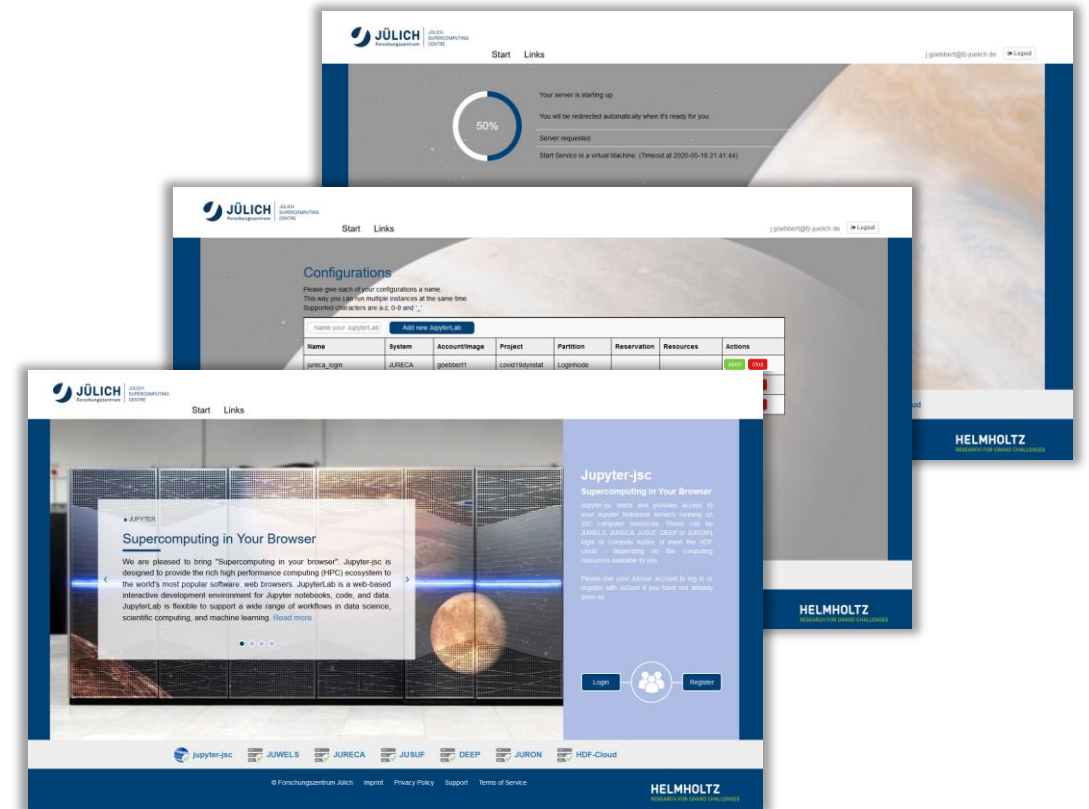
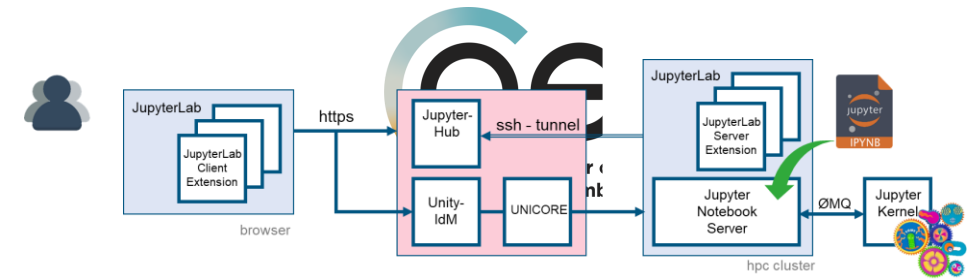
# JUPYTER-JSC WEBSERVICE

## Start your JupyterLab (the easy way)

### JupyterHub

is used to make Jupyter available to a group of HPC users.

- Creates/manages JupyterLabs for single users.
- Connects JupyterLabs to users via a configurable HTTP proxy.
- Supports custom spawners
  - UNICORE at JSC
- Supports custom authenticators
  - Unity-IdM at JSC



# JUPYTER-JSC WEBSERVICE

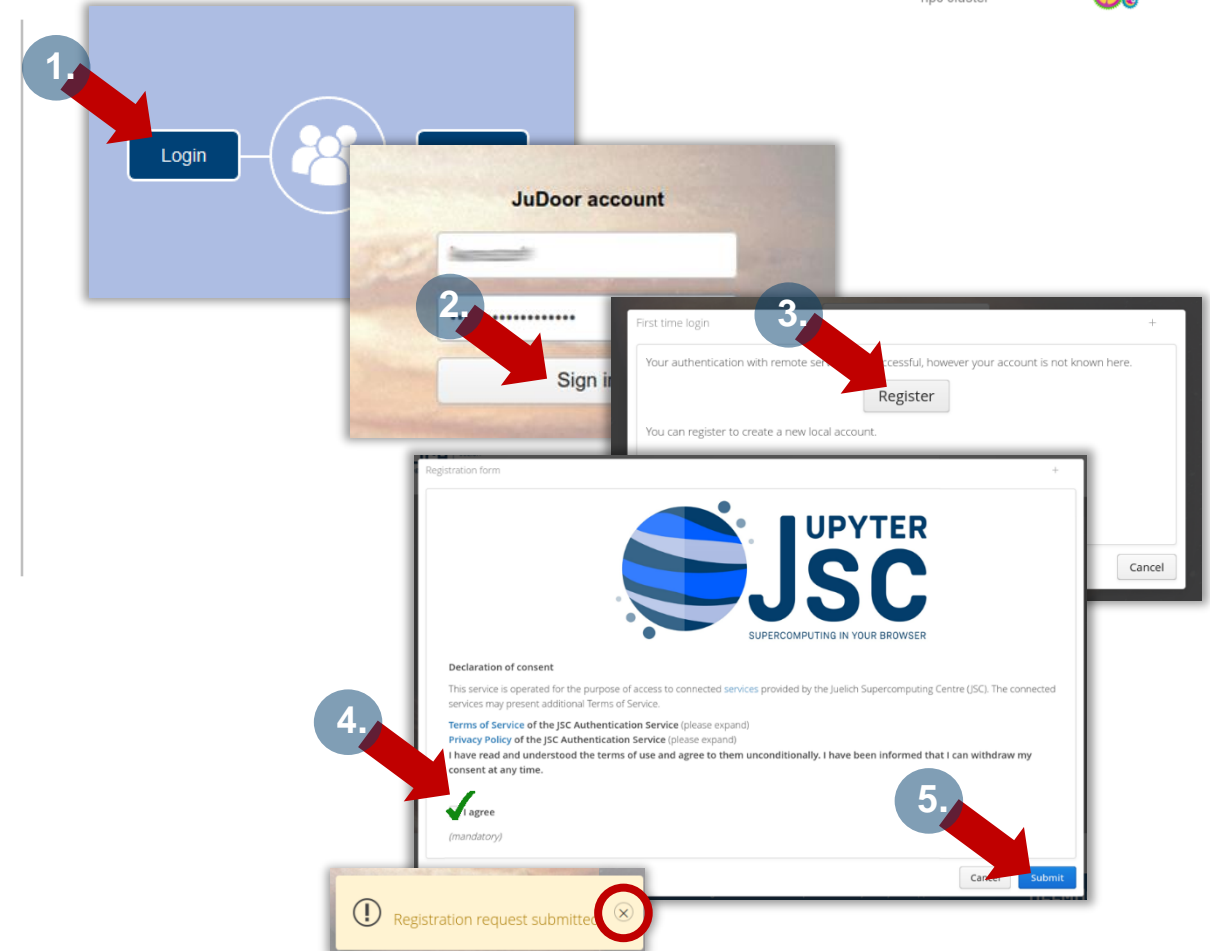
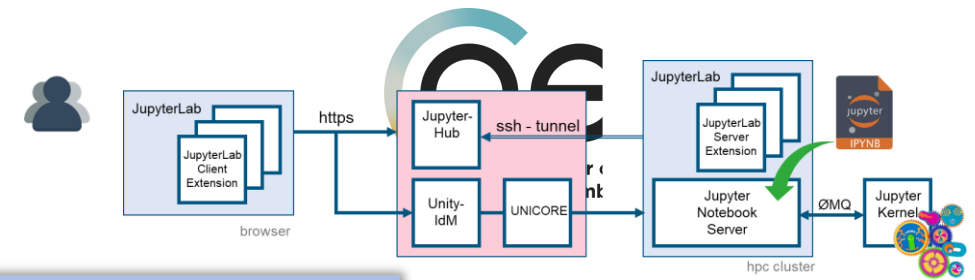
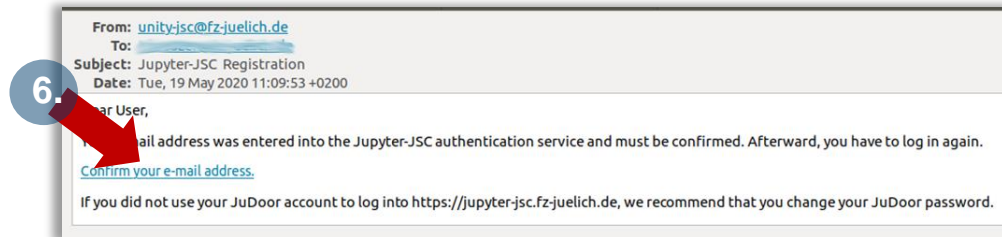
## First time login

=> <https://jupyter-jsc.fz-juelich.de>

### Jupyter-JSC first time login

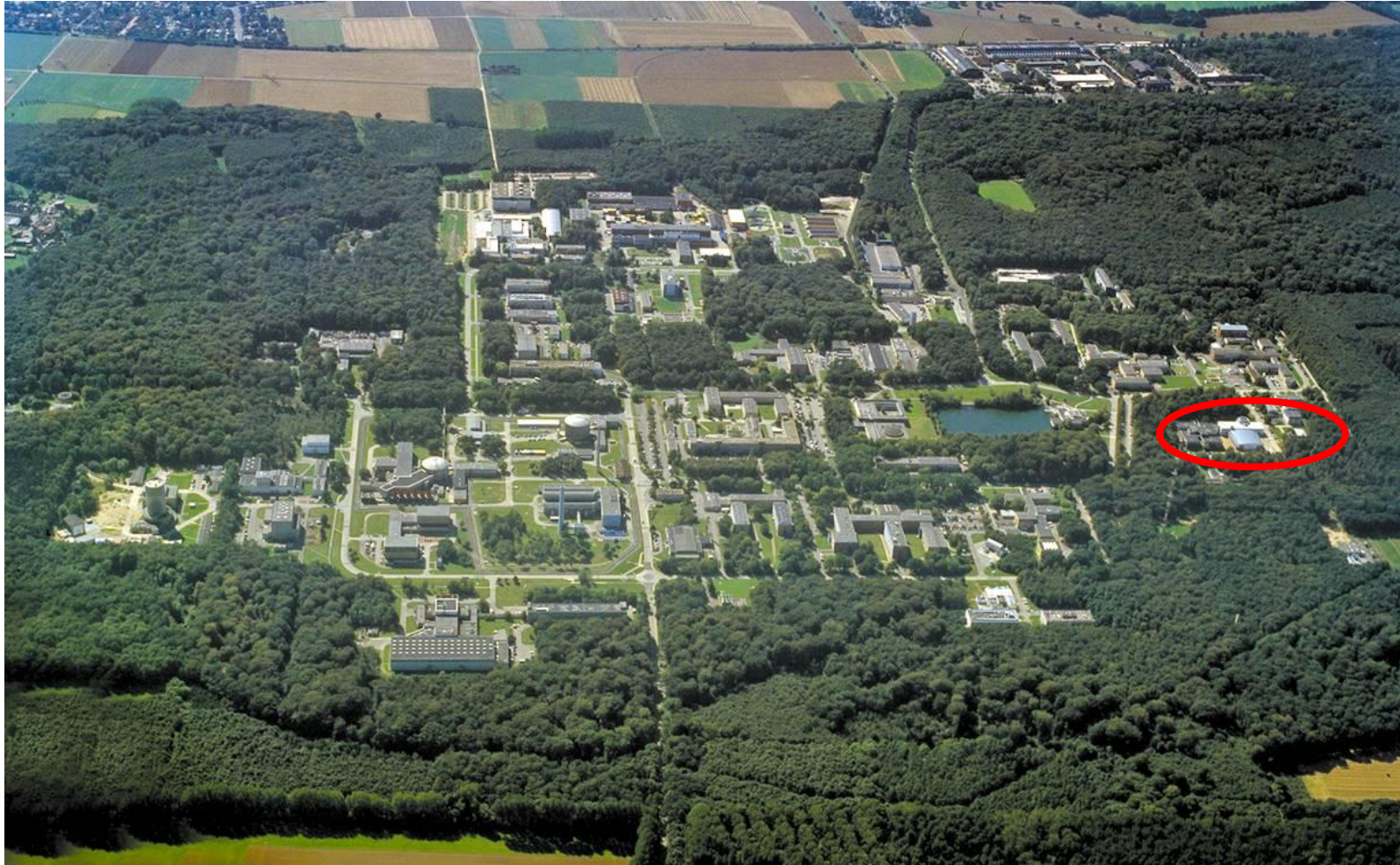
- Requirements:
  - Registered at [judoor.fz-juelich.de](https://judoor.fz-juelich.de)
    - (with “Connected Services” = jupyter-jsc)

1. Login at [jupyter-jsc.fz-juelich.de](https://jupyter-jsc.fz-juelich.de)
2. Sign in with your JSC account
3. Register to Jupyter-JSC
4. **Accept usage agreement**
5. Submit the registration
6. Wait for email and confirm your email address





# FORSCHUNGSZENTRUM JÜLICH



# JÜLICH SUPERCOMPUTING CENTRE



## Supercomputer operation for:

- Centre – FZJ
- Region – RWTH Aachen University
- Germany – Gauss Centre for Supercomputing  
John von Neumann Institute for Computing
- Europe – PRACE, EU projects

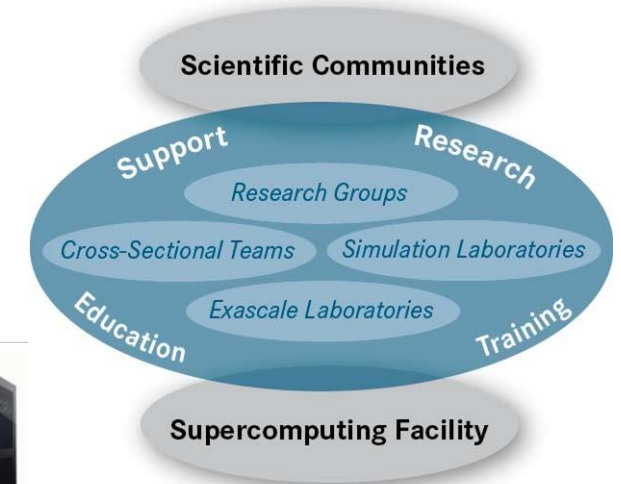
## Application support

- Unique support & research environment at JSC
- Peer review support and coordination

## R&D work

- Methods and algorithms, computational science, performance analysis and tools
- Computer architectures, Co-Design, Exascale Laboratories

## Education and Training







IBM Power 4+  
JUMP, 9 TFlop/s



IBM Blue Gene/L  
JUBL, 45 TFlop/s



IBM Blue Gene/P  
JUGENE, 1 PFlop/s



IBM Blue Gene/Q  
JUQUEEN  
5.9 PFlop/s



File Server



JUropa  
200 TFlop/s  
HPC-FF  
100 TFlop/s



JURECA (2015)  
2.2 PFlop/s



JUST Gen 5:  
100+ PB raw



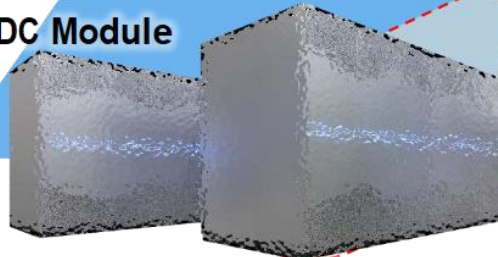
JURECA Booster  
(2017)  
5 PFlop/s



JUWELS Cluster  
(2018)  
12 PFlop/s



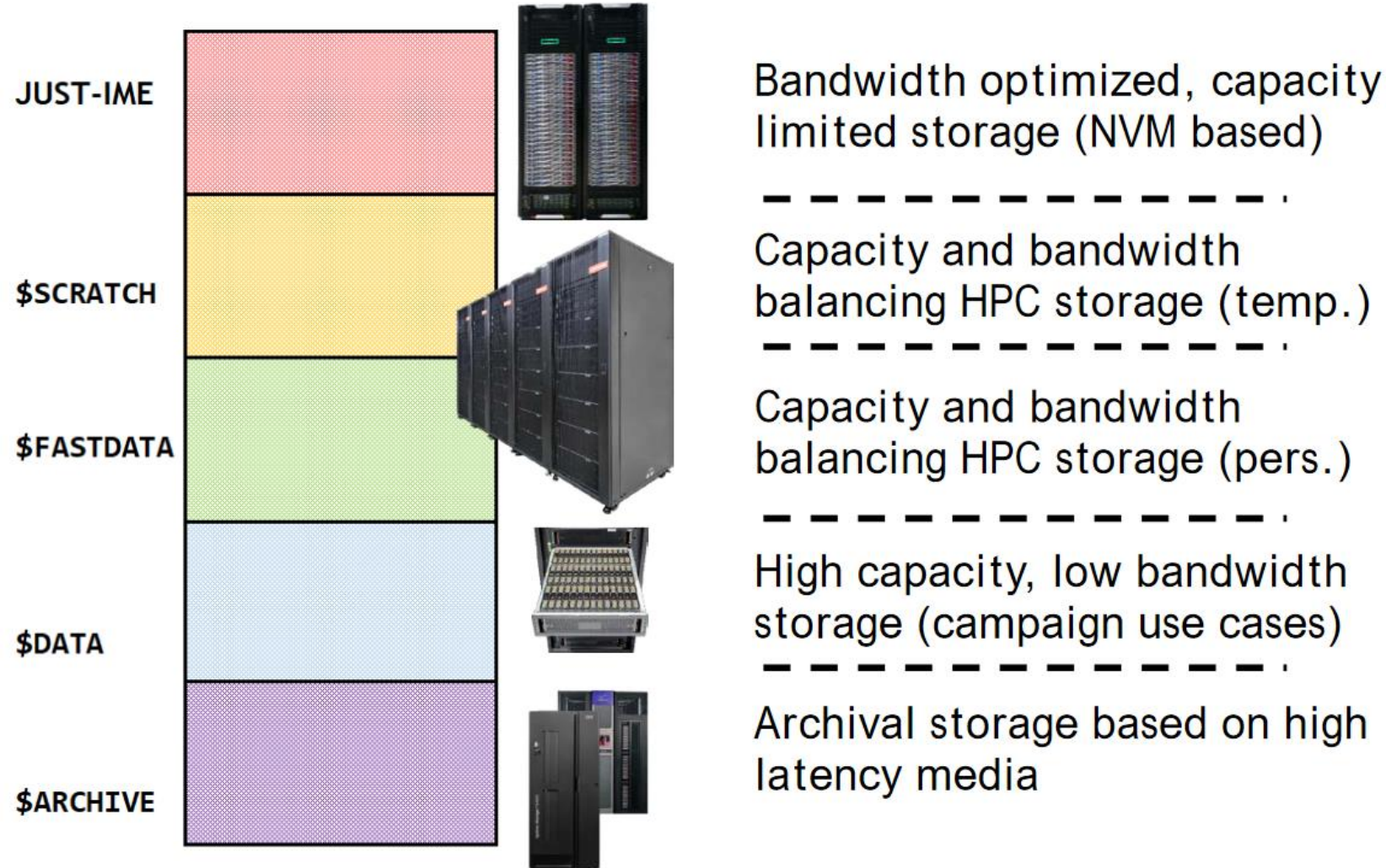
JUWELS Booster  
(2020)  
73 PFlop/s



JURECA DC Module  
(2020)



# JUST: MULTI-TIER STORAGE SYSTEM





# DOMAIN SPECIFIC SUPPORT

