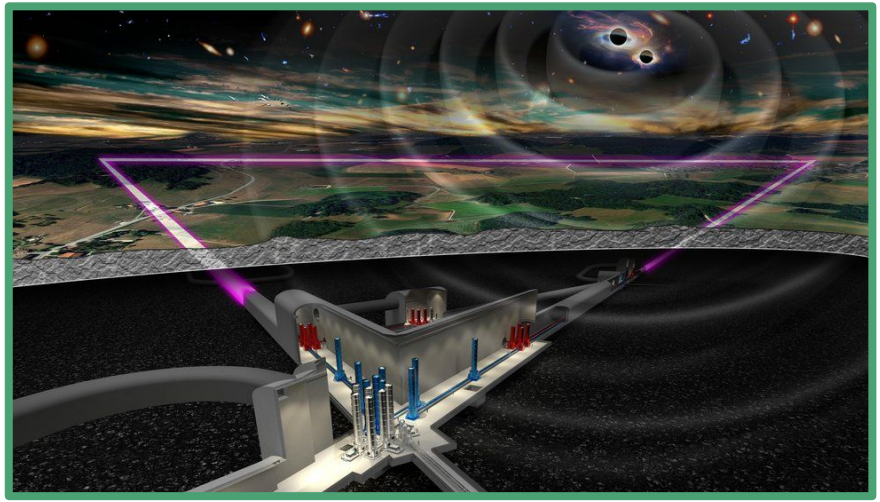
Building a Virtual Research Environment for the Einstein Telescope Project

Tommaso Diotalevi, Marco Lorusso, Stefano Dal Pra, Daniele Bonacorsi

ISGC2026 - Taipei (Taiwan), 20/03/2026

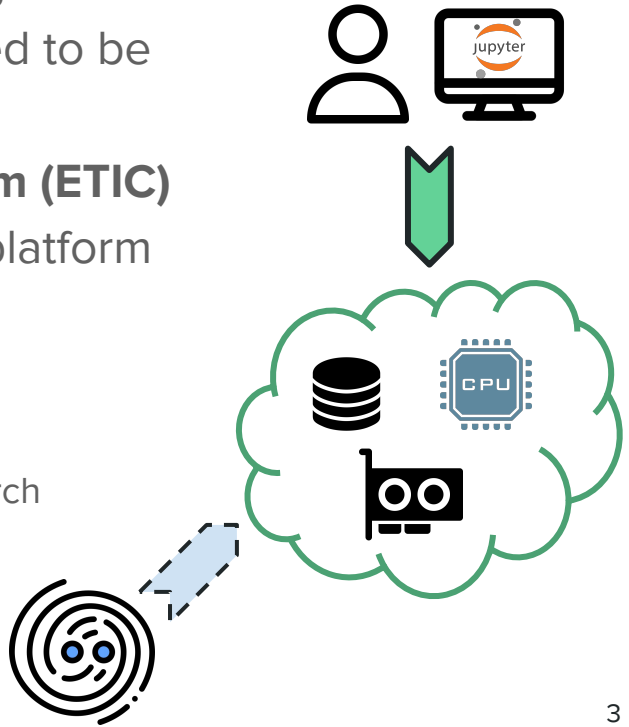
Introduction

- **Einstein Telescope** (ET) is a proposed next-generation gravitational-wave observatory in Europe.
  - It is designed to be about 10 times more sensitive than current detectors, such as LIGO (US) and Virgo (Italy).
- ET aims to observe **gravitational waves** from the earliest epochs of the universe and study **black holes**, neutron stars, and fundamental physics.
- 3 locations are under consideration:
  - Sardinia, Italy
  - Euregio Meuse-Rhine, Netherlands.
  - Saxony, Germany.



Introduction

- Once operational, the **Einstein Telescope** (ET) will generate a large amount of experimental data, which will need to be efficiently managed and analyzed.
- As part of the national **ET Infrastructure Consortium (ETIC)** project, the main goal is to build a heterogeneous platform capable of connecting the analysts with a cloud infrastructure providing modern facilities:
  - Based on open-source standards (Jupyter, Kubernetes, ...);
  - High-level libraries and frameworks, for a easy-to-use research environment;
  - Customizable by the user via containers.



Einstein Telescope in Bologna

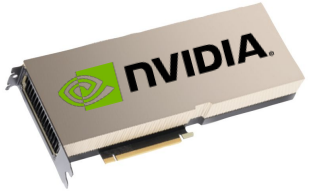
- The Bologna ET operational unit (BoET) is a joint effort between INFN and University of Bologna.
- **Scientific-technical goals**:
  - Development of a **computing cluster** (BETIF/DIFAET) for R&D through the setup of a heterogeneous infrastructure;
  - Development of electronic boards (in synergy with VIRGO) for super-attenuators;
  - DAQ test prototypes (developed for the VIRGO upgrade);
  - Time synchronization system based on White Rabbit;
  - Much more (theoretical studies, data analysis, …)

This talk!

Hardware specifications

- **CPUs**: 2 × Intel® Xeon® Platinum 8462Y+ (64C / 128T);

- **GPUs**: 2x NVidia® H100;



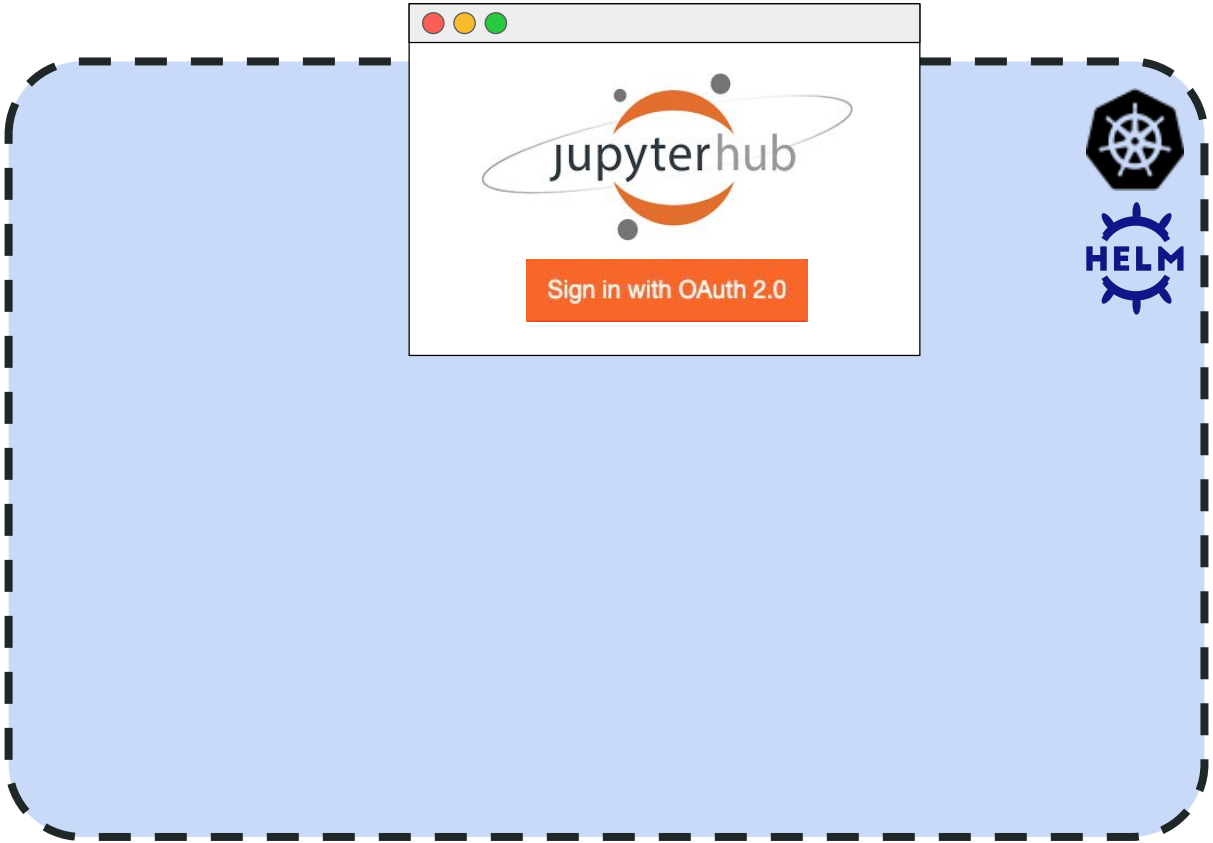
- **FPGAs**: 2x AMD® Alveo U55C

- Memory: 2TB DDR5;
- Networking: 2x10 Gbps Ethernet + 2x40 Gbps Fiber Optic;
- Storage: 4x15TB PCIe NVMe + 1TB SSD. Configured as RAID 10 for redundancy.

Our prototypal Virtual Research Environment

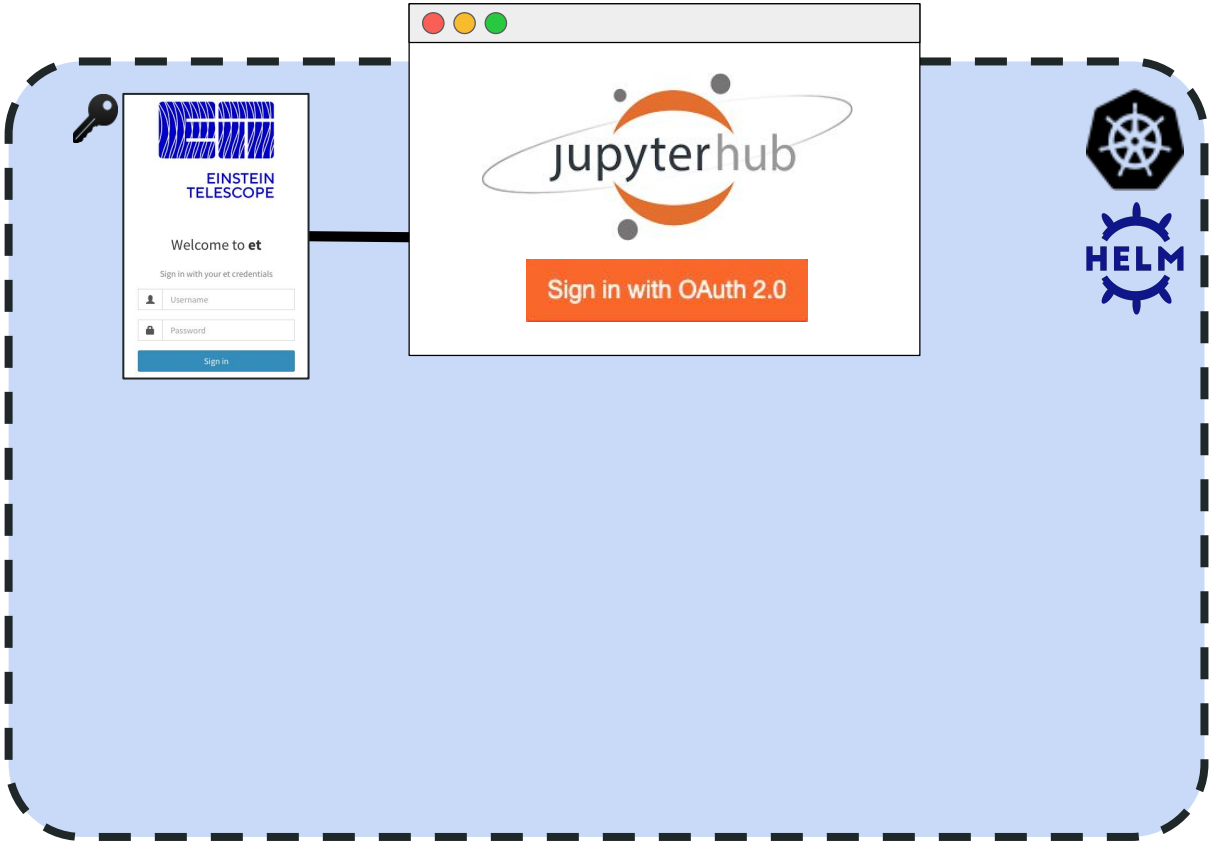
- Built directly from bare metal;
- Based on open-source **Kernel-based Virtual Machine** (KVM) virtualization module;
- Managed using **libvirt** API, creating a series of independent virtual nodes;
- GPU passthrough via **vfio-pci** dedicated driver
- Cluster managed with **Rancher Kubernetes Engine** (RKE2);
- Platform deployed using **HELM** package manager for K8s;
- User interface based on the **Jupyter** framework;
- Highly customizable user environment via **Docker** containers;
- Customizable python kernels, containerised and available using CERN's **CVMFS**;
- Data IO managed by **Rucio** directly in the platform interface.

Architecture



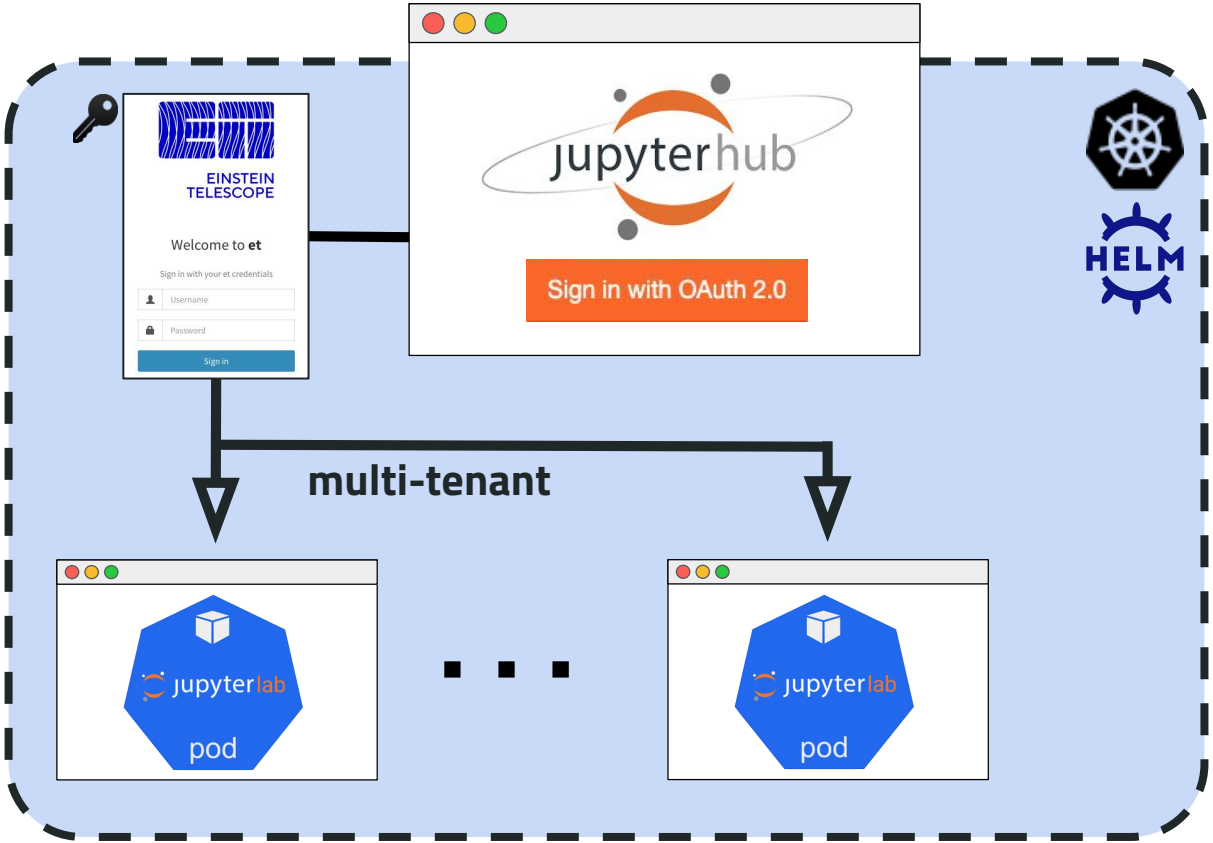
- Access to single hub entrypoint;

Architecture



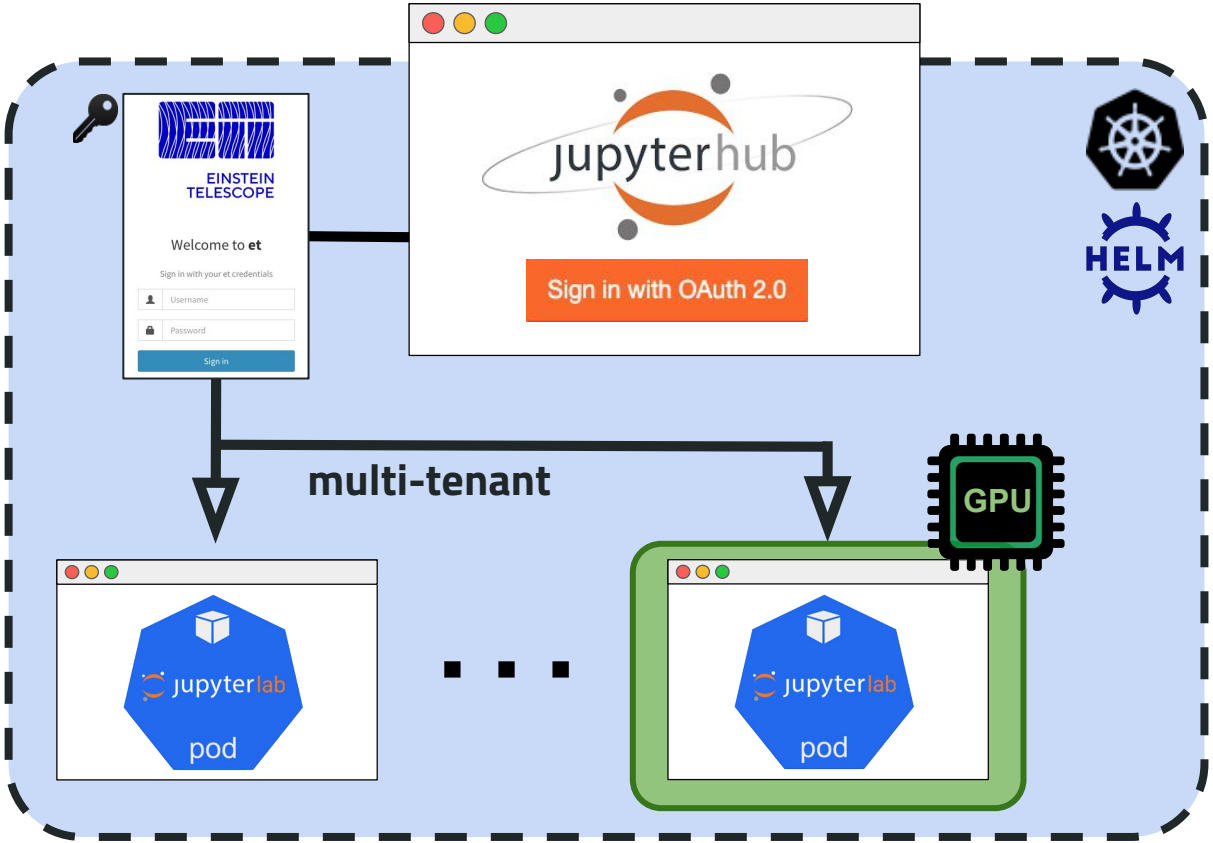
- Access to single hub entrypoint;
- Authn and authz token-based via INDIGO-IAM;

Architecture



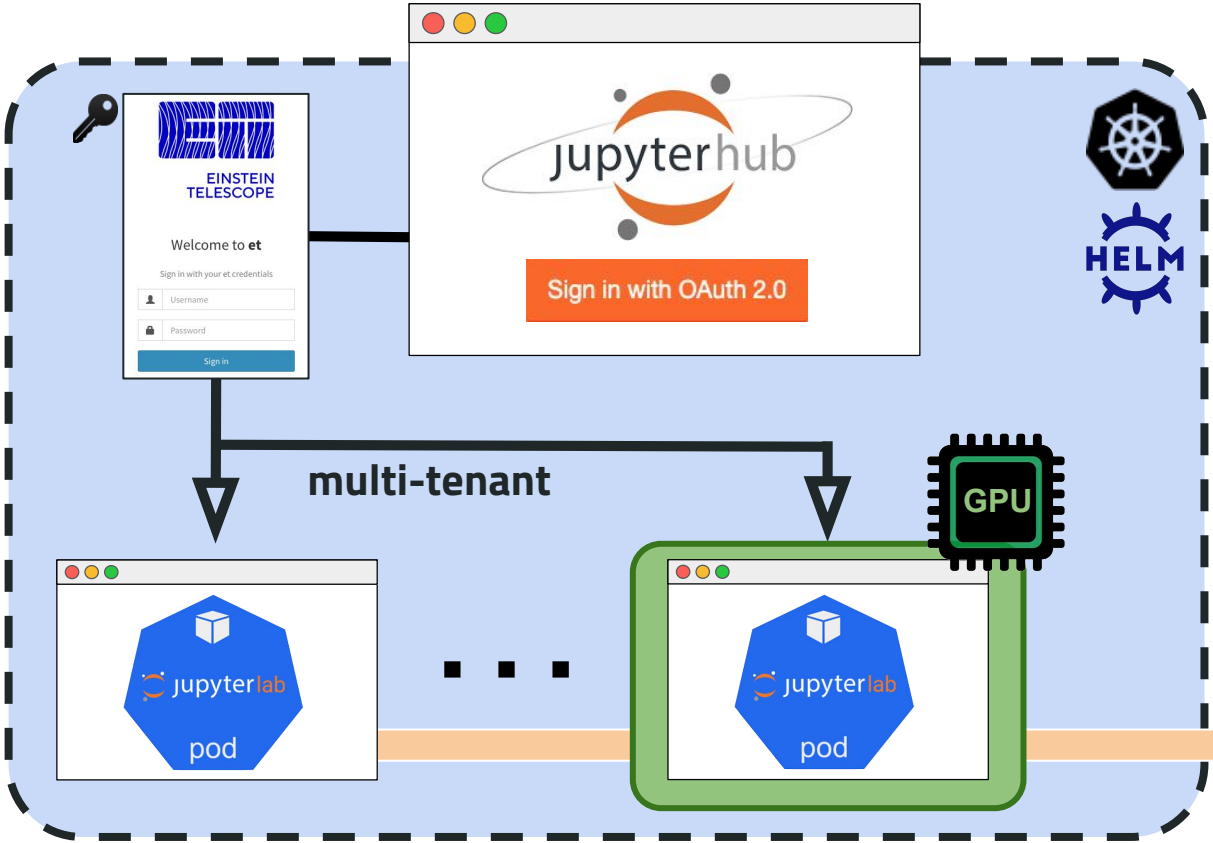
- Access to single hub entrypoint;
- Authn and authz token-based via INDIGO-IAM;
- Multi-user access to private JupyterLab sessions;

Architecture



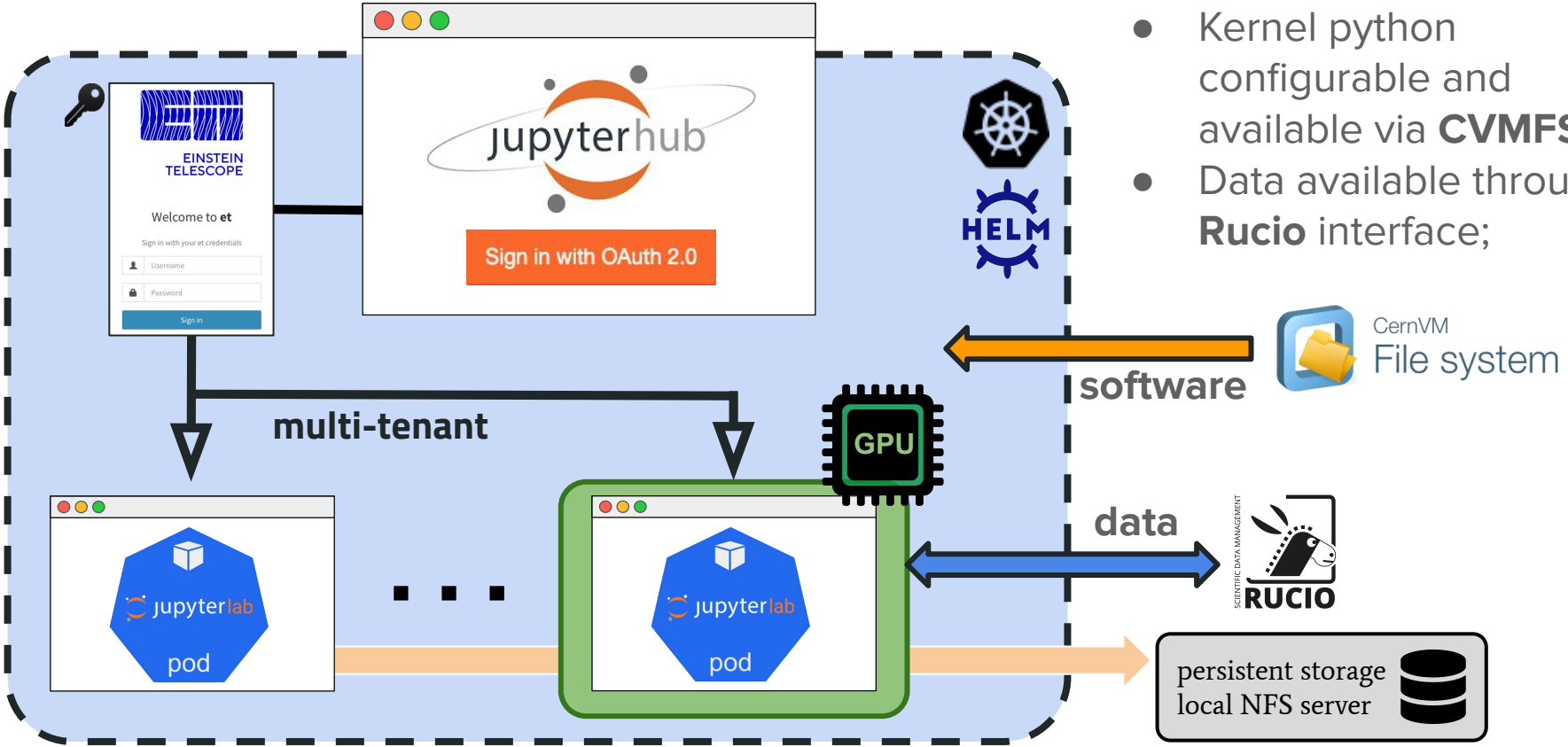
- Access to single hub entrypoint;
- Authn and authz token-based via INDIGO-IAM;
- Multi-user access to private JupyterLab sessions;
- GPU session available on demand;

Architecture



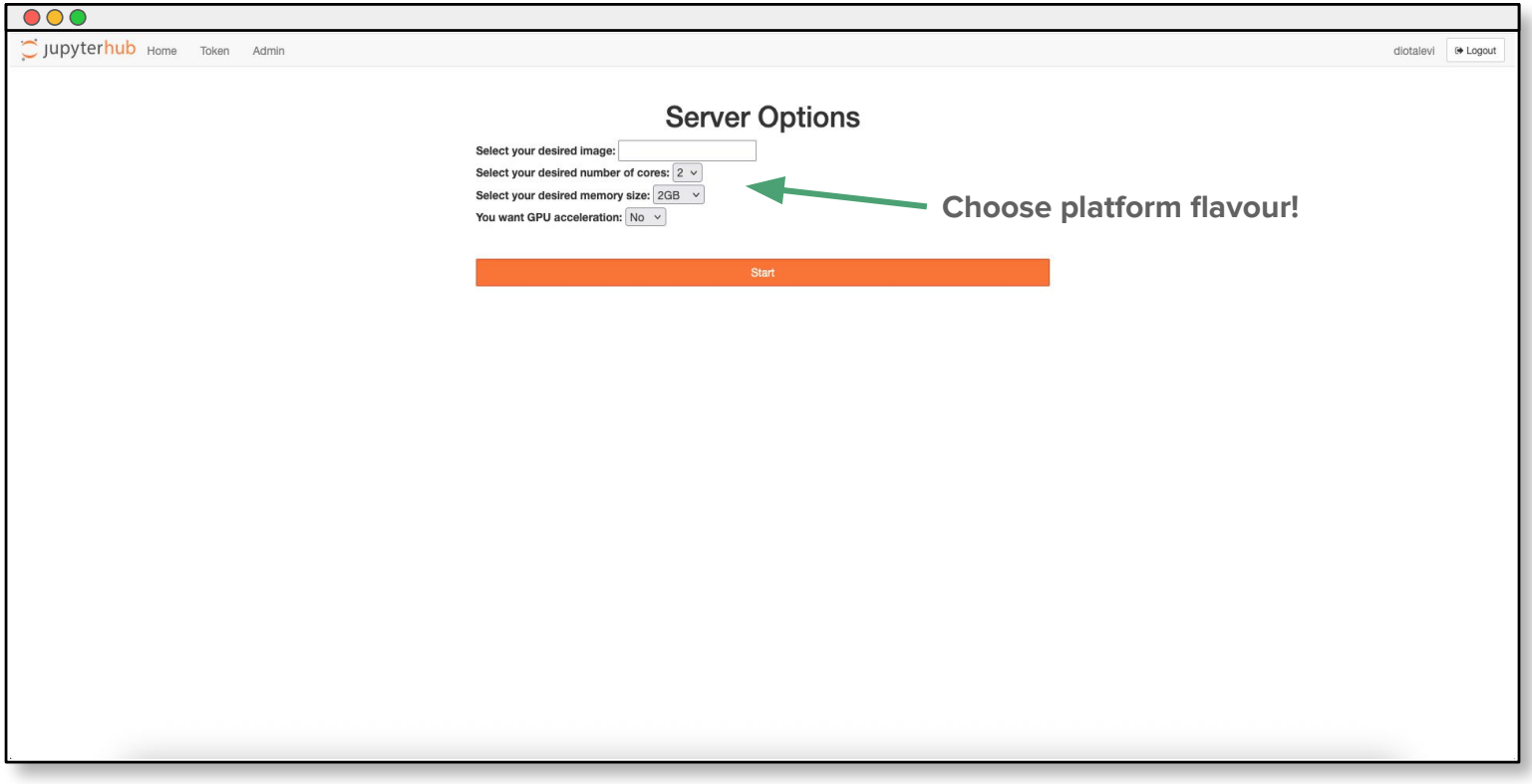
- Access to single hub entrypoint;
- Authn and authz token-based via INDIGO-IAM;
- Multi-user access to private JupyterLab sessions;
- GPU session available on demand;
- External storage for persistent user space;

Software and data access



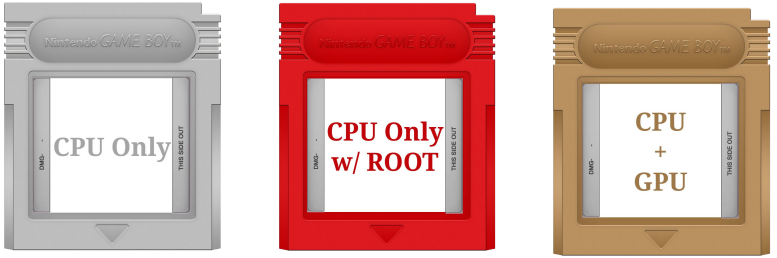
- Kernel python configurable and available via **CVMFS**;
- Data available through **Rucio** interface;

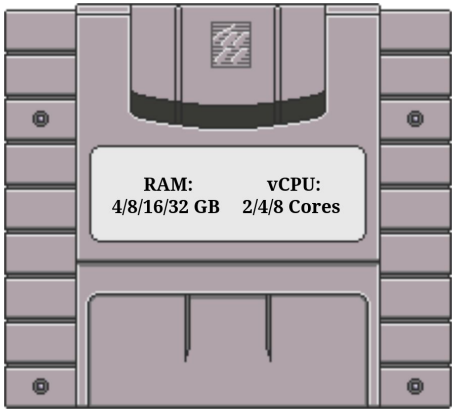
Welcome page (after authentication)



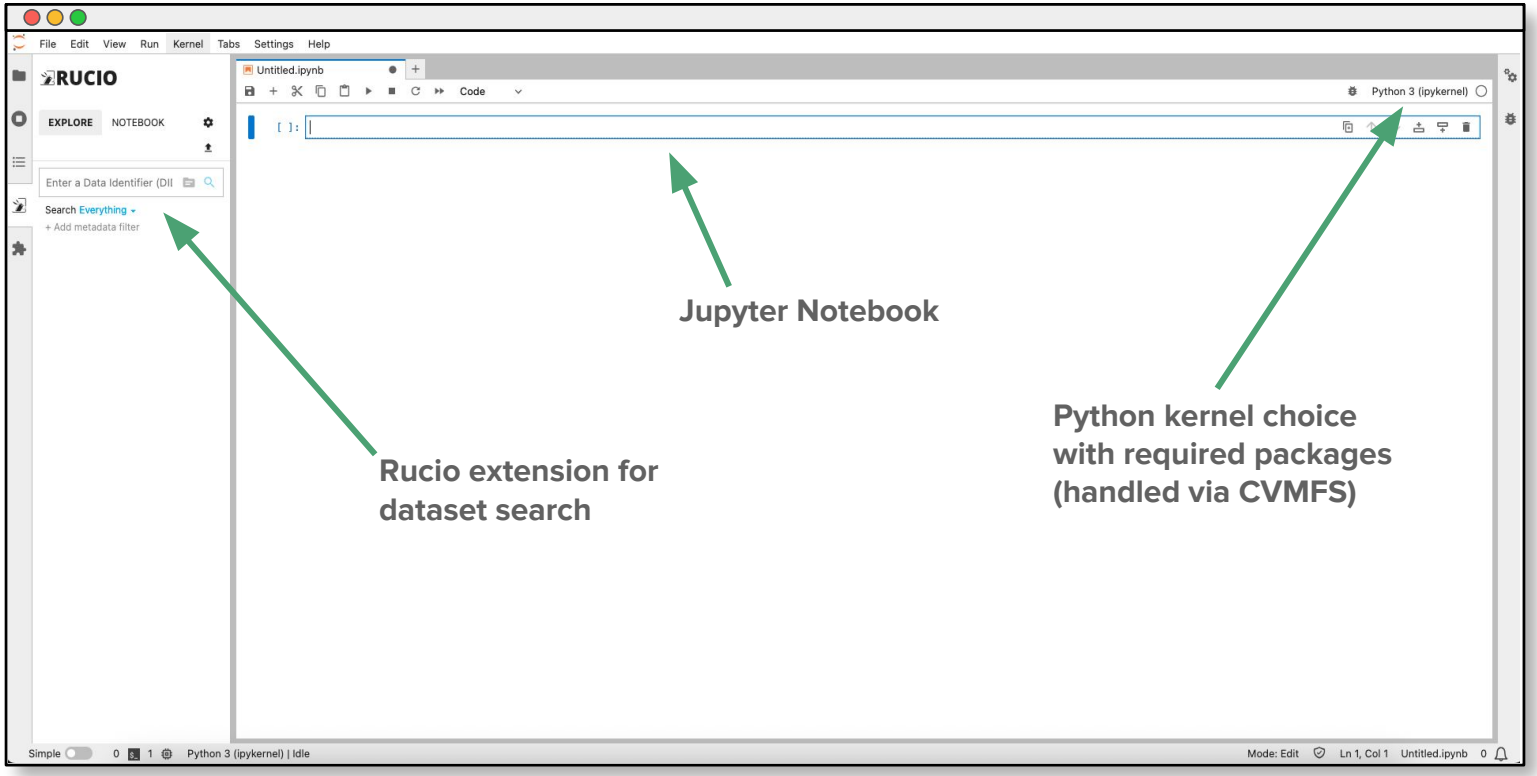
Platform flavours

- Fully customizable base images:
  - CPU only
  - CPU+GPU with CUDA 12.4 toolkit installed;
- Number of vCPU cores:
  - 2,4,8
- Memory size:
  - 4,8,16,32 GB
- GPU acceleration: yes or no (in that case the CPU+GPU image is necessary).



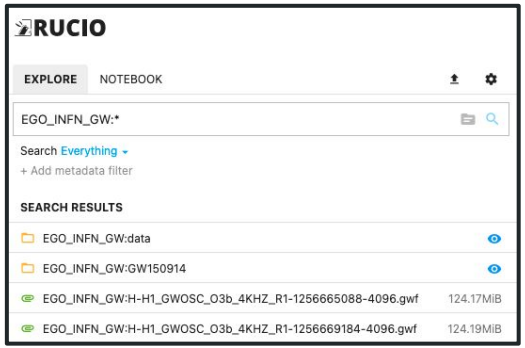


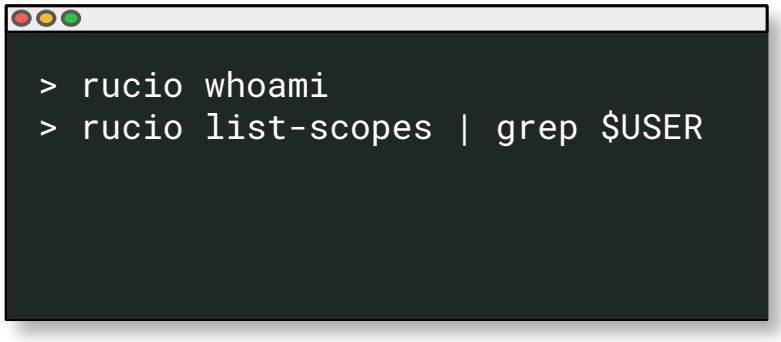
JupyterLab user session



Rucio data access

- Rucio client available on-platform, with the JupyterLab extension;
- Checks available copies of a chosen dataset:
  - provides access to its content directly, through a reference on a Jupyter notebook cell
- Authentication to Rucio Storage Elements (RSE) is possible via x509 proxies and JWT tokens;
- Available also via the CLI, for additional access and management.





Conclusions

- The Bologna ET group is actively committed to the development of a **Virtual Research Environment**, using cutting-edge hardware and open-source software/tools/technologies:
    - Heterogeneous computing resources completely transparent for the user.
      Currently working on **CPU/GPU** (integrating the FPGA resources as well);
    - Full customization for software & libraries for analysis, thanks to containerised configurations loaded through CVMFS;
    - Enforced **AAA** (Any Data, Anytime, Anywhere) **data access** with Rucio and network protocols (like XRootD and WebDAV), for stage-in/out operations;
    - The platform is available for authenticated (and authorized) ET users, via **IAM** groups.
- R&D will also be done on the available hardware: testing GPUs and FPGAs capabilities, with the future objective of expanding the cluster!

Thank you for the attention!

Check-out the infrastructure documentation here!