



# ESCAPE

European Science Cluster of Astronomy &  
Particle physics ESFRI research Infrastructures

## VRE introduction and updates

Enrique GARCIA on behalf of the VRE tech team

WP3 FG1 meeting – 17 June 2022



# VRE: Virtual Research Environment

- Project started July 2021 lead by Arturo Sanchez.
- After summer, new components integrated the team
  - Elena Gazzarrini (CERN, WP2 + EOSC Future)
  - Alba Vendrell (CERN, WP2)
  - Giovanni Guerrieri (Trieste+CERN)
  - Enrique Garcia
- Original idea: PoC platform/framework to connect the ESCAPE WPs outputs



# VRE current vision

- The ESCAPE VRE aims to be an **environment** that **integrates all** of the **ESCAPE tools** from WP2,3,4 and 5 in order to **provide a framework within which to deploy the Science Projects (TSP)** and their sub-projects.
  - It is not a goal to develop new components, but rather to glue them together into a coherent research environment.
  - It is focussed on the TSPs as cross-RI experiments and should be a collaborative online platform where EOSC Future members will be able develop those projects.
  - It **should expose the digital content of the scientific results** such that it can be easily findable, accessible and reusable.
  - It is a first proof of concept that will put into production and onboarding ESCAPE services in the context of EOSC.
- Thus we expect the VRE to be an access point where interested researchers (both inside and external to ESCAPE) can explore what is available and what is being done in ESCAPE related to the cross-RI science projects. **Ultimately this would be connected to the EOSC portal as an entry point to ESCAPE.**



# VRE components

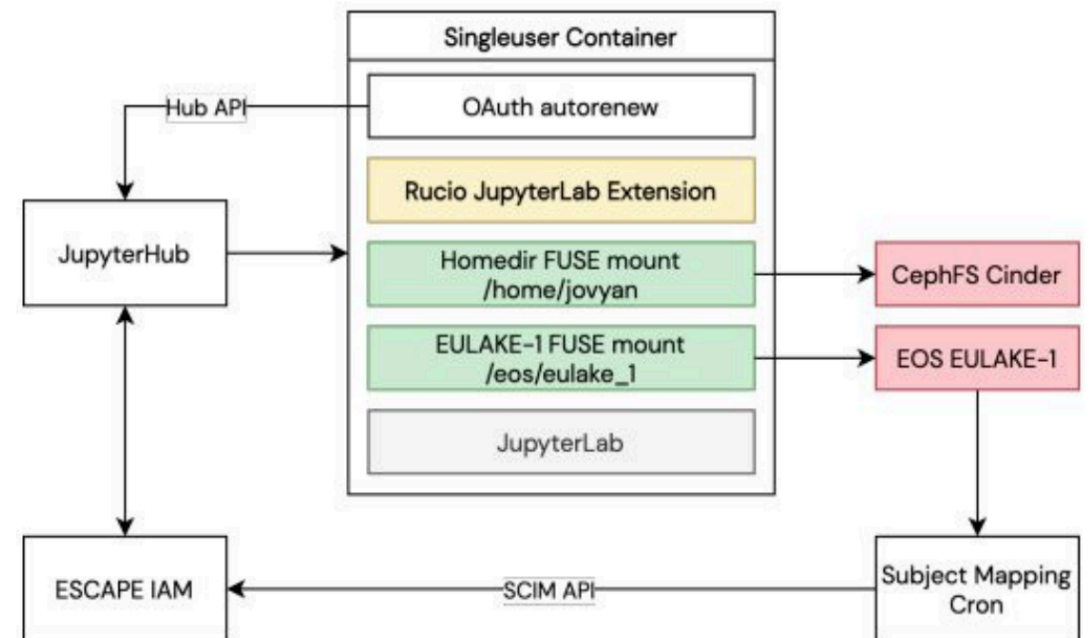
- VRE main building block: **DLaaS** (Data Lake as a Service).
  - Developed by WP2 to:
    - Provide easier interaction with the RUCIO instance (distributed storage)
    - Connect DL with an storage element (CERN RSE) + computing resource (Open stack VM)
  - A (Jupyter) env that connects the ESCAPE Data Lake with a computing facility

The goal of the service is to abstract the complexities of the Data Lake from the scientists. This way, scientists can focus their time on doing science instead of data procurement.



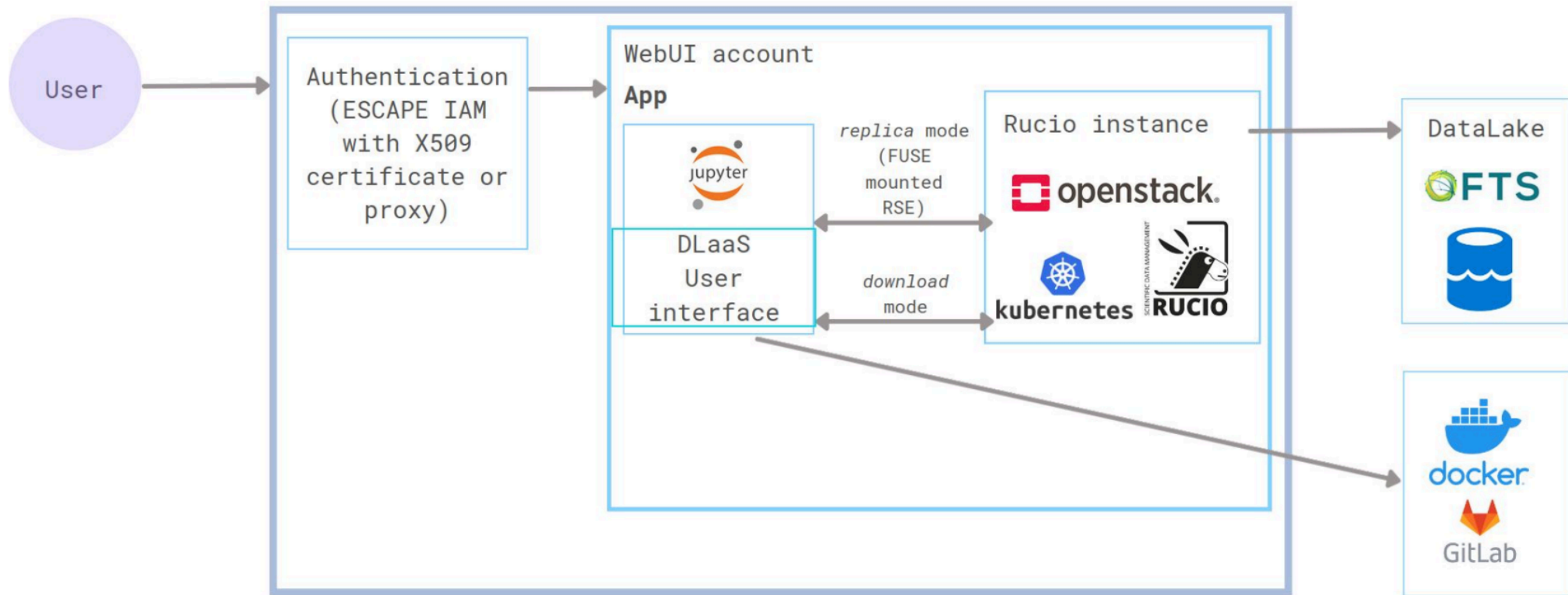
# DLaaS infrastructure

- K8s cluster deployed in a Openstack VM @ CERN
  - Connects RUCIO instance + RSE + monitoring + JupyterHub server + ...
- OAuth authentication using ESCAPE IAM.
- Uses Rucio Jupyterlab extension <https://github.com/rucio/jupyterlab-extension>



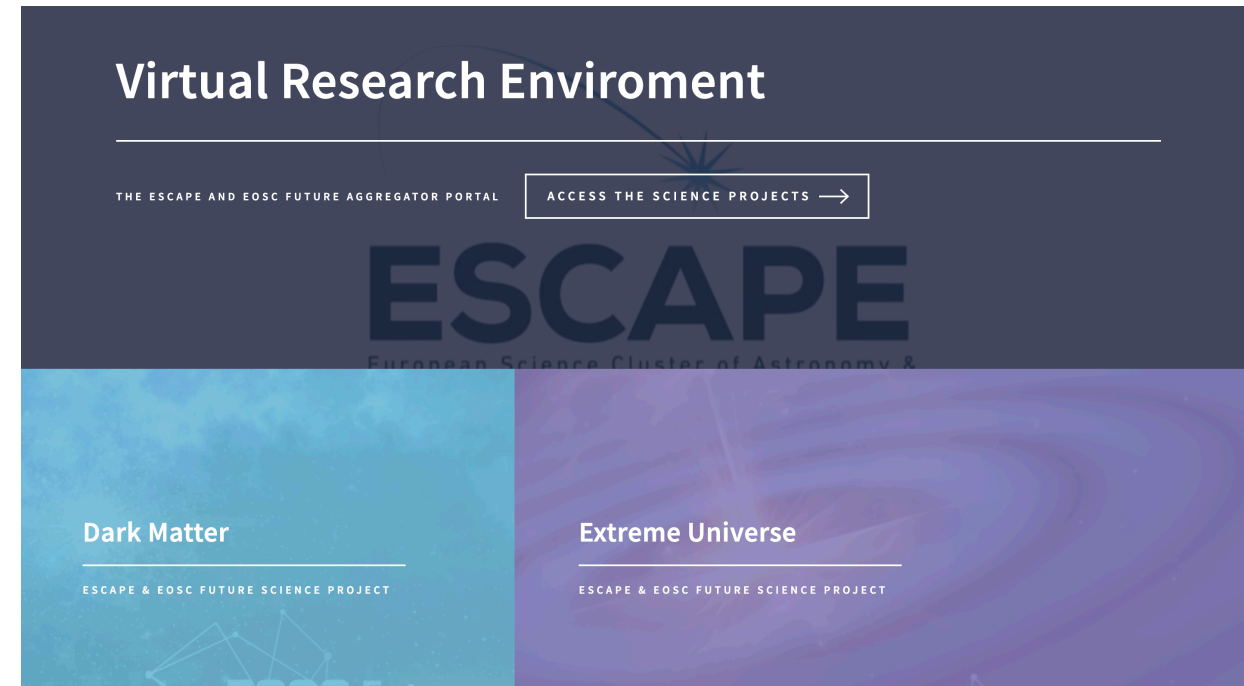


# VRE current status

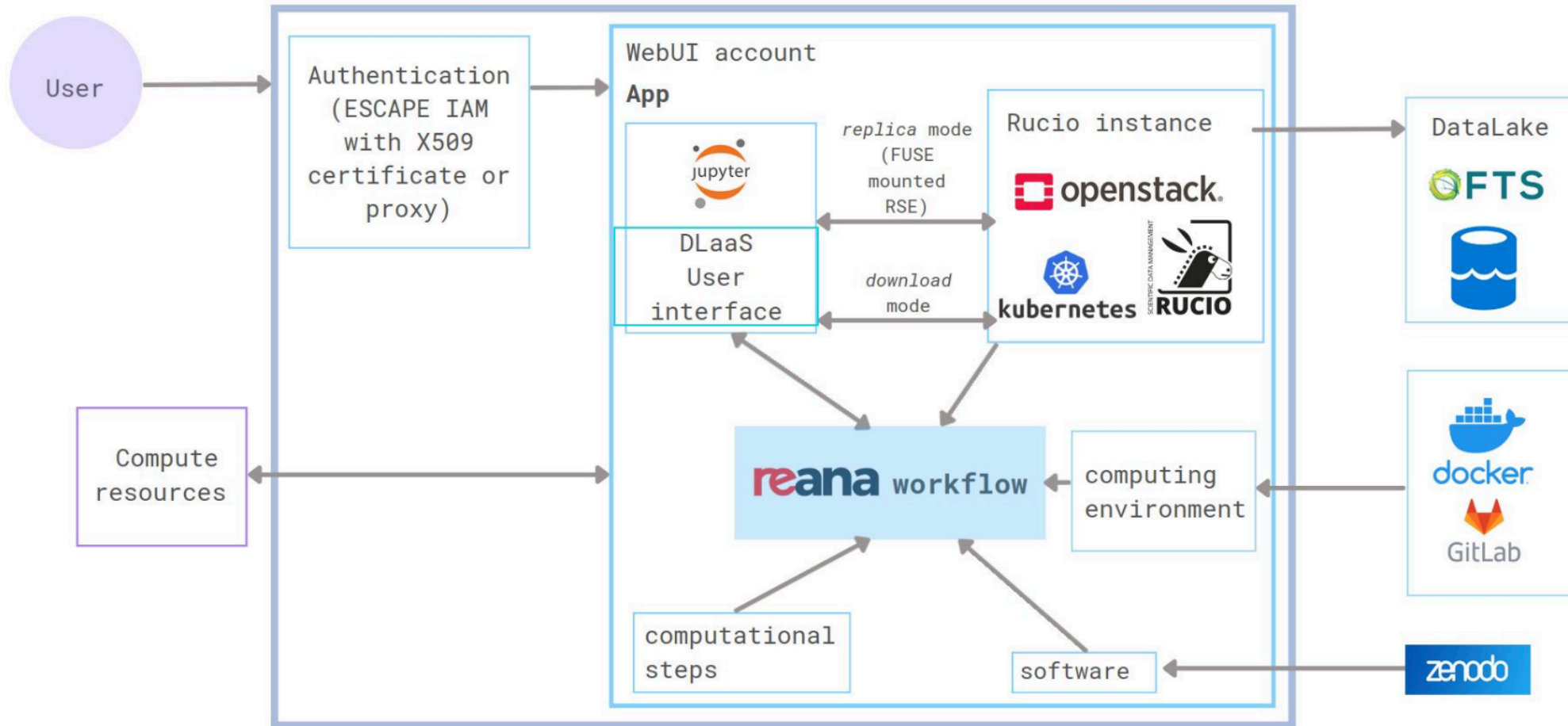


# VRE current status

- Landing page - explaining the idea and linking the SP projects:
  - <https://escape2020.pages.in2p3.fr/virtual-environment/home/>
- Centralised GitLab instance for the Post Doc devs
  - <https://gitlab.in2p3.fr/escape2020/virtual-environment>
- Where scientific env are also 'containerised' to be included later into the VRE as an environment
  - <https://gitlab.in2p3.fr/escape2020/virtual-environment/docker-images>
- Bi-weekly meeting with EU + DM postdocs
  - <https://indico.in2p3.fr/category/1033/>



# VRE vision goals (PoC)





# VRE future steps

- Little scalability → we would need to deploy the whole infrastructure in EOSC resources
- WIP X509 certificated to be substituted (most likely) by tokens
- WIP EOSC integration: Federated AAI
  - That would allow access to embargo/restricted data
- WIP Integration with ESAP, OSSR and VO



# VRE – OSSR integration: GSoC proposal

- Project presented on 03/04/2022 WP2 call.
- Google offers an online “Summer student” program
  - <https://summerofcode.withgoogle.com/>
  - Bring new contributors to Open Source projects.
- ESCAPE OSSR is participating in the program
  - Through the HSF (HEP Software Foundation)
    - <https://hepsoftwarefoundation.org/>
  - “Repository as a Service”
    - [https://hepsoftwarefoundation.org/gsoc/2022/proposal\\_ESCAPE-repository-as-a-service.html](https://hepsoftwarefoundation.org/gsoc/2022/proposal_ESCAPE-repository-as-a-service.html)



# Context / Background

- Building on top and from the DLaaS idea



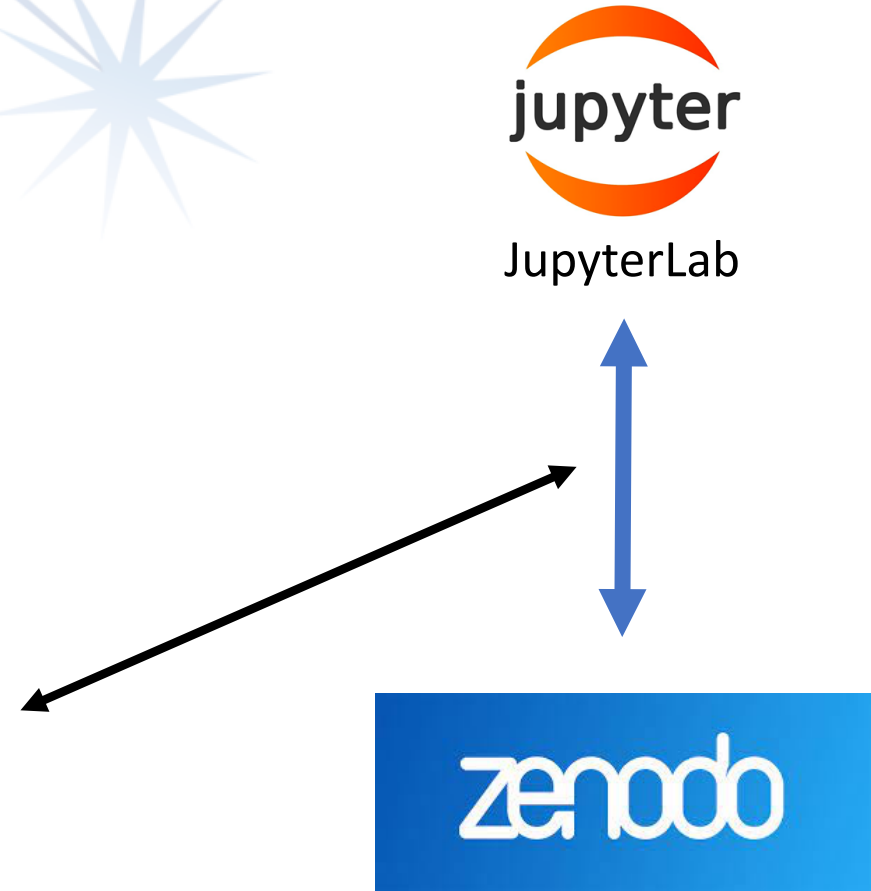
DLaaS  
JupyterLab

## Data Lake as a Service



# "OSSR as a Service"

- Independent JupyterLab extension
  - Installable on any JL environment
    - JupyterHub
    - Personal computer
- Interface that connects JLab with the OSSR through a user friendly GUI
  - Allow searching and listing any OSSR entry
  - Would allow bringing (`wget`) any OSSR entry content
  - Would allow installing any repository in the working environment
  - Would allow publishing back to the OSSR



# (DL+ OSSR)aaS

● Connecting 2 ESCAPE services



DLaaS + OSSRaaS  
JupyterLab

DL + OSSR as a Service

