university of groningen

R3B at project ESCAPE; a case study to practice open science for FAIR/GSI experiments

Maisam M. Dadkan

Posdoc fellow at project ESCAPE

Fairness conference, May 2022

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

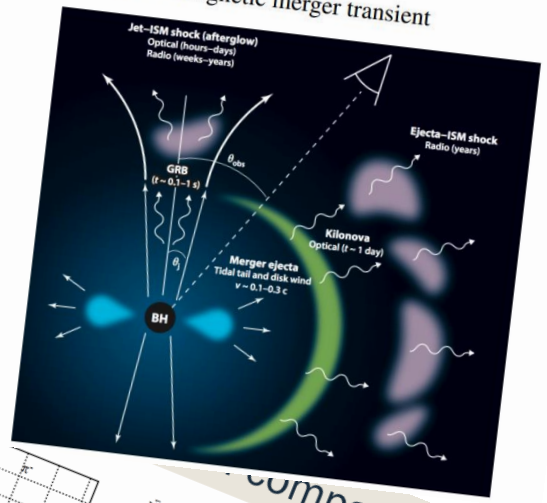
Electromagnetic Counterparts of Neutron Star Mergers: Signatures of Heavy r-Process Nucleosynthesis

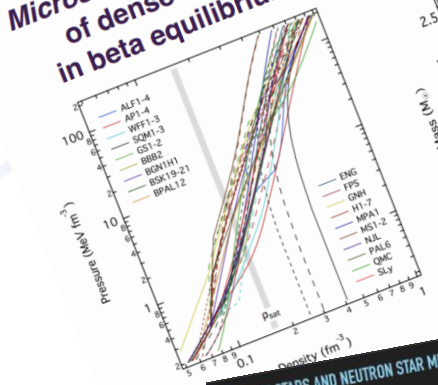
Andreas Flörs (GSI)
Luke Shingles (GSI, QUB), Gerrit Leck (GSI), Gabriel Martínez-Pinedo (GSI)
Ricardo Ferreira da Silva (LIP)

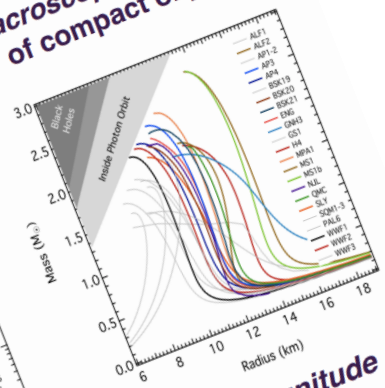
GSI

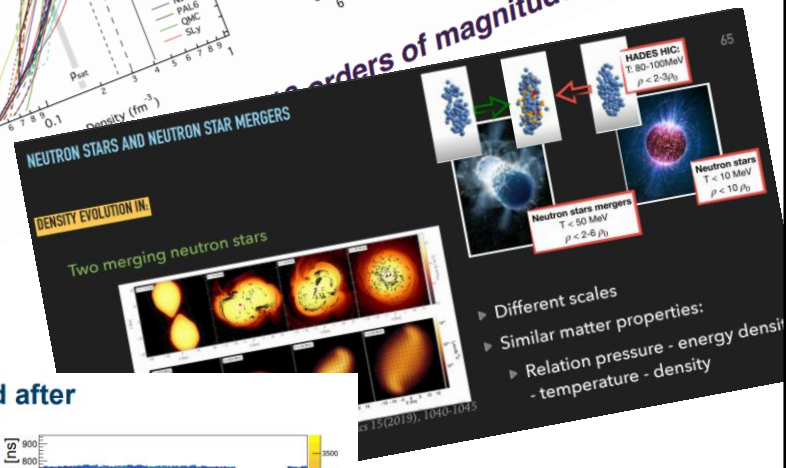
Neutron Star Mergers and Kilonovae

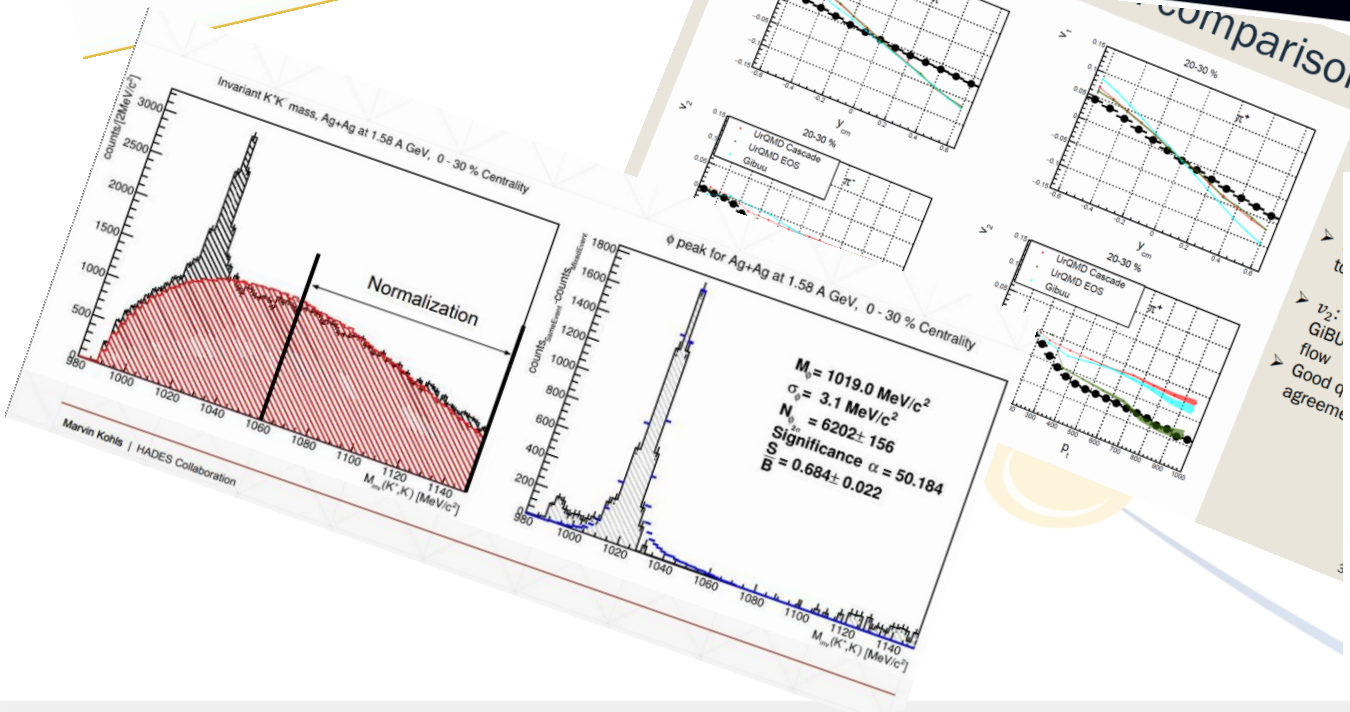
- kilonova: electromagnetic merger transient

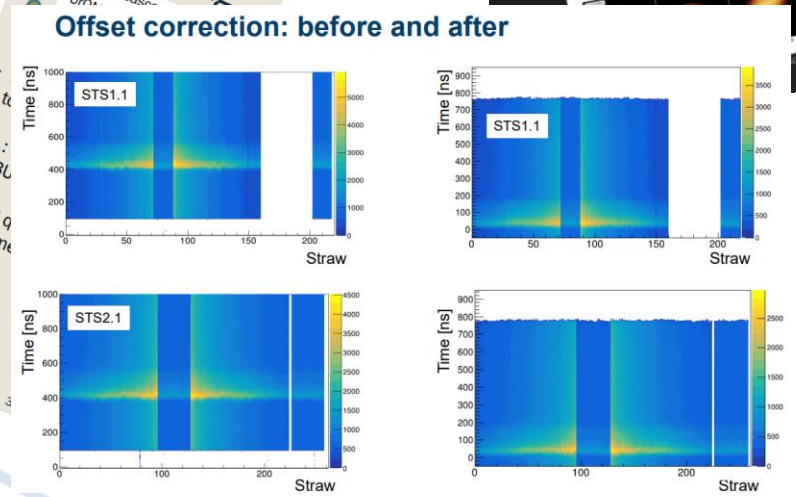












Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures

Neutron Star Mergers and Kilonovae

- kilonova: electromagnetic merger transient

Microscopic properties of dense matter in beta equilibrium

Macroscopic properties of compact objects

Partial wave dispersion relation approach:

- $\sigma/f_0(500)$ in $\pi\pi$ scattering (real & lattice data); $f_0(980)$ in $\{\pi\pi, K\bar{K}\}$
- $\kappa/K^*(700)$ in πK scattering (real & lattice data)

Can search for resonances & bound states given the data input

$\sqrt{s_p} = 458(10)^{+7}_{-15} - i\,256(9)^{+5}_{-8}$ MeV

$\sqrt{s_{f_0(980)}} = 993(2)^{+2}_{-1} - i\,21(3)^{+2}_{-4}$ MeV

$\sqrt{s_p} = 498(21)^{+12}_{-19} - i\,138(13)^{+5}_{-10}$ MeV

$\sqrt{s_p} = 758(5)(0)$ MeV bound state

$\bar{B}\bar{B}$ Decomposition in energy scan

- $B\bar{B}$, $B^*\bar{B}$, $B^*\bar{B}^*$ cross section at various energies
- Higher statistic to improve fits
- JHEP06(2021)137

$B\bar{B}$ $B^*\bar{B}$ $B^*\bar{B}^*$

Orders of magnitude

NEUTRON STAR MERGERS

Merging neutron stars

HADES HIC: T: 80-100MeV $\rho < 2\text{-}3\rho_0$

Neutron stars mergers T < 50 MeV $\rho < 2\text{-}6\ \rho_0$

Neutron stars T < 10 MeV $\rho < 10\ \rho_0$

- Different scales
- Similar matter properties:
- Relation pressure - energy density - temperature - density

Yield Extraction

Double Gaussian function is used for signal

$Fit(m) = Ae^{\frac{-1}{2}\frac{(m-m_0)^2}{\sigma_1^2}} + Be^{\frac{-1}{2}\frac{(m-m_0)^2}{\sigma_2^2}} + pol2(m)$

1. Exclude signal region (m_0<0.43485 & m_0>0.56135) and fit background with $pol2(m)$
2. Use background fit parameters as initial values for next iteration, where signal (double Gaussian) fit function has fixed $m_0 = 0.4976$ GeV/c^2 and widths σ_1=0.004 GeV, σ_2=0.007 GeV
3. Use fit parameters as initial values for unconstrained fit to the whole inv. mass range

Normalization

Marvin Kohls | HADES Collaboration

Step 1 Step 2 Step 3

Divide (p_T,y) phase space into 6x6 bins

Offset correction: before and after

STS1.1 STS2.1

Time [ns] Straw

Current results (different collaborations agree within errors):

$T_c(\mu_B = 0) = 158.0 \pm 0.6$ MeV $\kappa_2 = 0.0153 \pm 0.0018$ $\kappa_4 = 0.00032 \pm 0.00067$

T_c [MeV] μ_B [MeV]

crossover line: $\mathcal{O}(\mu_B^4)$ constant: ϵ freeze-out: STAR ALICE

$n_S = 0$, $\frac{n_Q}{n_B} = 0.4$

Analytical continuation NLO Taylor-expansion LO Taylor-expansion

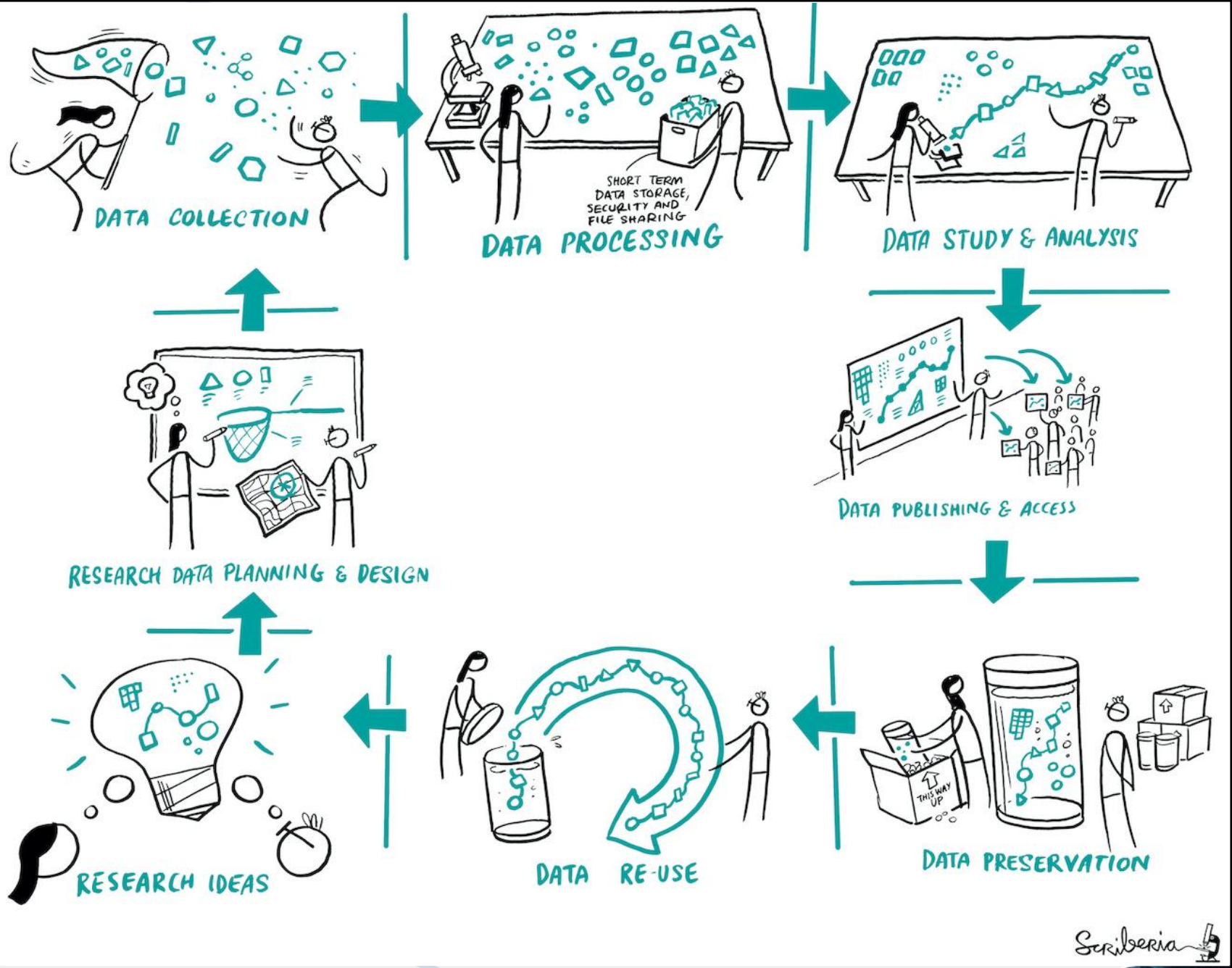
Dyson-Schwinger: hep-ph/1906.11644

Chemical freezeout: nucl-th/0511071v3 nucl-th/1212.2431 hep-ph/1403.4903 nucl-th/1512.08025 nucl-ex/1701.07065

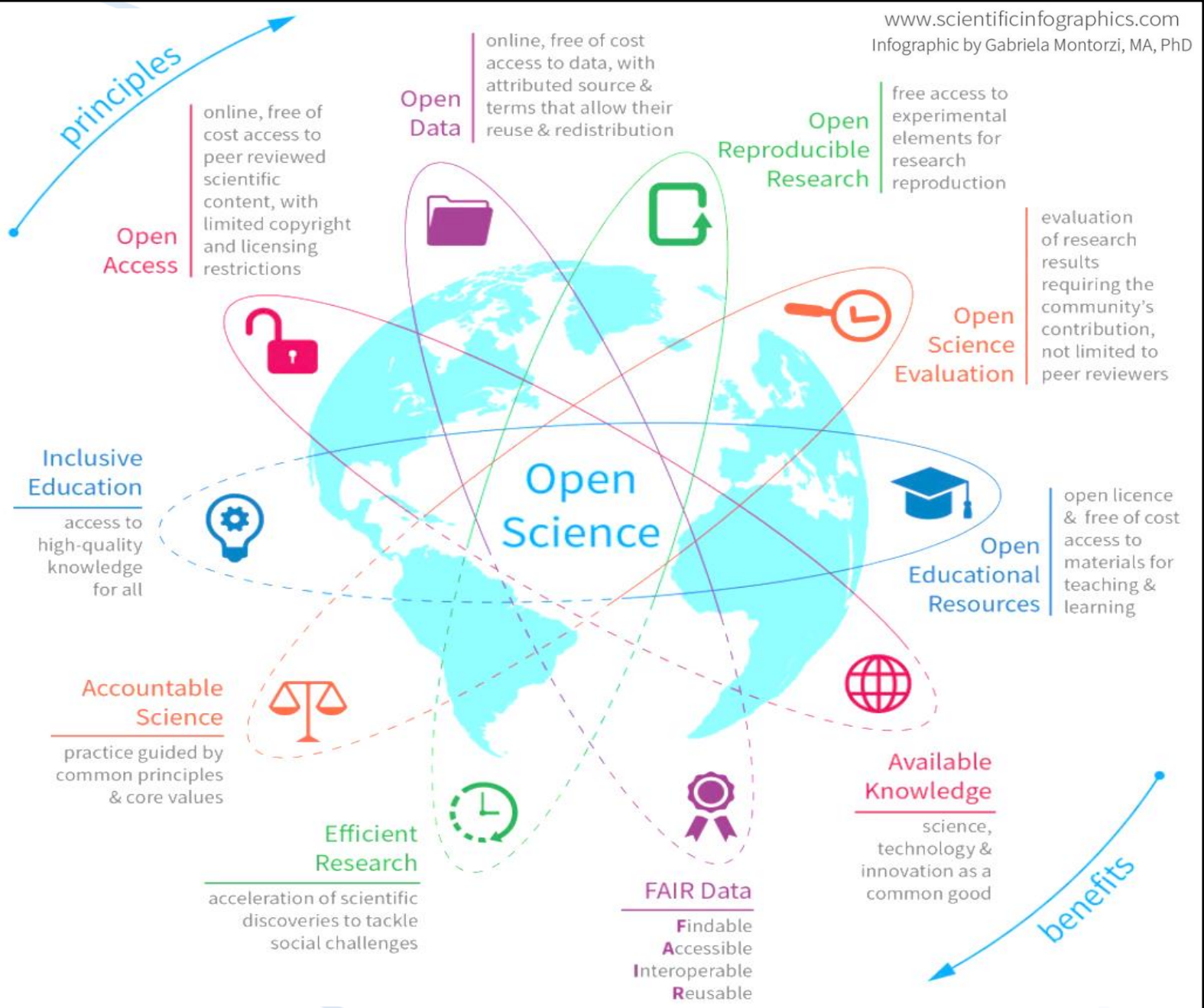
Funded by the European Union's Horizon 2020 - Grant N° 824064

Data = Data+analysis workflow

- Is…
  - your data findable?
  - your data accessible?
  - your data reusable?



Funded by the European Union's Horizon 2020 - Grant N° 824064



Funded by the European Union's Horizon 2020 - Grant N° 824064

Time to unlock the potential of the digital age:

- The key is the internet – digitization
- Open science requires systems thinking
- The value is making new knowledge from connections
- Standards to make (machine) interoperable
- Minimize technical, legal, financial and linguistic barriers
- Collaboration not Competition for Global Goods
- A reset for scientific careers; new incentive measures needed

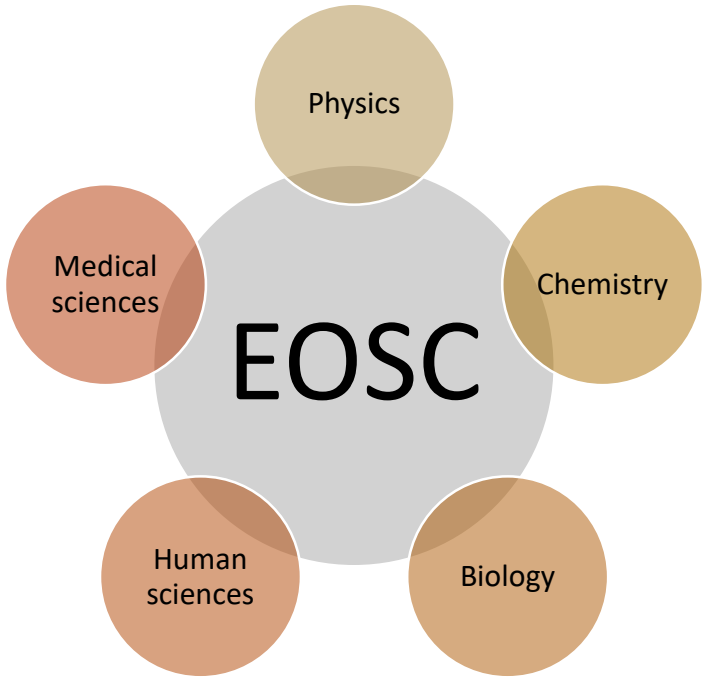
Funded by the European Union's Horizon 2020 - Grant N° 824064

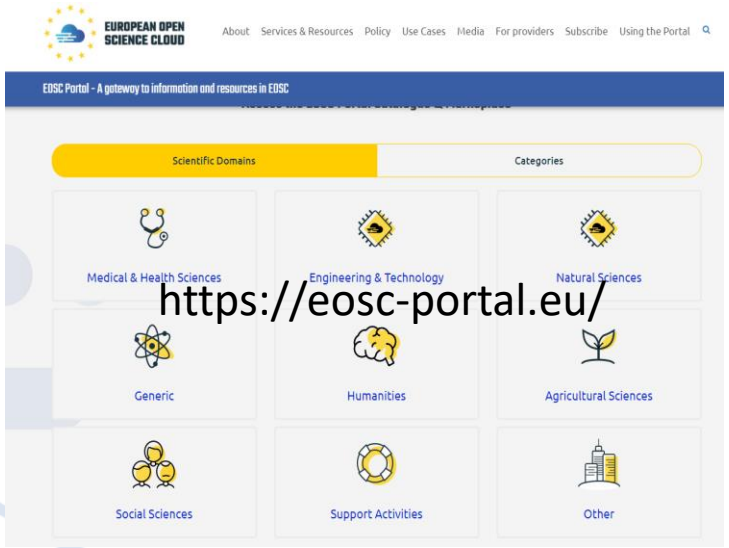
EOSC & ESCAPE

European Open Science Cloud (EOSC)

- Proposed in 2016
- **1.7 million** European researchers
- **€360** million in **48** sub-projects (under Horizon 2020)

- Universal access to data through a **single online platform** based on FAIR
- Make **synergy** between different **scientific research communities** in EU
- **Federating** existing data resources and Research Infrastructures (RIs)
- Accelerating innovations and tools for **computing** and **big-data management**



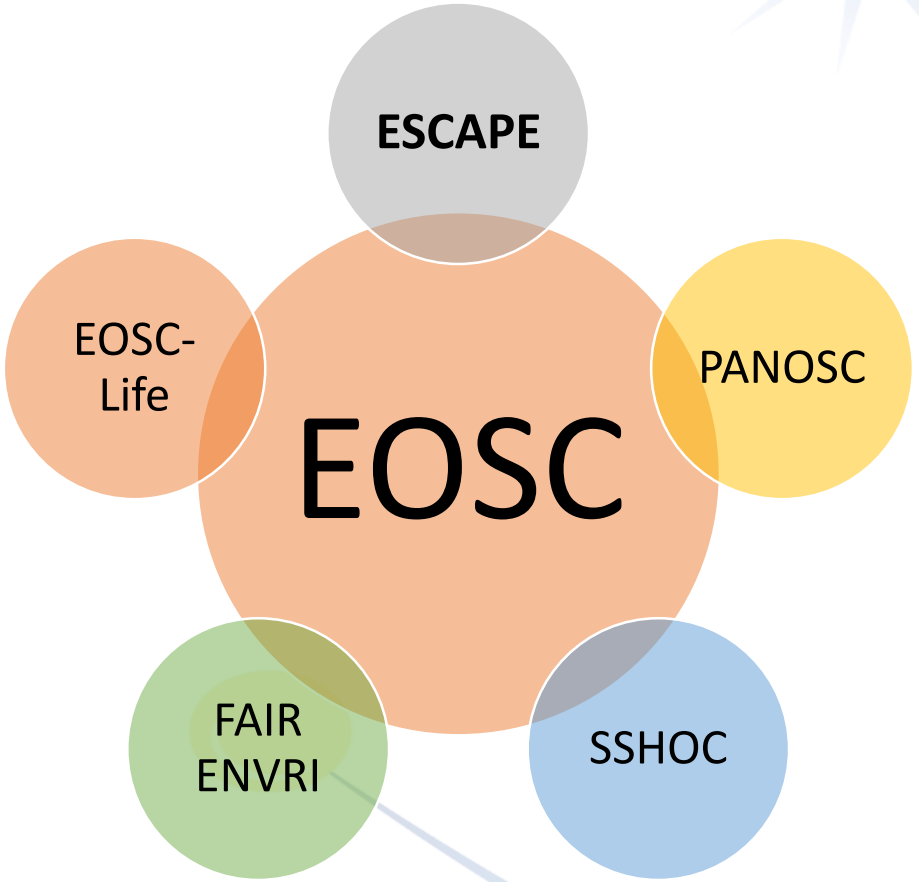


https://eosc-portal.eu/

Funded by the European Union's Horizon 2020 - Grant N° 824064

E S C A P E

(**E**uropean **S**cience **C**luster of
Astronomy & **P**article physics **E**SFRIs)

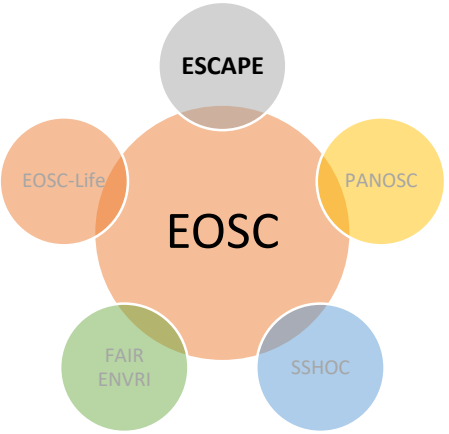


Funded by the European Union's
Horizon 2020 - Grant N° 824064

E S C A P E

(**E**uropean **S**cience **C**luster of
Astronomy & **P**article physics **E**SFRIs)

EOSC & ESCAPE

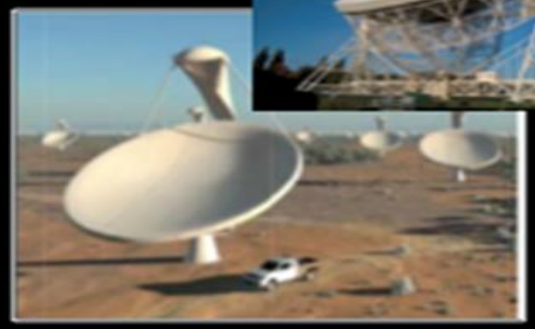


- **31** partners including
- **10** ESFRI projects & landmarks: CTA, EST, FAIR, HL-LHC, KM3NeT, SKA, LSST, VIRGO, ESO, JIVE
- **2** pan-European International Organizations: CERN & ESO with their world-class established infrastructures, experiments and observatories
- **2** European Research Infrastructures: EGO and JIV-ERIC
- **1** involved initiative/infrastructure: EURO-VO
- **4** supporting European consortia: APPEC, ASTRONET, ECFA and NuPECC

- Budget: **16 M€**
- Started: **1/2/2019**
- Duration: **48 months** (end date 31/1/2023)

Funded by the European Union's Horizon 2020 - Grant N° 824064

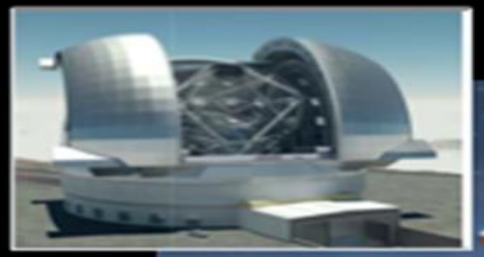
Radio



JIVE-VLBI

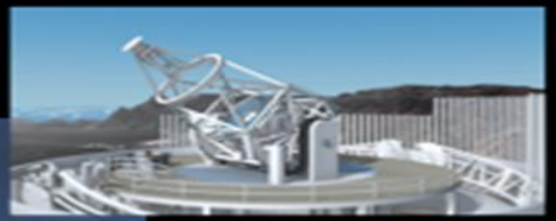
SKA

Visible light



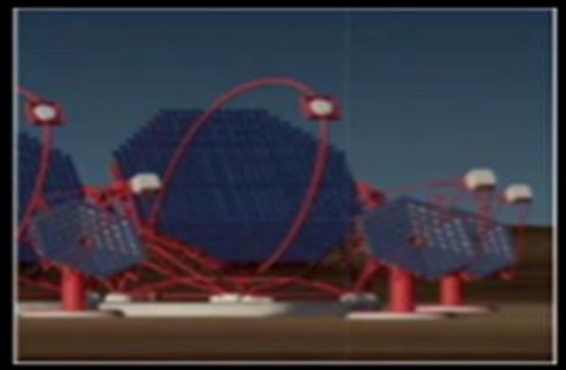
ELT

ESO



EST

Gamma rays



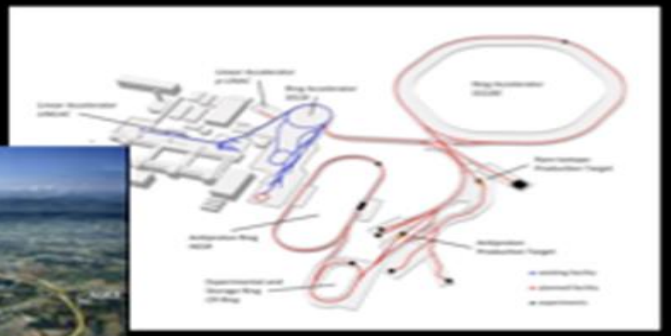
CTA

Accelerator-based
Particle Physics

HL-LHC

CERN

Accelerator-based
Nuclear Physics

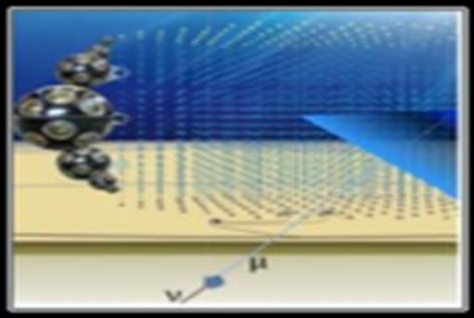


FAIR

Gravitational
Waves

EGO-VIRGO

Cosmic-rays
Neutrinos



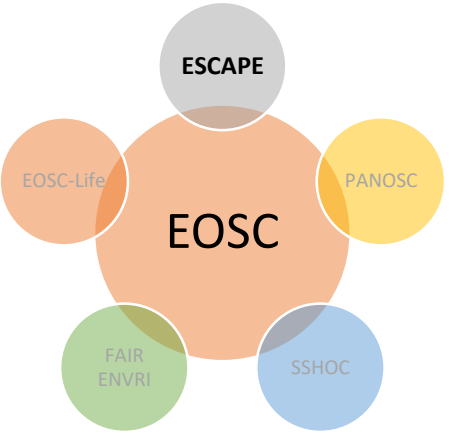
KM3NeT

Funded by the European Union's Horizon 2020 - Grant N° 824064

E S C A P E

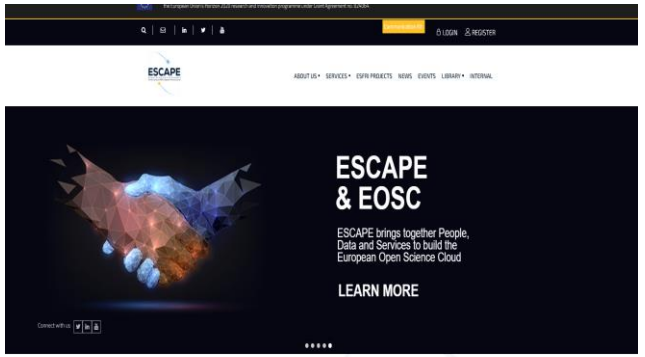
(**E**uropean **S**cience **C**luster of
Astronomy & **P**article physics **E**SFRIs)

EOSC & ESCAPE



- Address the **Open Science challenges** shared by ESFRIs in **astronomy** and **particle physics**
- Developing solutions for the **large data sets** handled by the ESFRI facilities
- Facilitate **interdisciplinary** and **networked research** between different sciences
- **Connecting** ESFRI's projects to **EOSC**
- **Educate** and **train** the scientific and wider user communities

https://projectescape.eu



Select ESCAPE Services

Funded by the European Union's Horizon 2020 - Grant N° 824064

EOSC & ESCAPE

Data Lake:
Build a scalable, federated, data infrastructure as the basis of open science for the ESFRI projects within ESCAPE.

Software Repository:
Repository of "scientific software" as a major component of the "data" to be curated in EOSC.

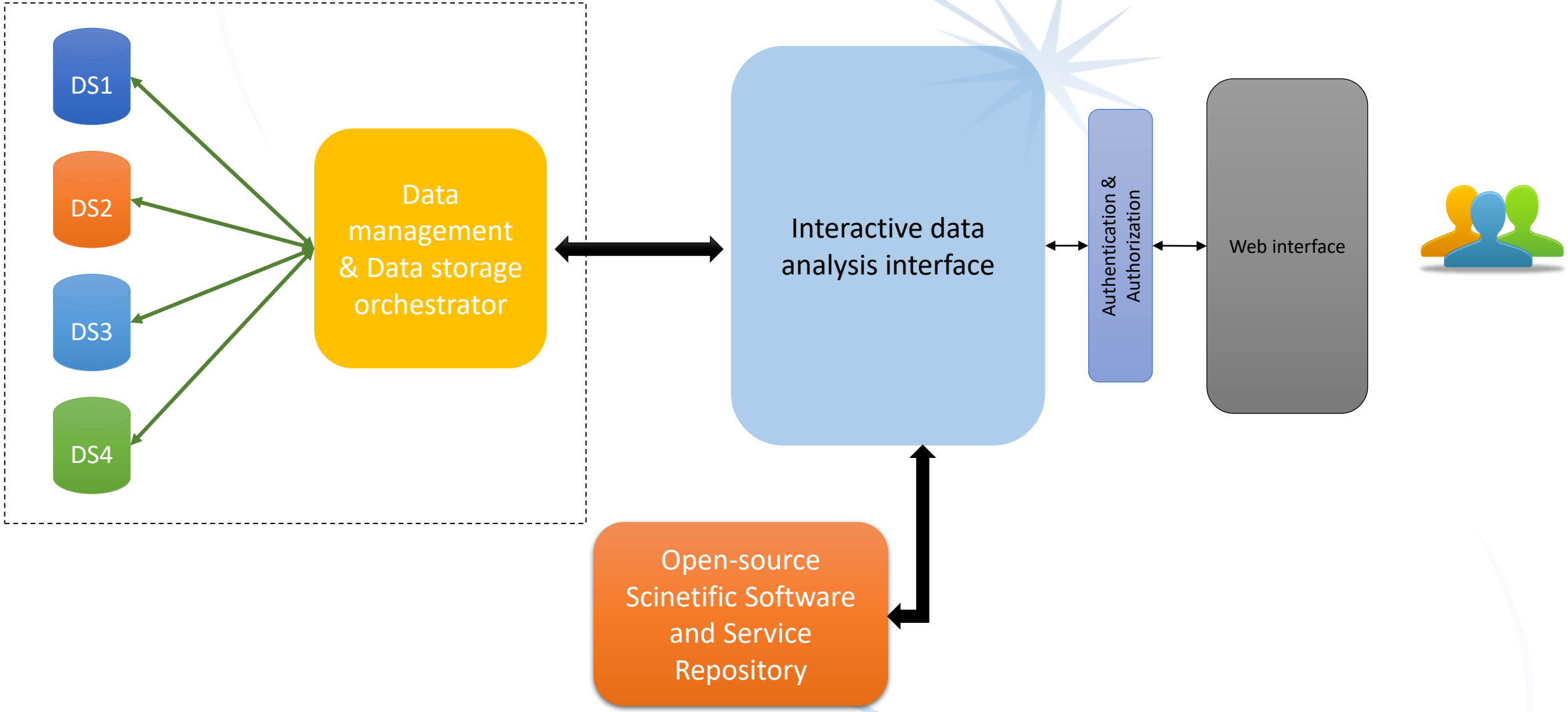
Science Platforms:
Flexible science platforms to enable the open data analysis tailored by and for each facility as well as a global one for transversal workflows.

Citizen Science:
Open gateway for citizen science on ESCAPE data archives and ESFRI community

Virtual Observatory:
Extend the VO *FAIR* standards, methods and to a broader scientific context; prepare the VO to interface the large data volumes of next facilities.

Funded by the European Union's Horizon 2020 - Grant N° 824064

EOSC & ESCAPE

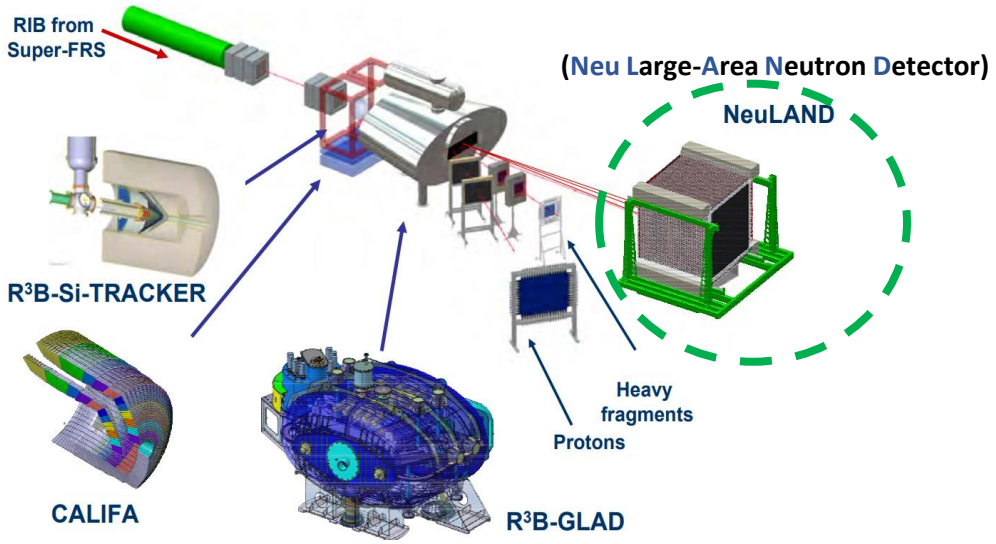


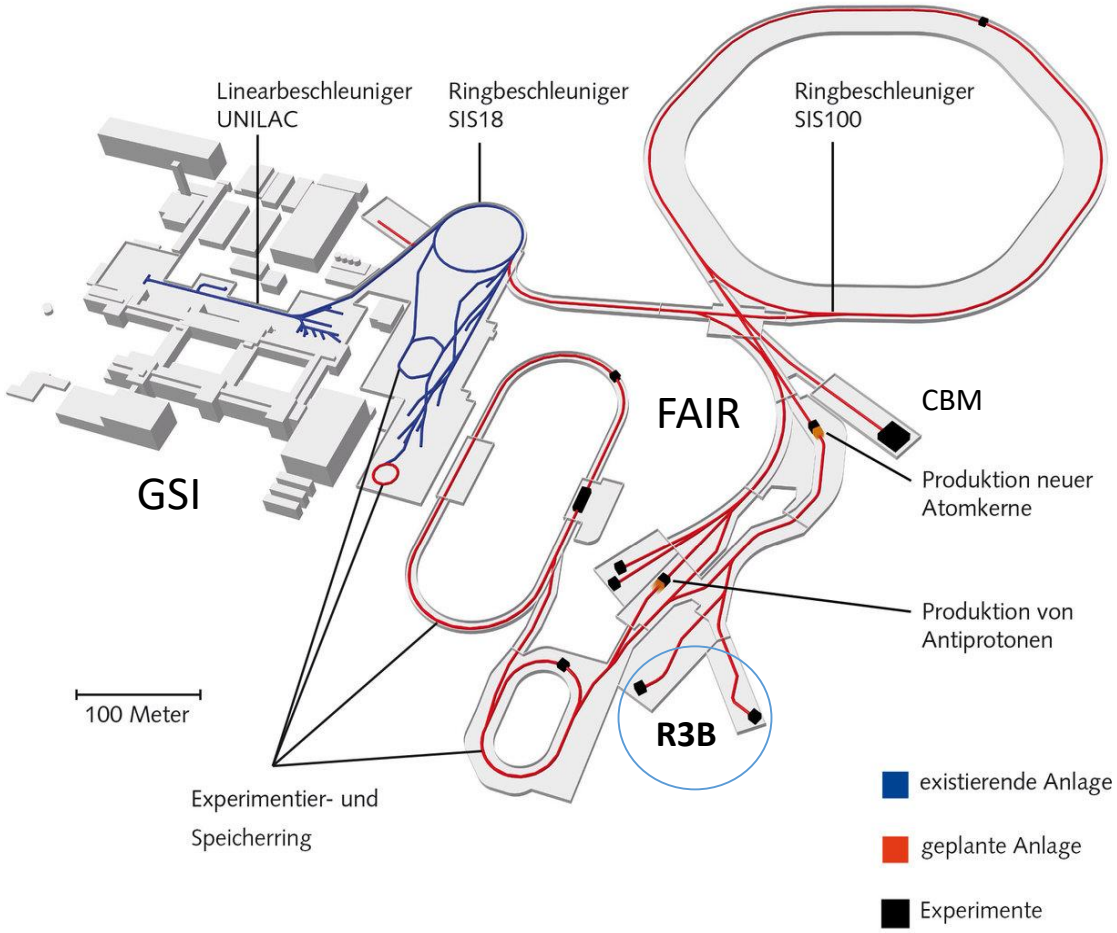
Funded by the European Union's Horizon 2020 - Grant N° 824064

R3B-NeuLAND

R³B

Reactions with Relativistic Radioactive Beams



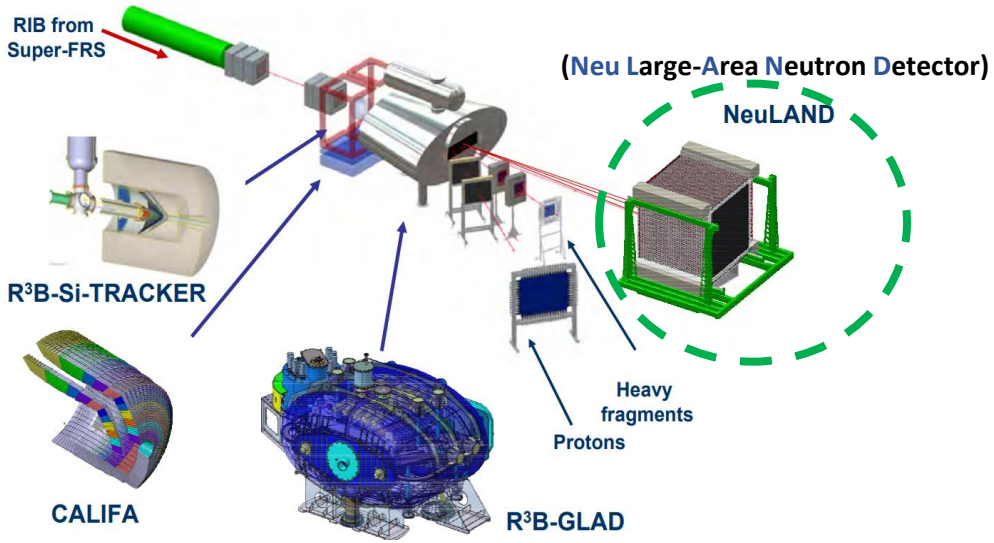


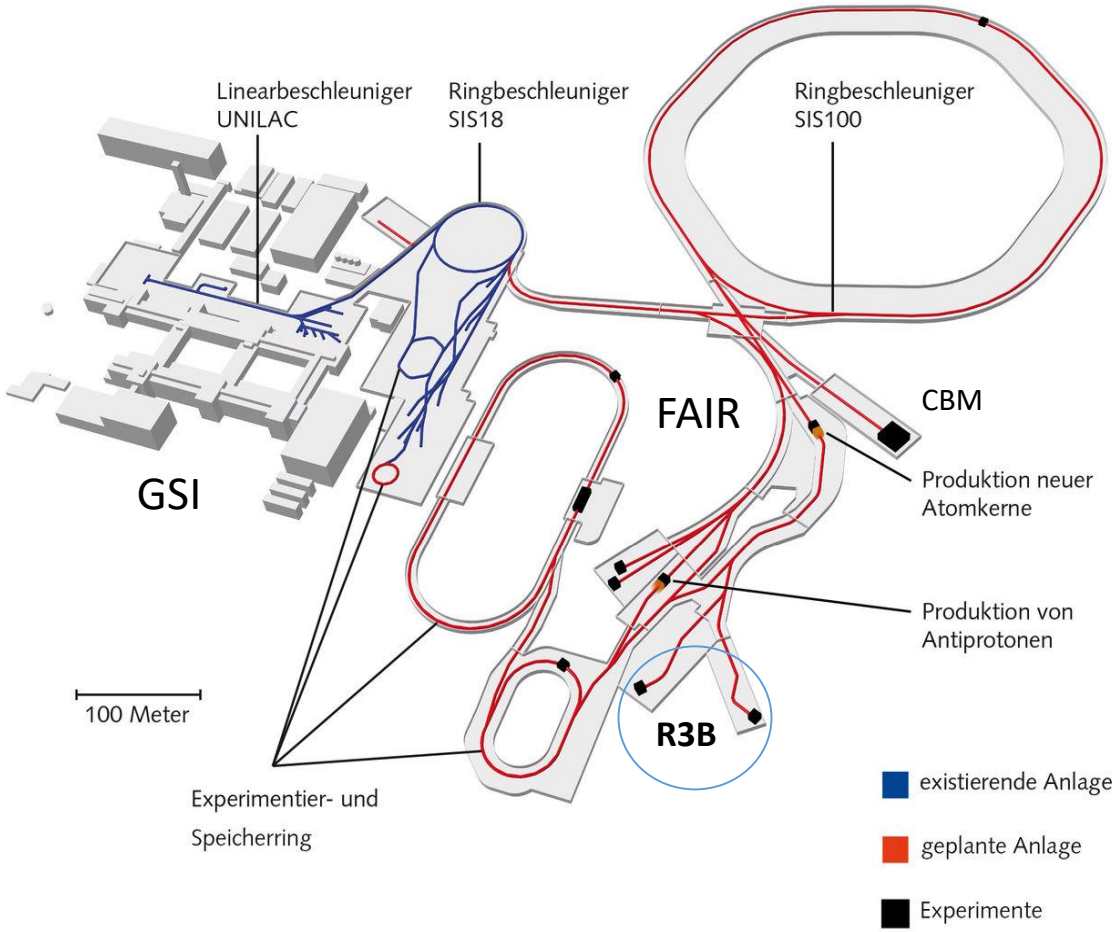
Funded by the European Union's
Horizon 2020 - Grant N° 824064

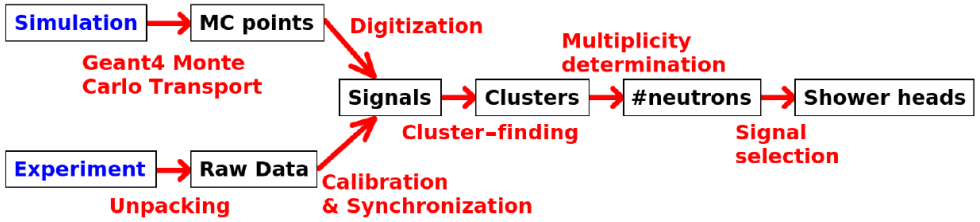
R3B-NeuLAND

R³B

Reactions with Relativistic Radioactive Beams





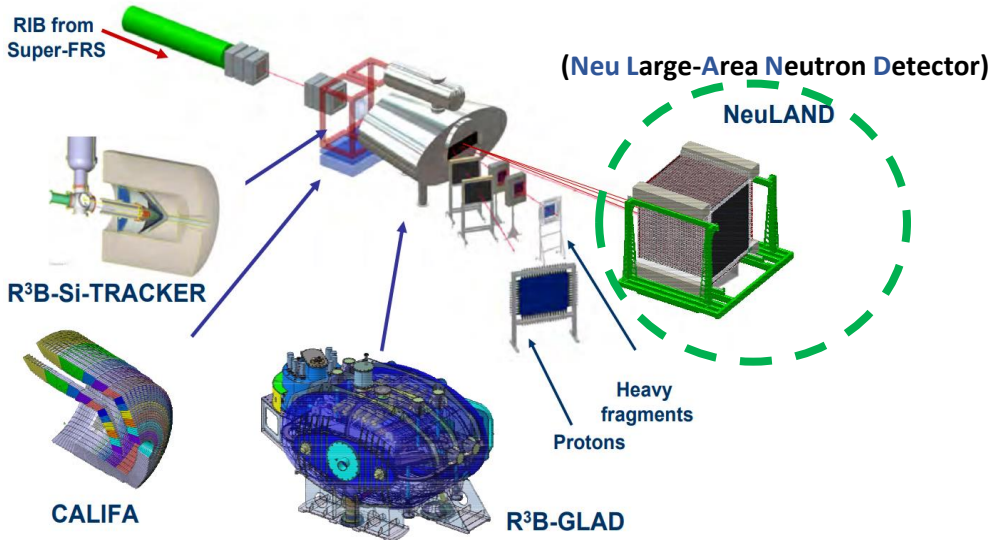


- Multiplicity determination
- Shower head determination

Funded by the European Union's Horizon 2020 - Grant N° 824064

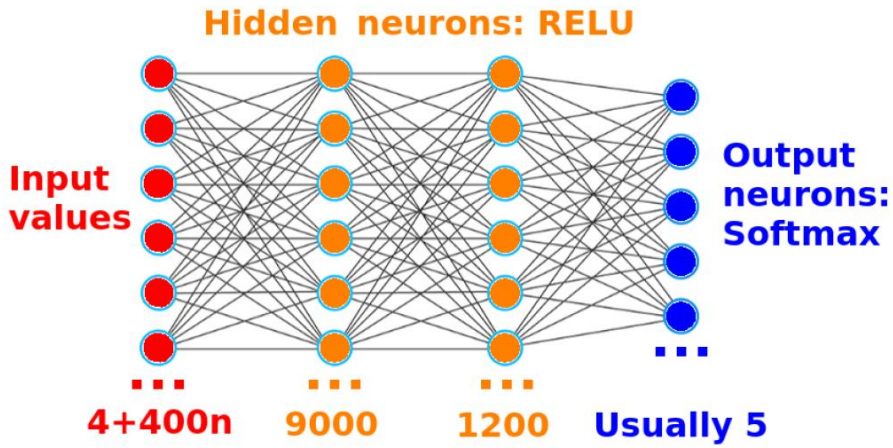
R3B

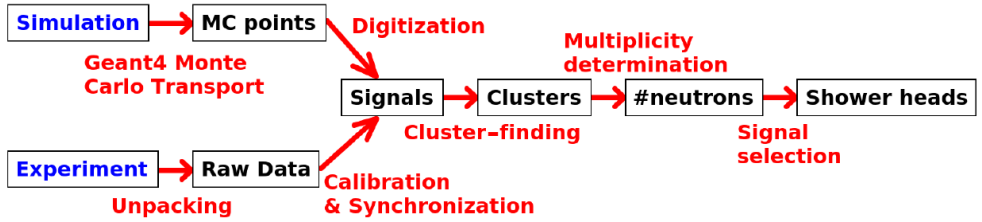
Reactions with Relativistic Radioactive Beams



Analysis methods:

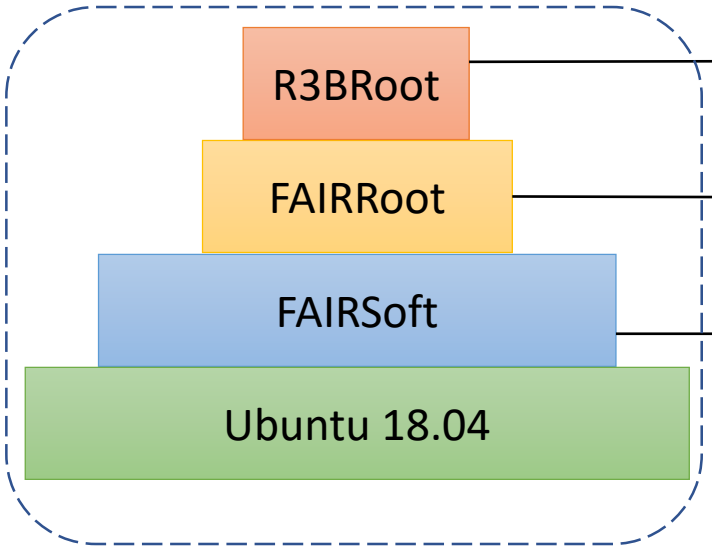
- Technical Design Report (TDR)
- Deep Neural Network (DNN)





- Multiplicity determination
- Shower head determination

Funded by the European Union's
Horizon 2020 - Grant N° 824064



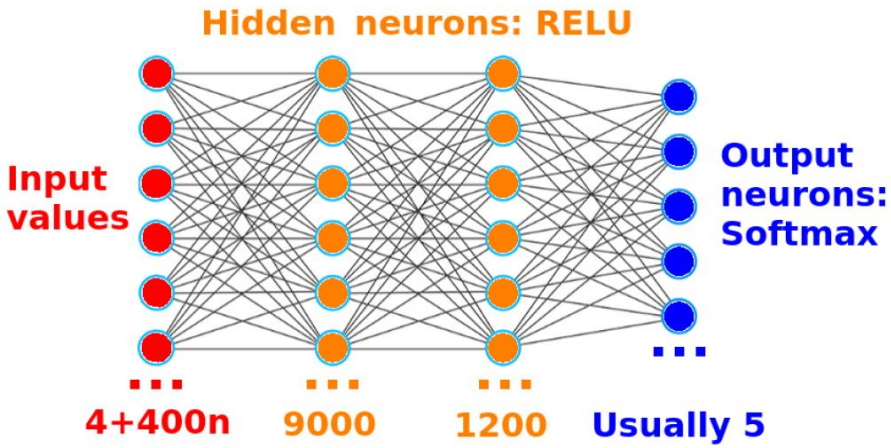
Simulation & analysis libraries for R3B experiment

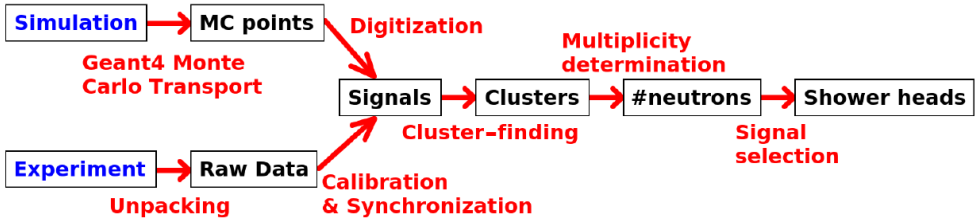
Simulation libraries for FAIR-GSI experiments

Root, Geant3/4,…

Analysis methods:

- Technical Design Report (TDR)
- Deep Neural Network (DNN)

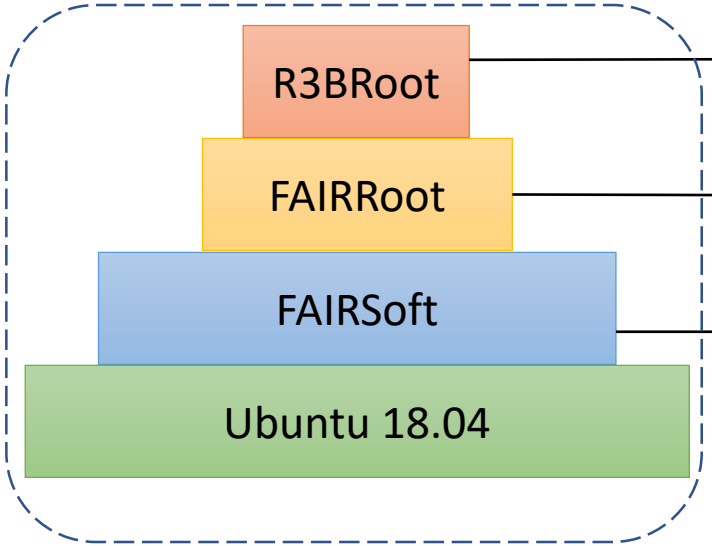




- Multiplicity determination
- Shower head determination

Funded by the European Union's Horizon 2020 - Grant N° 824064

- Complicated installation
- Version conflict
- Difficult to run the simulation
- Not user-friendly



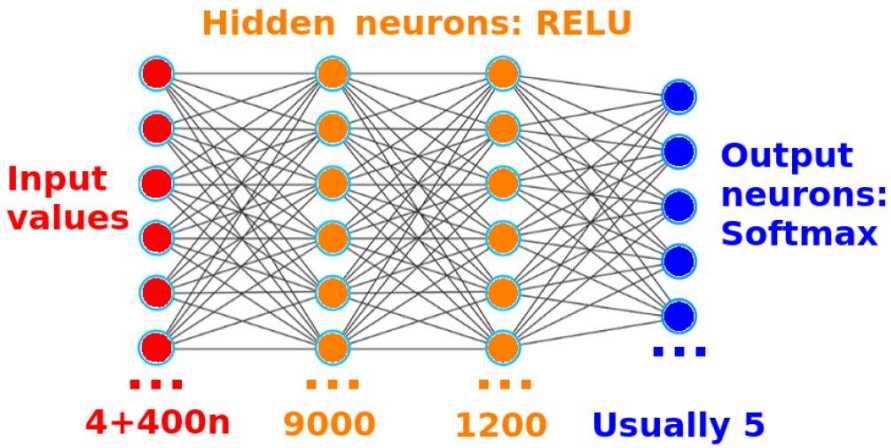
Simulation & analysis libraries for R3B experiment

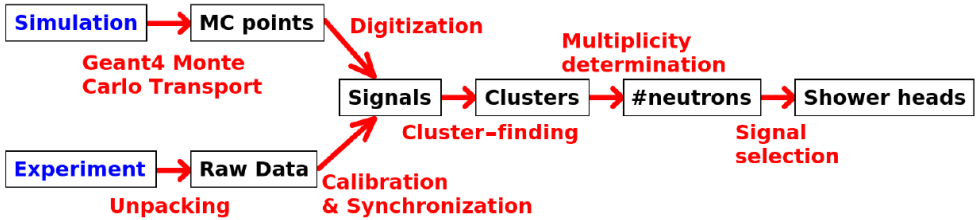
Simulation libraries for FAIR-GSI experiments

Root, Geant3/4,…

Analysis methods:

- Technical Design Report (TDR)
- Deep Neural Network (DNN)



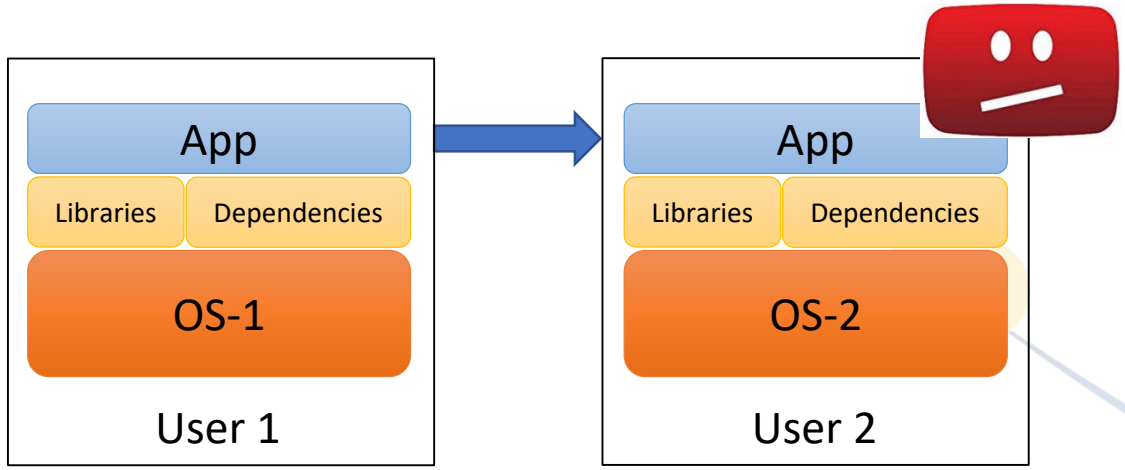


- Multiplicity determination
- Shower head determination

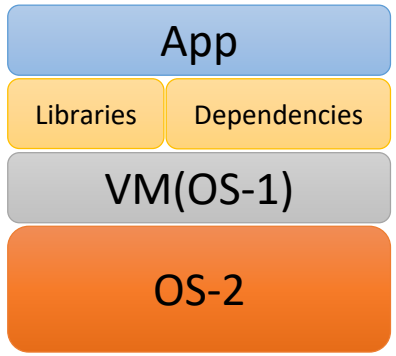
Funded by the European Union's Horizon 2020 - Grant N° 824064

R3BRoot software portability

Compatibility and portability issues



Virtual Machine?!

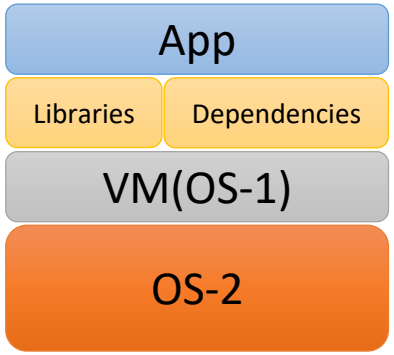


Funded by the European Union's Horizon 2020 - Grant N° 824064

R3BRoot software portability

1. Compatibility and portability issues
2. Configuration is time consuming
3. Scaling
4. Resource utilization

Virtual Machine?!



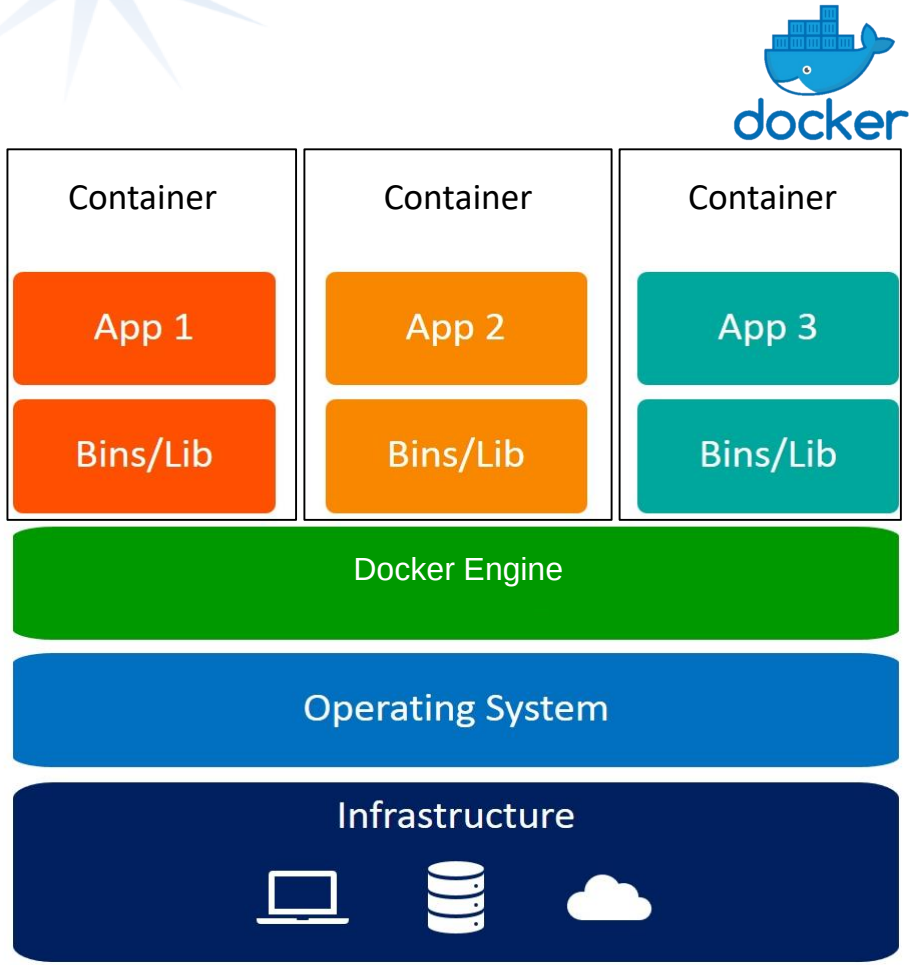
User 2

Funded by the European Union's Horizon 2020 - Grant N° 824064

Containerization:

A standard OS-virtualization technology for **packing up software** and all the **dependencies** to **run applications on different environments**.

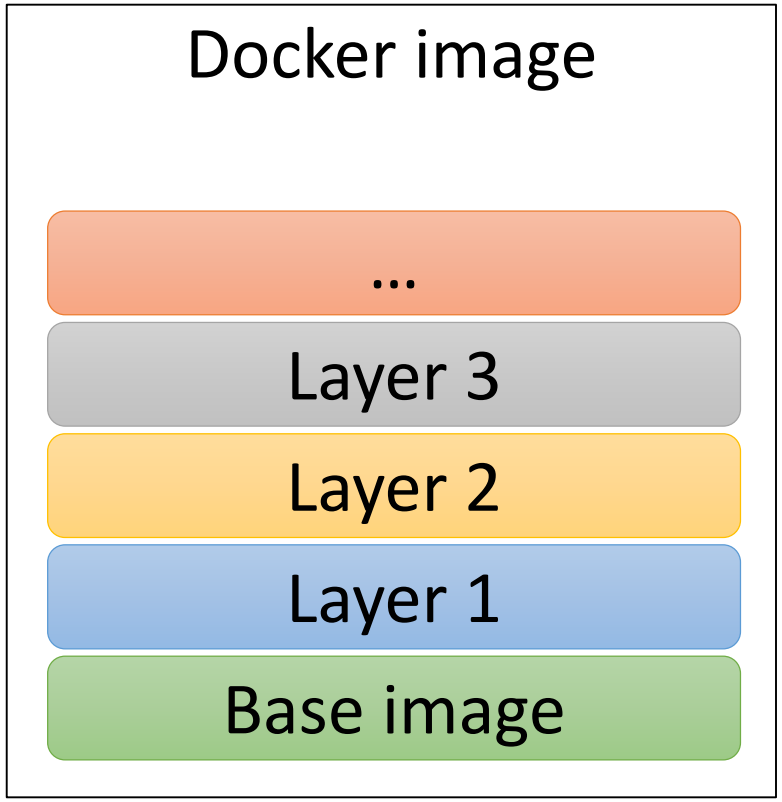
- Lighter than VM
- Faster configuration
- Less resource utilization
- Portability
- Micro services & Scaling



Funded by the European Union's Horizon 2020 - Grant N° 824064

R3BRoot software portability

Containerization:



Docker file

Activities Text Editor di 17:39

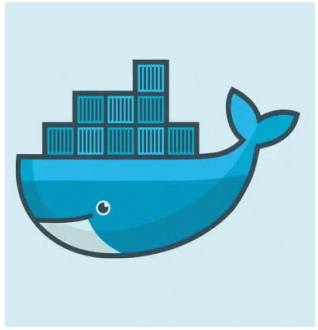
Open Fairsoft.Dockerfile
~/r3bdocker/r3bdocker_final/Version1.4

```
1 ARG BASE_IMAGE=ubuntu:18.04
2 FROM $BASE_IMAGE
3
4 MAINTAINER Maisam M. Dadkan, Email:maisam.m.dadkan@gmail.com
5 #
6 # Install dependencies
7 #
8
9 RUN     apt-get update && \
10        apt-get install -y cmake cmake-data g++ gcc gfortran \
11        debianutils build-essential make patch sed \
12        libx11-dev libxft-dev libxext-dev libxpm-dev libxmu-dev \
13        libglu1-mesa-dev libgl1-mesa-dev \
14        libncurses5-dev curl libcurl4-openssl-dev bzip2 libbz2-dev gzip unzip tar \
15        subversion git xutils-dev flex bison lsb-release python-dev python3-dev\
16        libc6-dev-i386 libxml2-dev wget libssl-dev libkrb5-dev \
17        automake autoconf libtool zlib1g-dev \
18        libreadline-dev libsqlite3-dev llvm \
19        libncursesw5-dev xz-utils liblzma-dev python-openssl python3-openssl &&\
20        apt-get clean && rm -rf /var/lib/apt/lists/*
21
22 #
23 # Steal newer CMake
24 #
25 WORKDIR /opt
26 RUN wget https://github.com/Kitware/CMake/releases/download/v3.16.4/cmake-3.16.4-Linux-x86_64.sh -O cmake.sh && chmod +x cmake.sh && \
27     mkdir /opt/cmake-3.16.4 && ./cmake.sh --skip-license --prefix=/opt/cmake-3.16.4 && \
28     rm cmake.sh
29 ENV PATH="/opt/cmake-3.16.4/bin:${PATH}"
30
31 #copy modified source code of fairsoft and install it
32 WORKDIR /tmp/fairsoft-build
33 COPY    ./fairsoft.conf ./fairsoft.conf
34 COPY    ./FairSoft-jun19p2 ./FairSoft
35 RUN     cd FairSoft &&  ./configure.sh ../fairsoft.conf && ./make_clean.sh all && rm -rf /tmp/fair*
36
37 ENV SIMPATH="/opt/fairsoft_jun19p2"\
38     G4INCLDATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/G4INCL1.0"\
39     G4LEVELGAMMADATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/PhotonEvaporation5.3"\
40     G4RADIOACTIVEDATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/RadioactiveDecay5.3"\
41     G4PIIDATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/G4PII1.3"\
42     G4SAIDXSDATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/G4SAIDDATA2.0"\
43     G4ABLADATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/G4ABLA3.1"\
44     G4REALSURFACEDATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/RealSurface2.1.1"\
45     G4NEUTRONHPDATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/G4NDL4.5"\
46     G4PARTICLEXSDATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/G4PARTICLEXS1.1"\
47     G4ENSDFSTATEDATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/G4ENSDFSTATE2.2"\
48     G4LEDATA="/opt/fairsoft_jun19p2/share/Geant4-10.5.1/data/G4EMLOW7.7"
49
```


Funded by the European Union's Horizon 2020 - Grant N° 824064

R3BRoot software portability

A standard OS-virtualization platform for **packing up software** and all the **dependencies** to **run applications on different environments**.

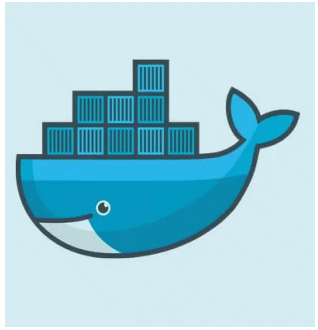


- Reusability
- Accessibility
- Interoperability
- Portability
- Micro services & Scaling

Funded by the European Union's Horizon 2020 - Grant N° 824064

R3BRoot software portability

A standard OS-virtualization platform for **packing up software** and all the **dependencies** to **run applications on different environments**.

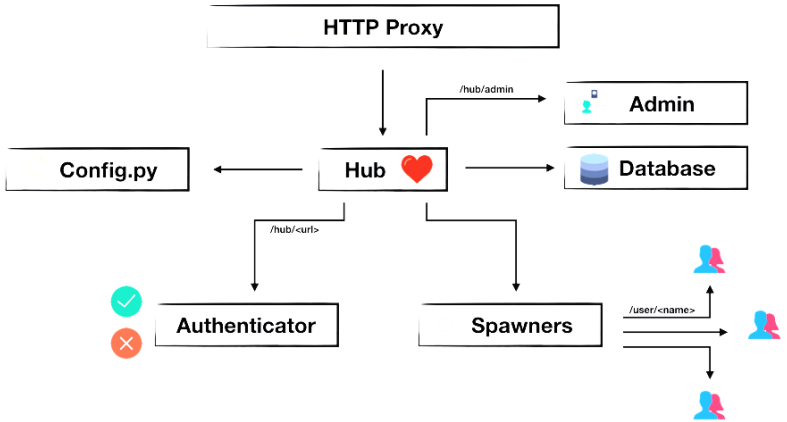


A **Web IDE** for **live code**, equations, **visualizations** and **narrative text (Jupyter Notebook)**

Brings the power of **notebooks** to **groups of users** & handles users access to **computational environments** and ... **(Jupyter Hub)**

- Reusability
- Accessibility
- Interoperability
- Portability
- Micro services & Scaling

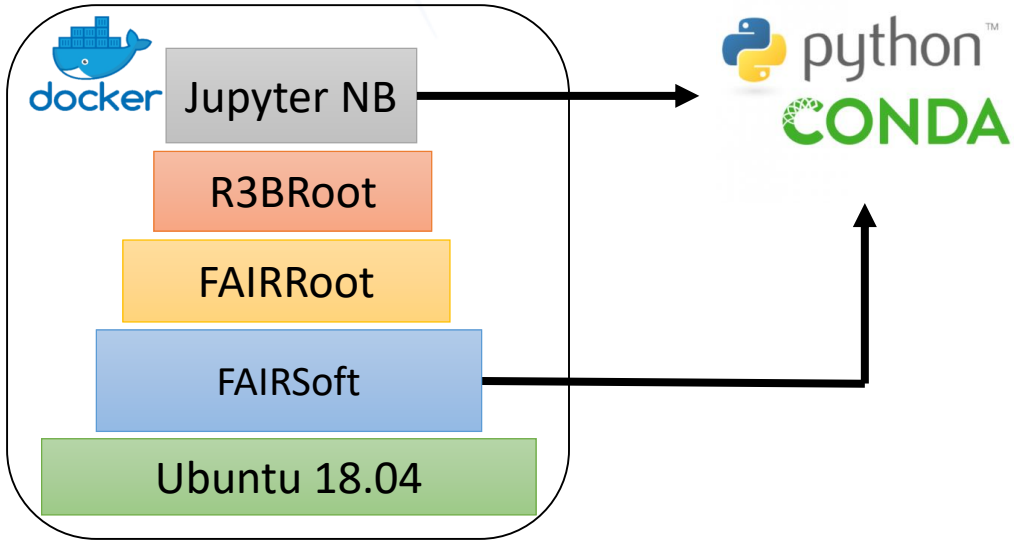
JupyterHub



Funded by the European Union's Horizon 2020 - Grant N° 824064

R3BRoot software portability

✓ **Standard Python environment**

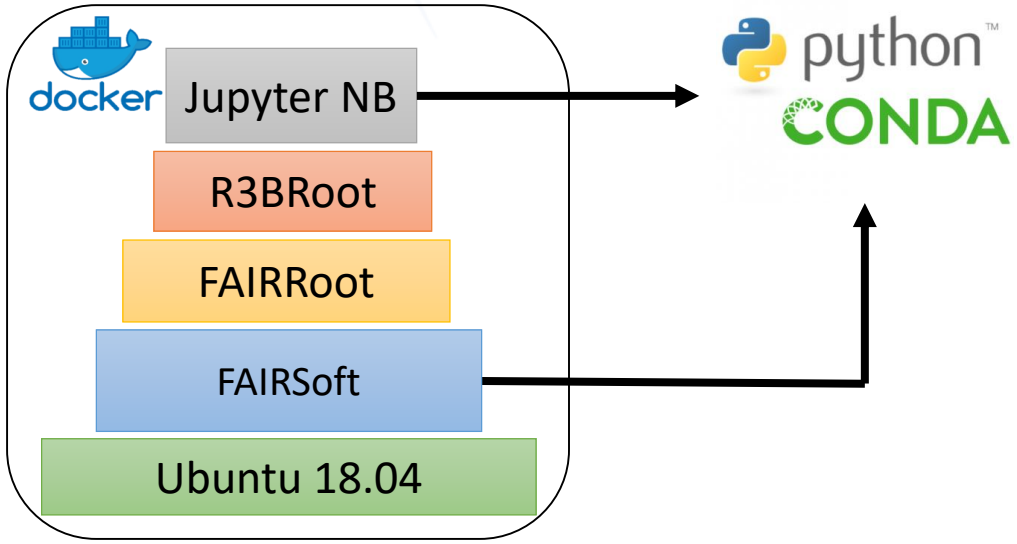


- New FAIRSoft (some bugs are fixed)
- PyRoot and Jupyter NB use the same python version (ensure compatibility)
- Docker images are made for each part (can be used in other FAIR projects)
- Some security issues were resolved (root access limited)

Funded by the European Union's Horizon 2020 - Grant N° 824064

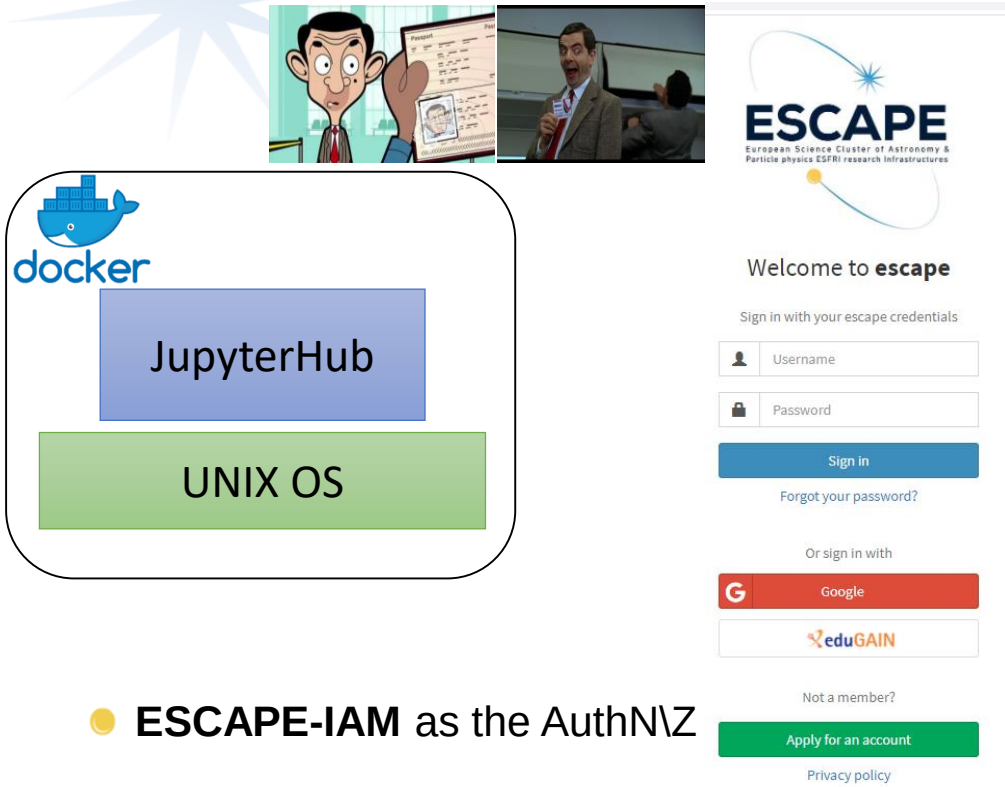
R3BRoot software portability

✓ Standard Python environment



- New FAIRSoft (some bugs are fixed)
- PyRoot and Jupyter NB use the same python version (ensure compatibility)
- Docker images are made for each part (can be used in other FAIR projects)
- Some security issues were resolved (root access limited)

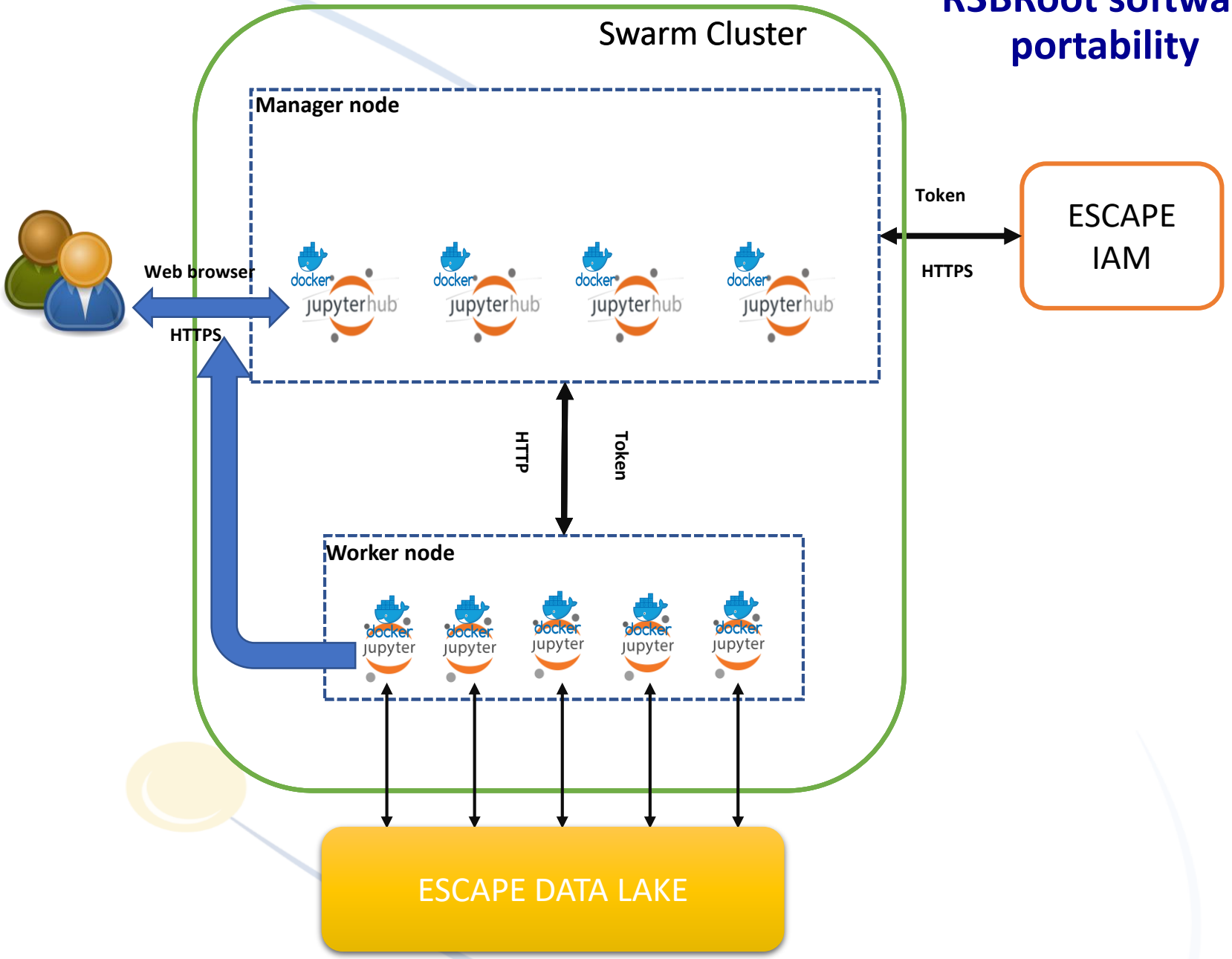
✓ Integrated with ESCAPE AuthN/Z (IAM)



- **ESCAPE-IAM** as the AuthN\Z
- To take advantage of other ESCAPE services
- Enabling the end-to-end AuthN/Z
- Supported **token-based** AuthN/Z

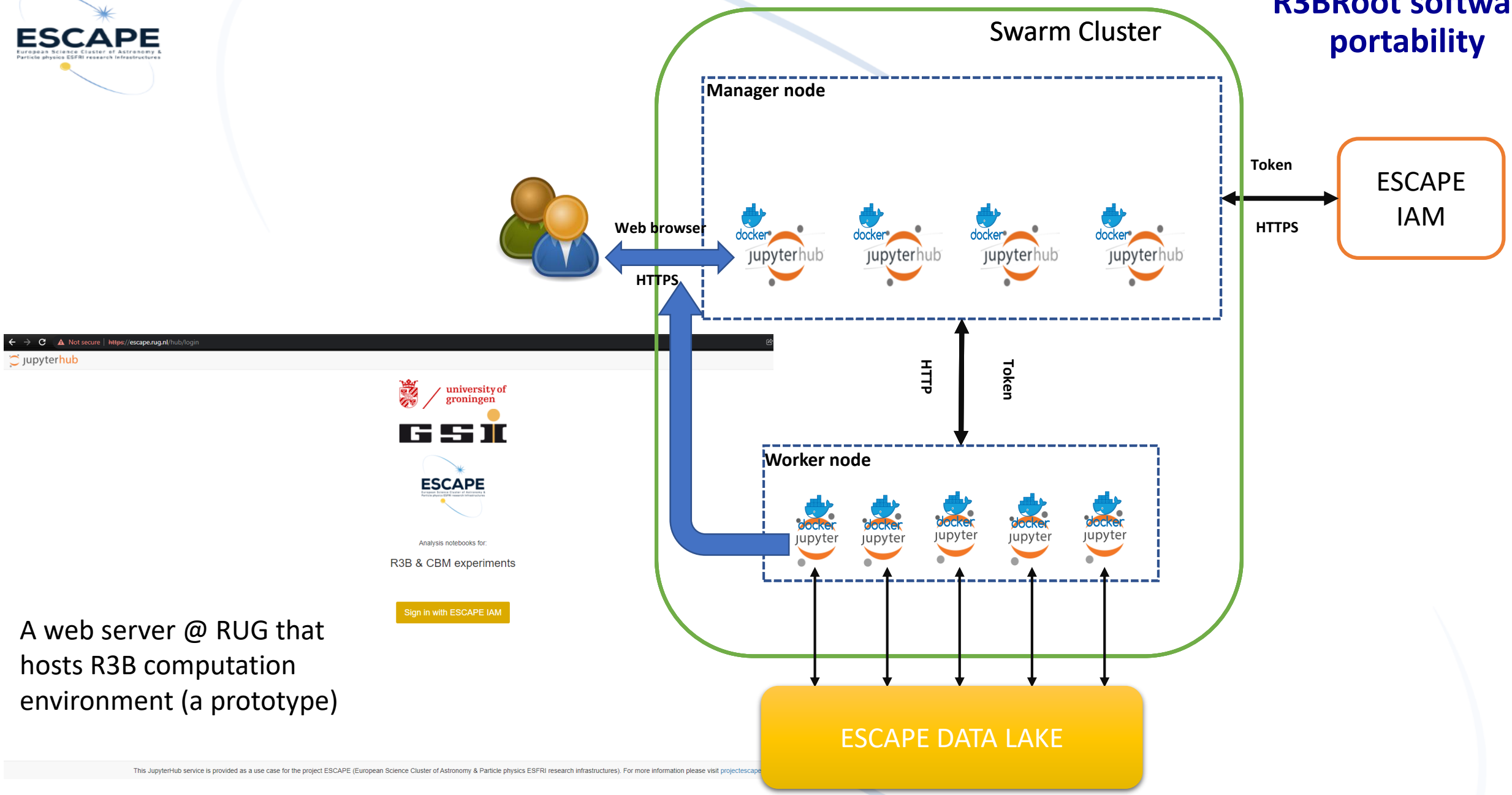
Funded by the European Union's Horizon 2020 - Grant N° 824064

R3BRoot software portability



Funded by the European Union's Horizon 2020 - Grant N° 824064

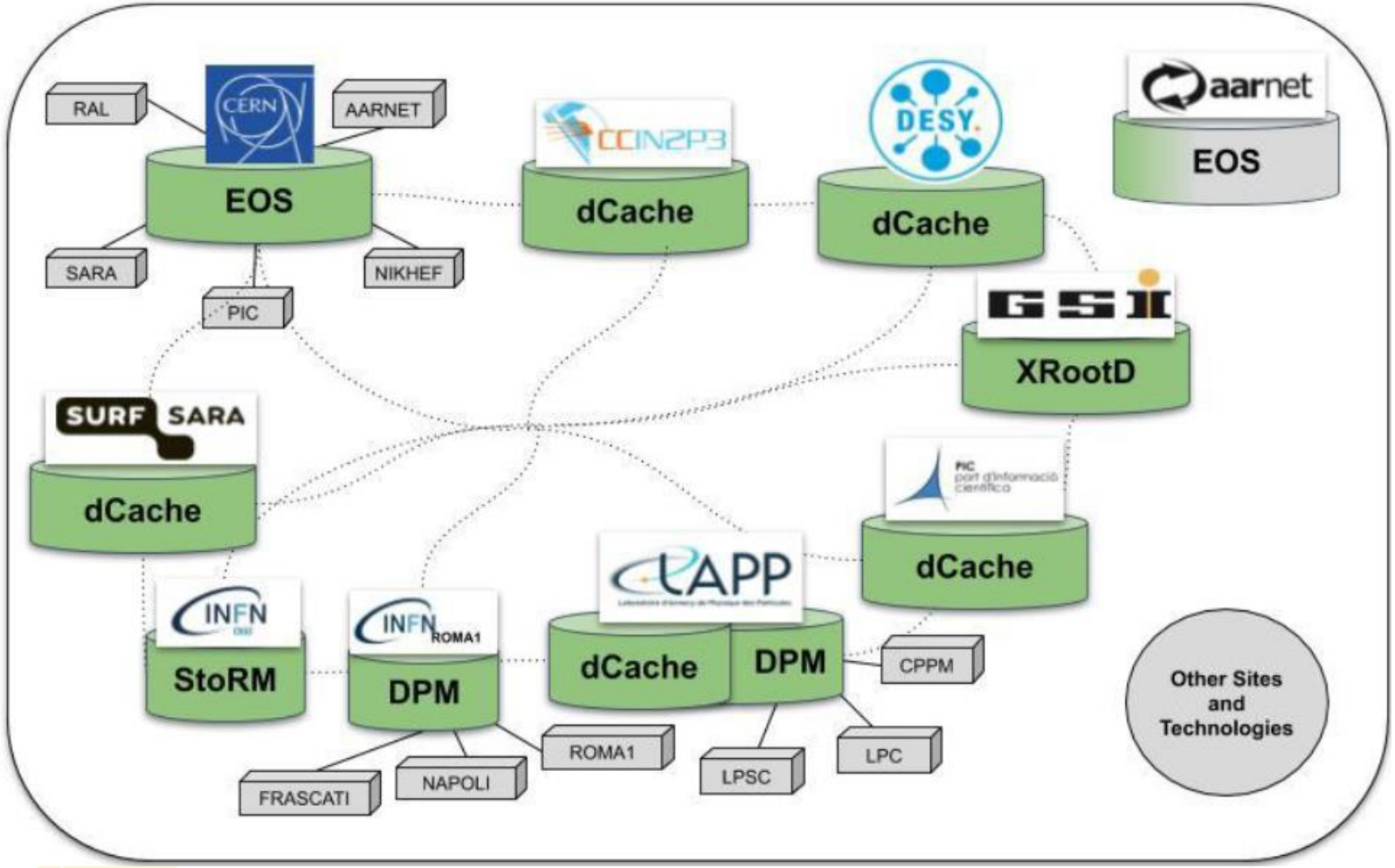
R3BRoot software portability



A web server @ RUG that hosts R3B computation environment (a prototype)

Funded by the European Union's Horizon 2020 - Grant N° 824064

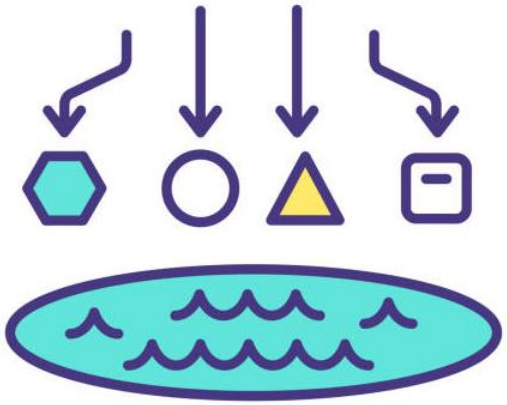
Data Lake



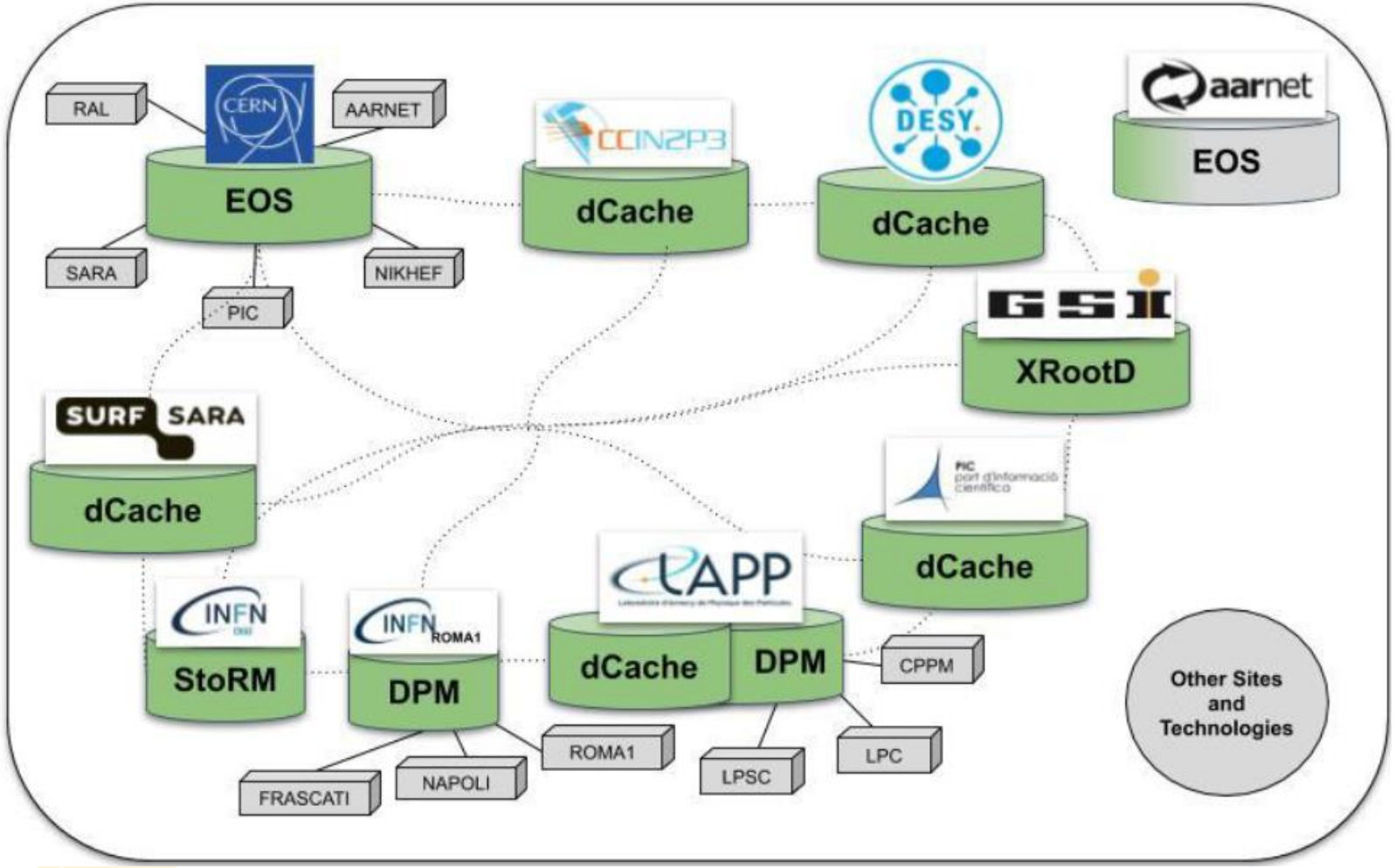
- Distributed across the world
- Using different technologies for data management
- Each one has a different regulation and policy

Funded by the European Union's Horizon 2020 - Grant N° 824064

Data Lake



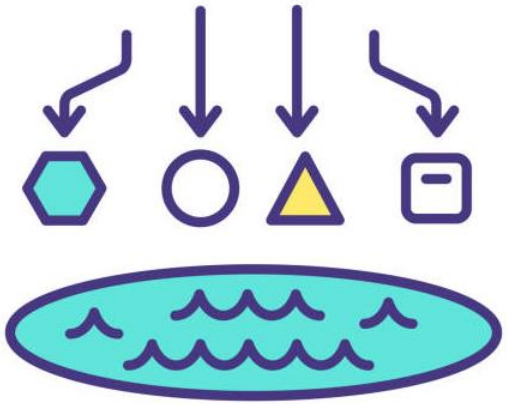
Data Lake



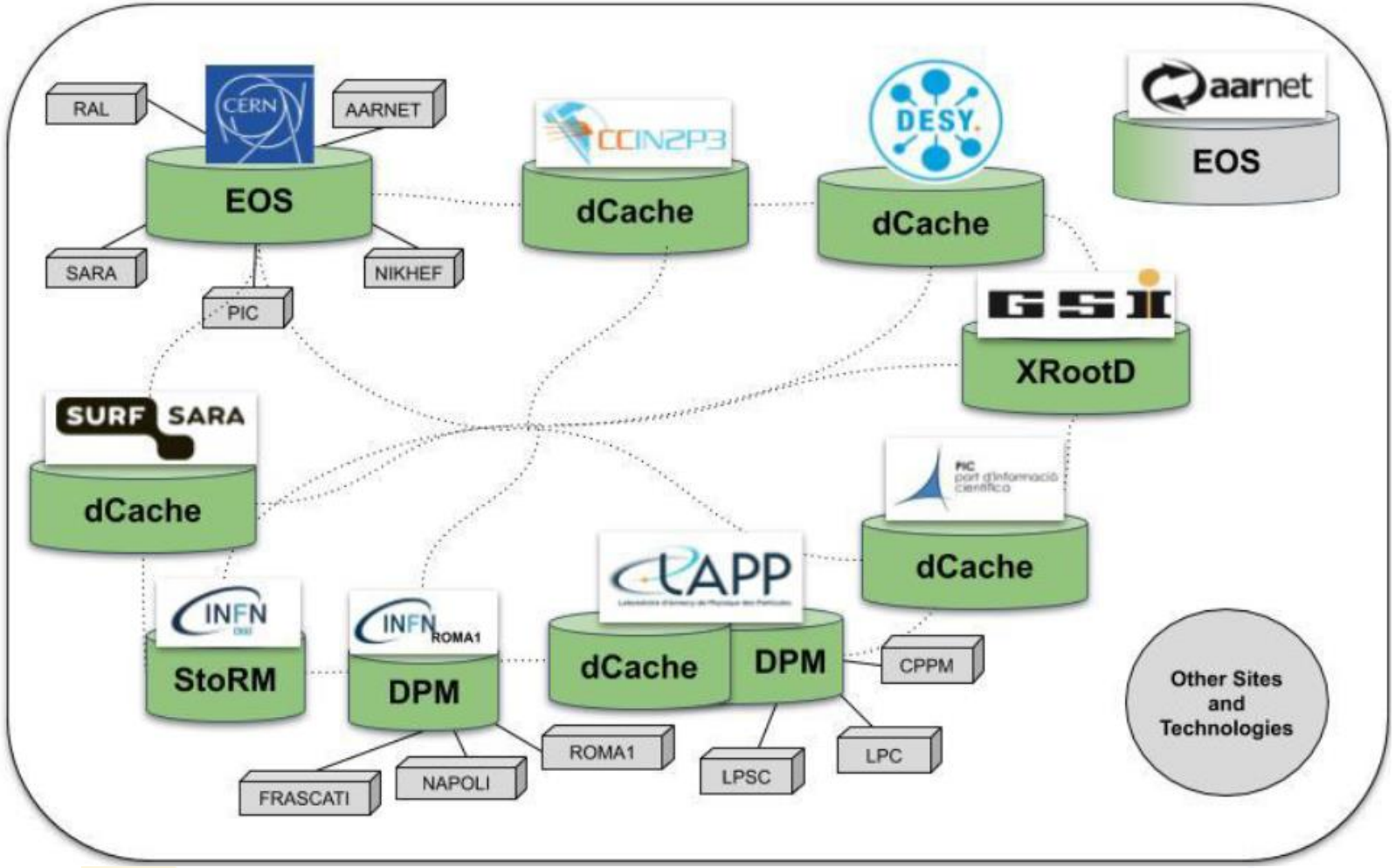
- Distributed across the world
- Using different technologies for data management
- Each one has a different regulation and policy

Funded by the European Union's Horizon 2020 - Grant N° 824064

Data Lake



Data Lake

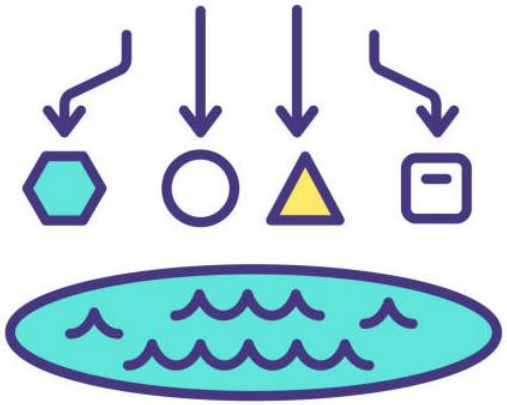


- Data Orchestration
  - Rucio customised to the needs of ESCAPE community

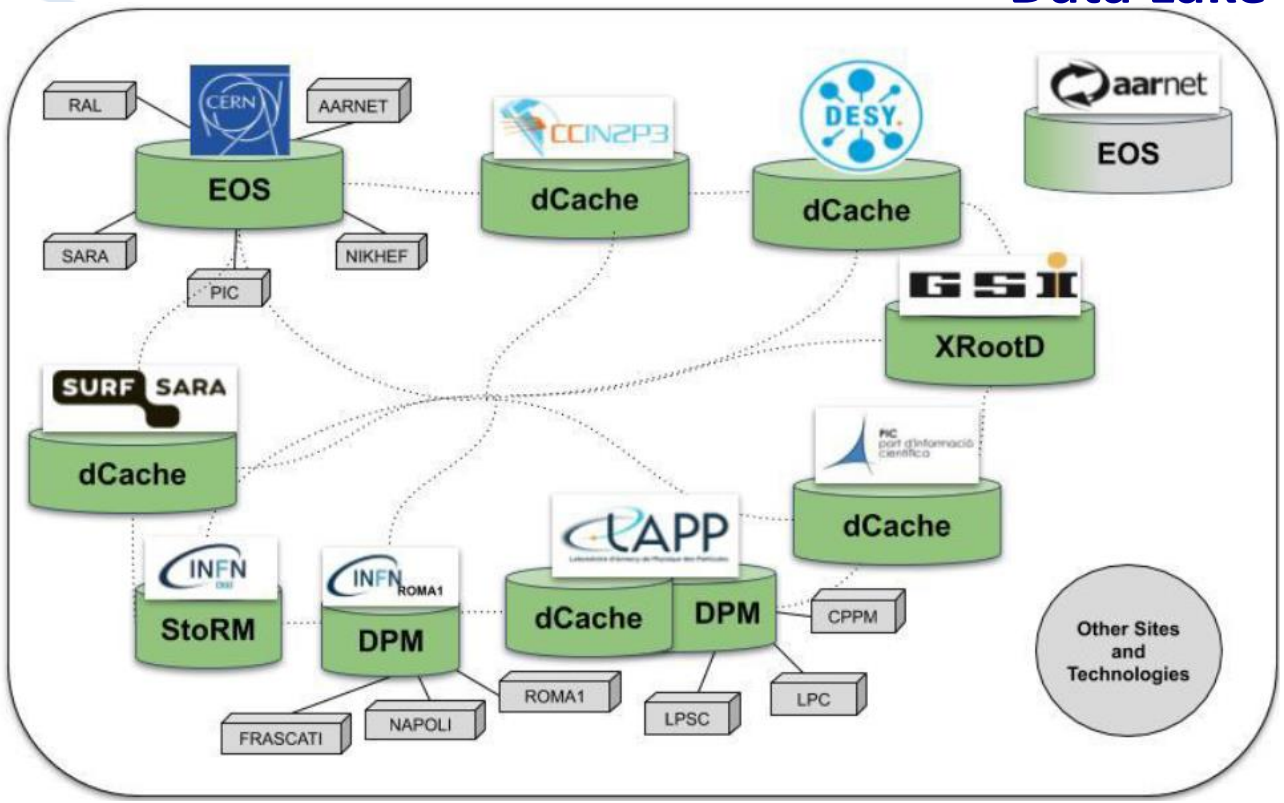
- File Transfer Service → WLCG FTS
  - continuous testing for both FTS and file access libraries

Funded by the European Union's Horizon 2020 - Grant N° 824064

Data Lake



Data Lake



- Data Orchestration
  - Rucio customised to the needs of ESCAPE community

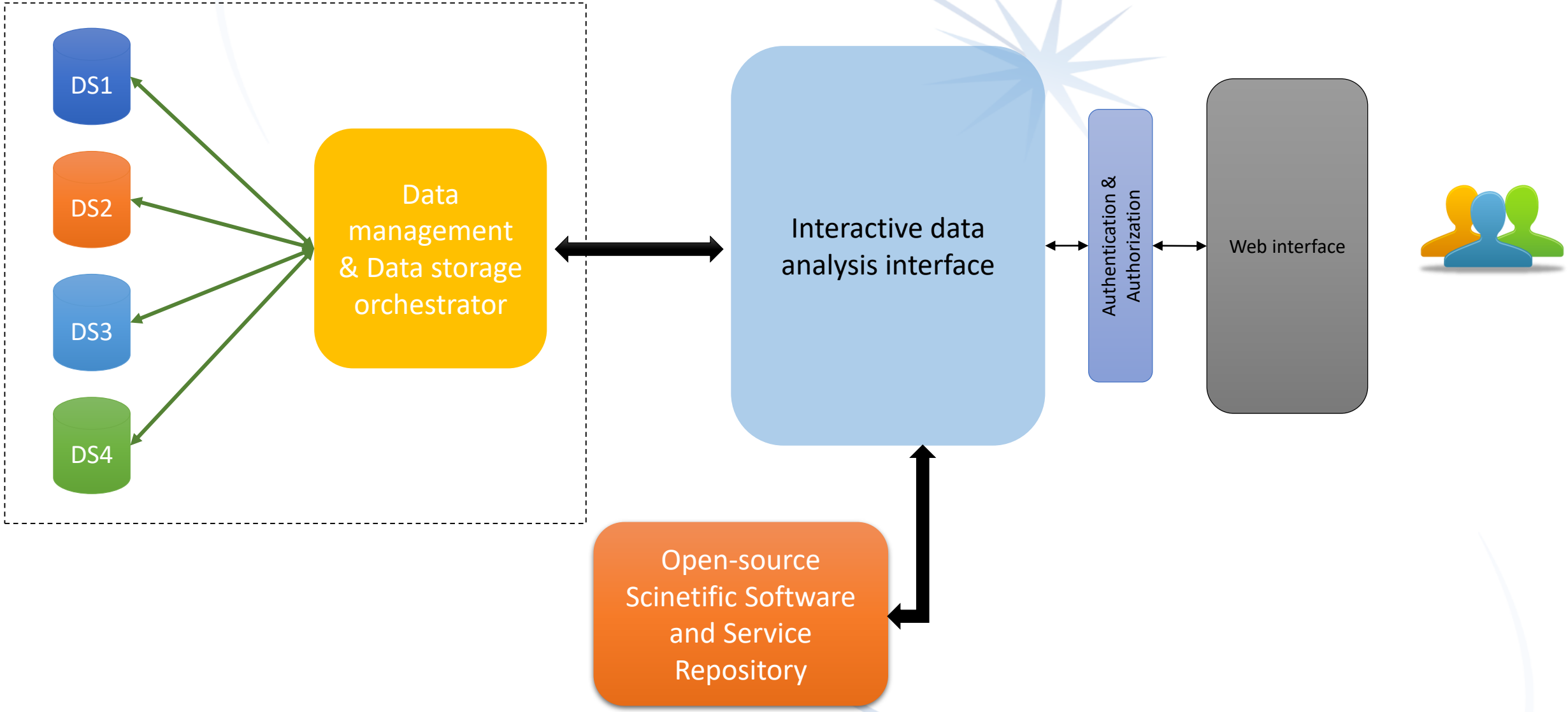
- File Transfer Service → WLCG FTS
  - continuous testing for both FTS and file access libraries

For ATLAS experiment:

- 1B+ files, 505 PB of data, 400+ Hz interaction rate
- 120 data centres, 5 HPCs, 600 storage areas
- 500 Petabytes/year transferred & deleted
- 2.5 Exabytes/year uploaded & downloaded

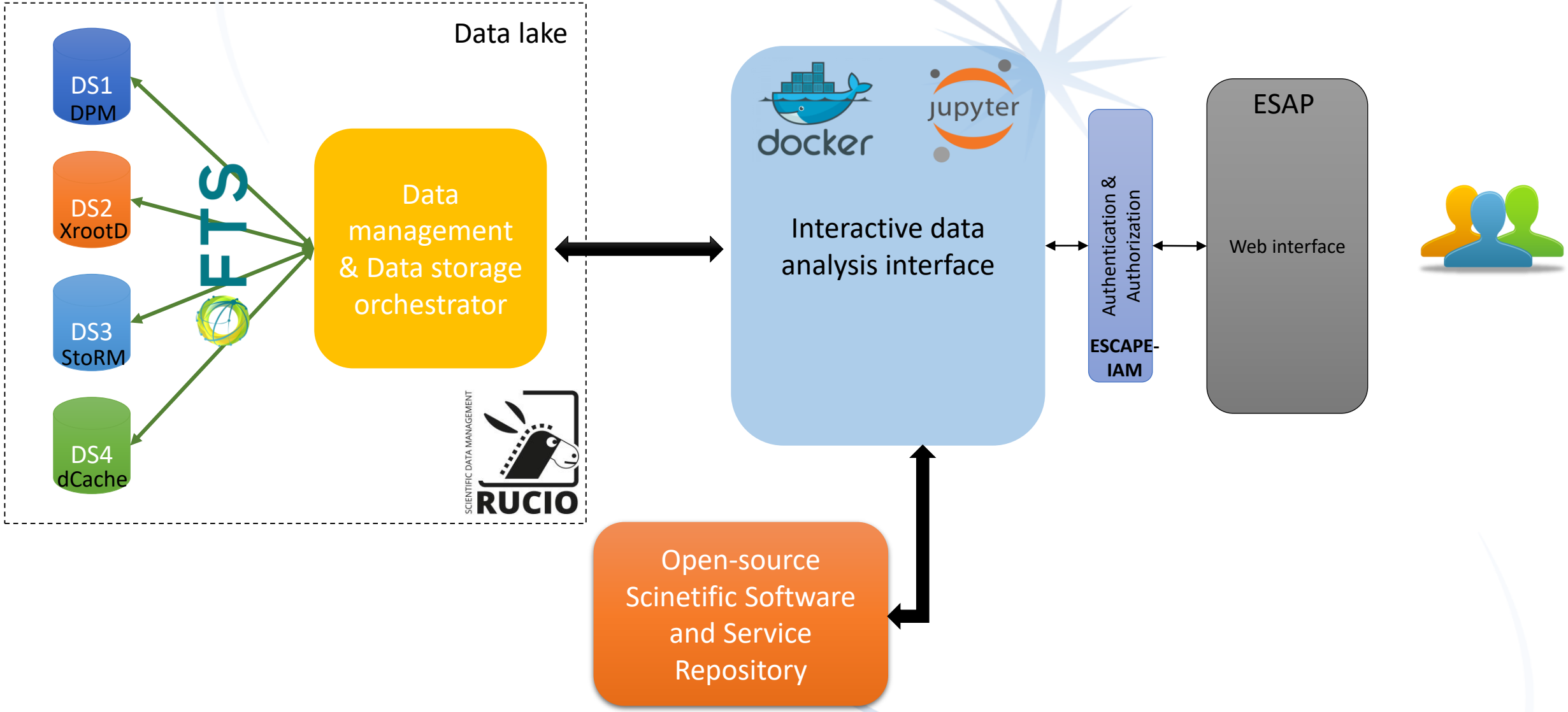
Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Final Shape



Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Final Shape



Funded by the European Union's Horizon 2020 - Grant N° 824064

Data Lake as a Service

- The Data Lake is a place where experiments can ‘dump’ the data
- … and scientists can ‘fish’ data from
- Challenge: making sure the scientists can ‘fish’ easily.

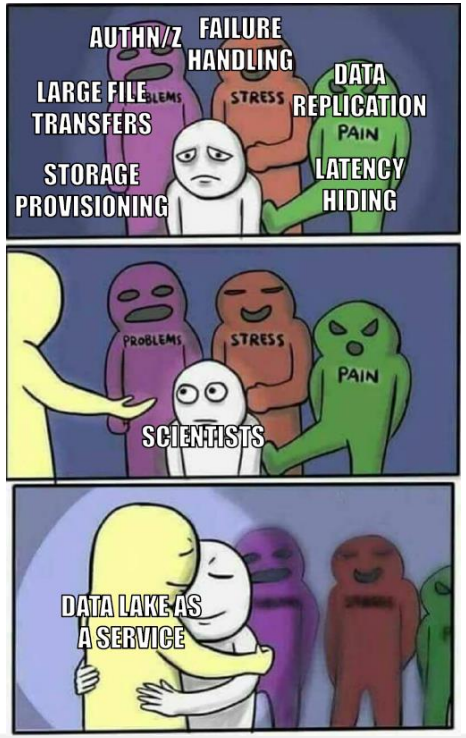
Funded by the European Union’s Horizon 2020 - Grant N° 824064

Data Lake as a Service

- The Data Lake is a place where experiments can ‘dump’ the data
- … and scientists can ‘fish’ data from
- Challenge: making sure the scientists can ‘fish’ easily.

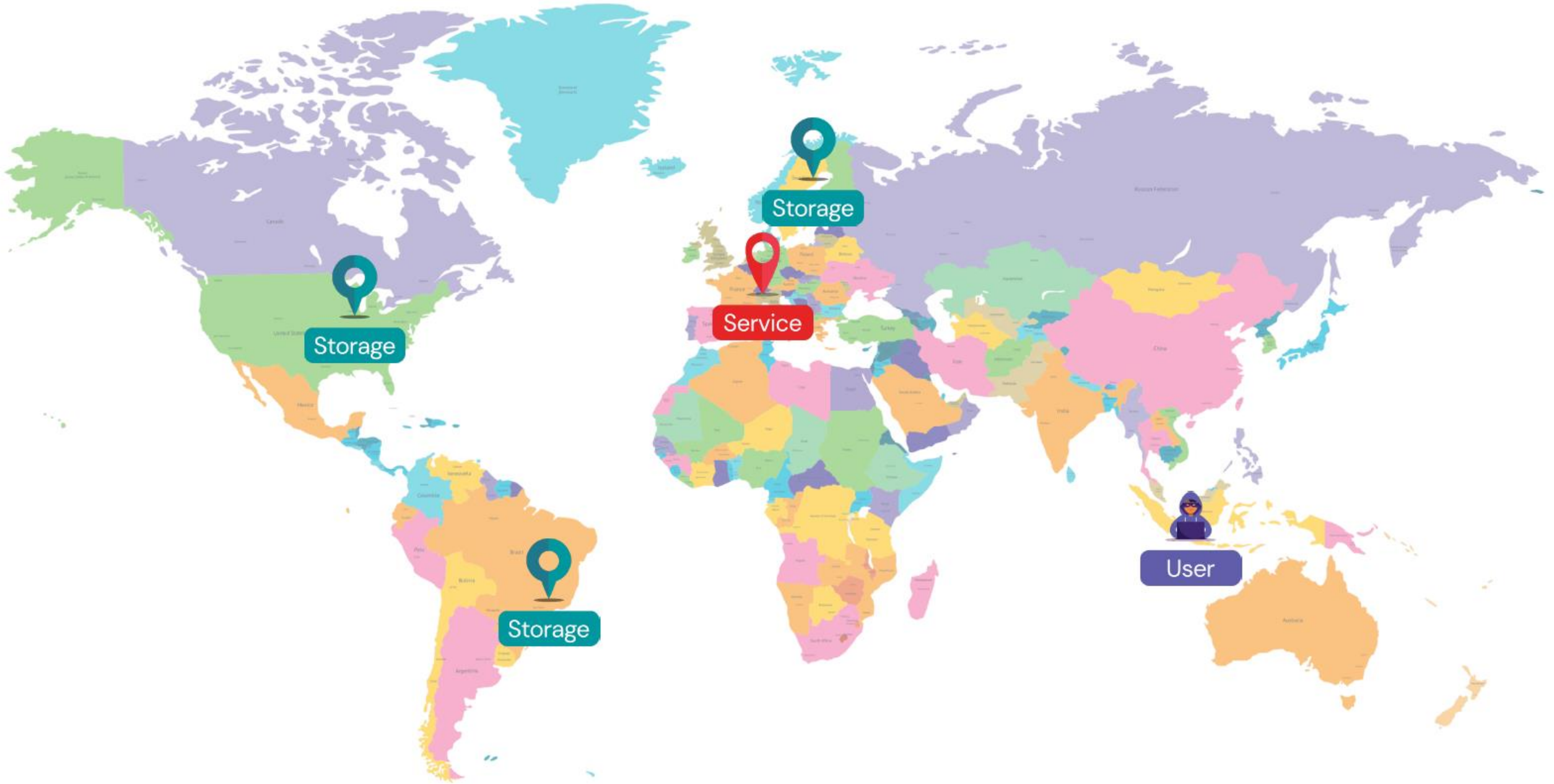
Making ‘data fishing’ easier!

Data Lake as a Service (DLaaS)

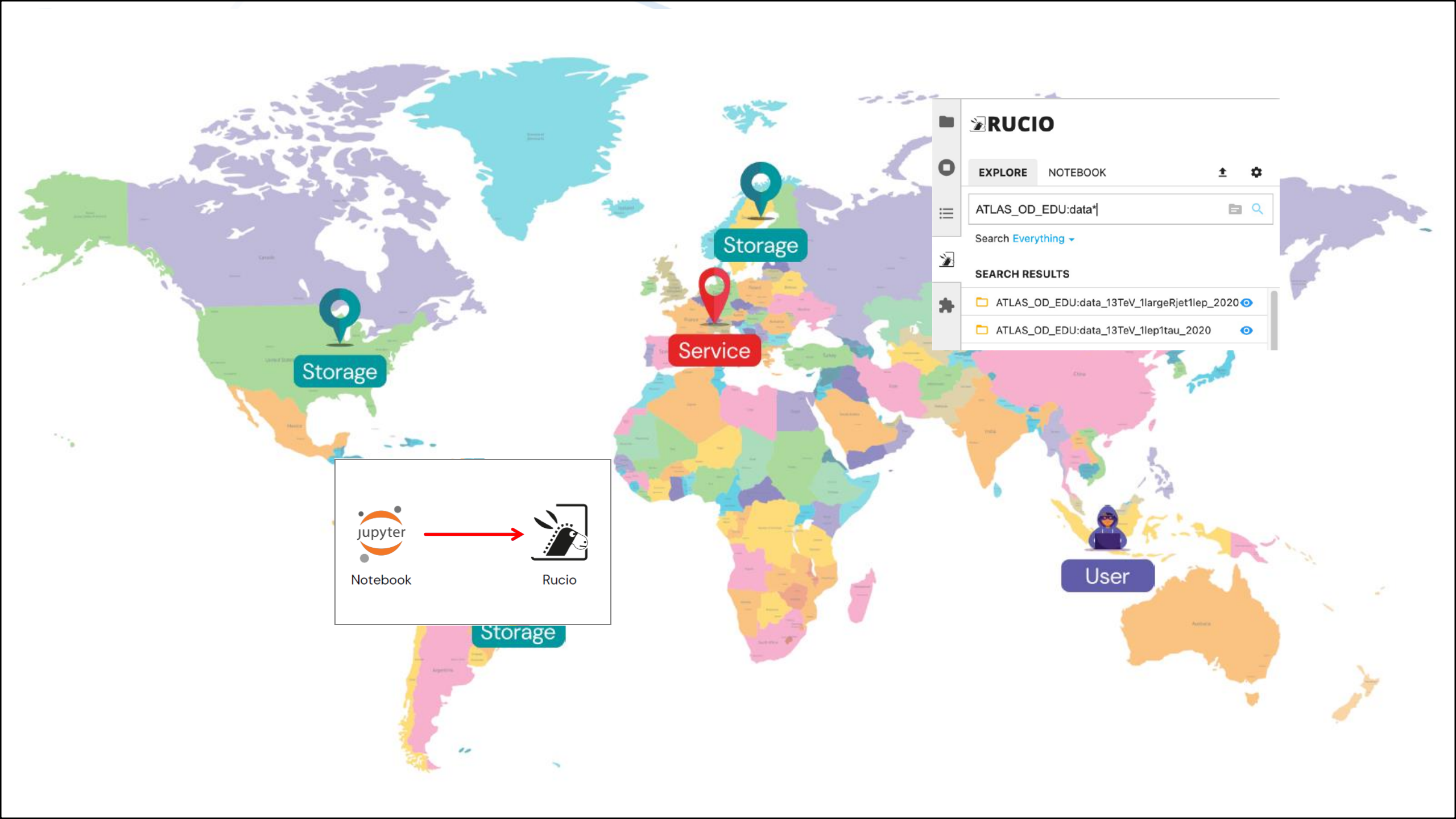


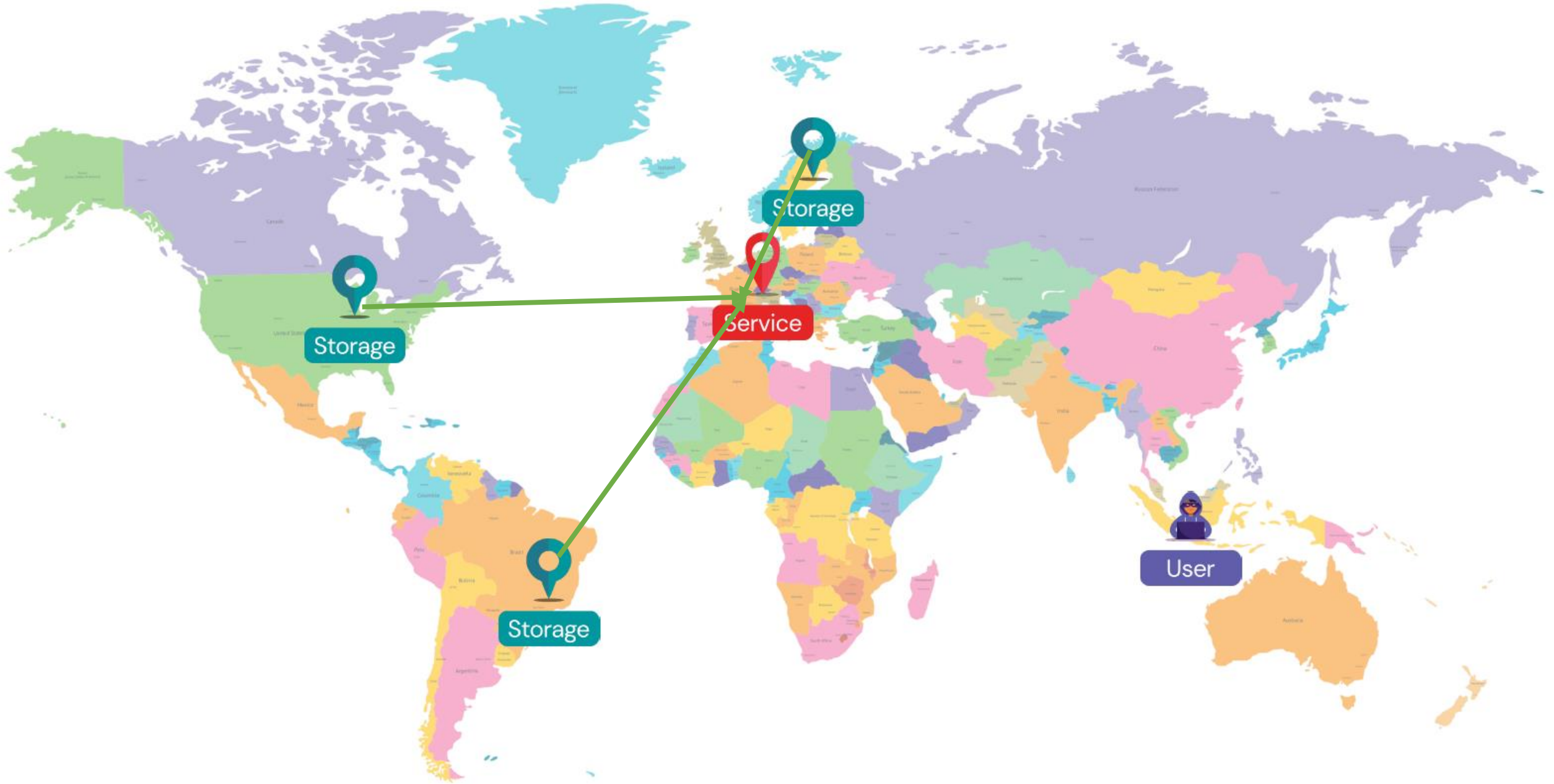
- The Data Lake has a lot of moving parts
- The goal of the service is to abstract the complexities of the Data Lake from the scientists.
- This way, scientists can focus their time on doing science instead of data procurement.

Funded by the European Union’s Horizon 2020 - Grant N° 824064


Storage
Storage
Service
Storage
Storage
User

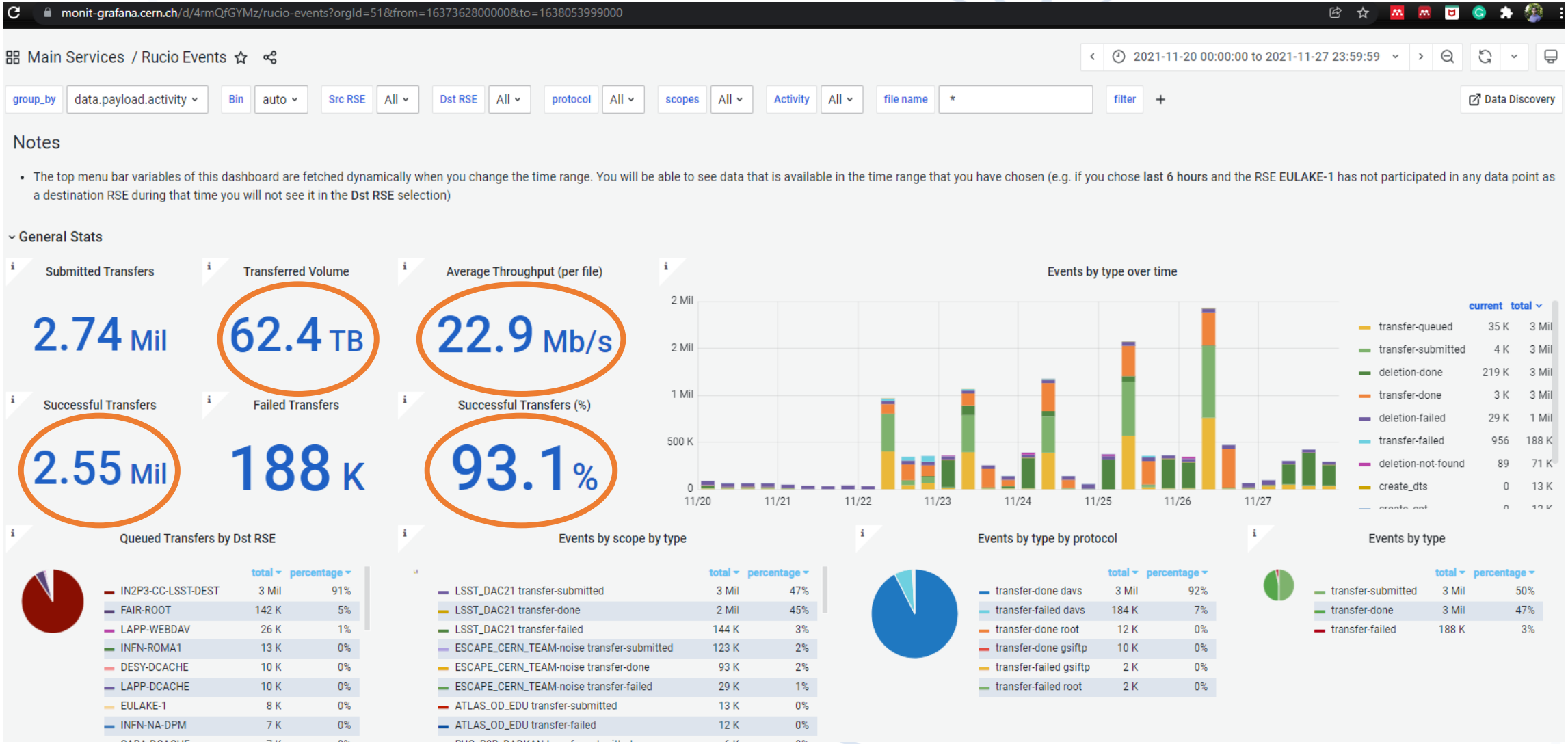

JupyterHub
https://escape-notebook.cern.ch/
jupyterhub
Storage
Storage
Service
Storage
User


RUCIO
EXPLORE
NOTEBOOK
ATLAS_OD_EDU:data*
Search Everything
SEARCH RESULTS
ATLAS_OD_EDU:data_13TeV_1largeRjet1lep_2020
ATLAS_OD_EDU:data_13TeV_1lep1tau_2020
Storage
Storage
Service
Storage
User
jupyter
Notebook
Rucio


Storage
Storage
Service
Storage
Storage
User

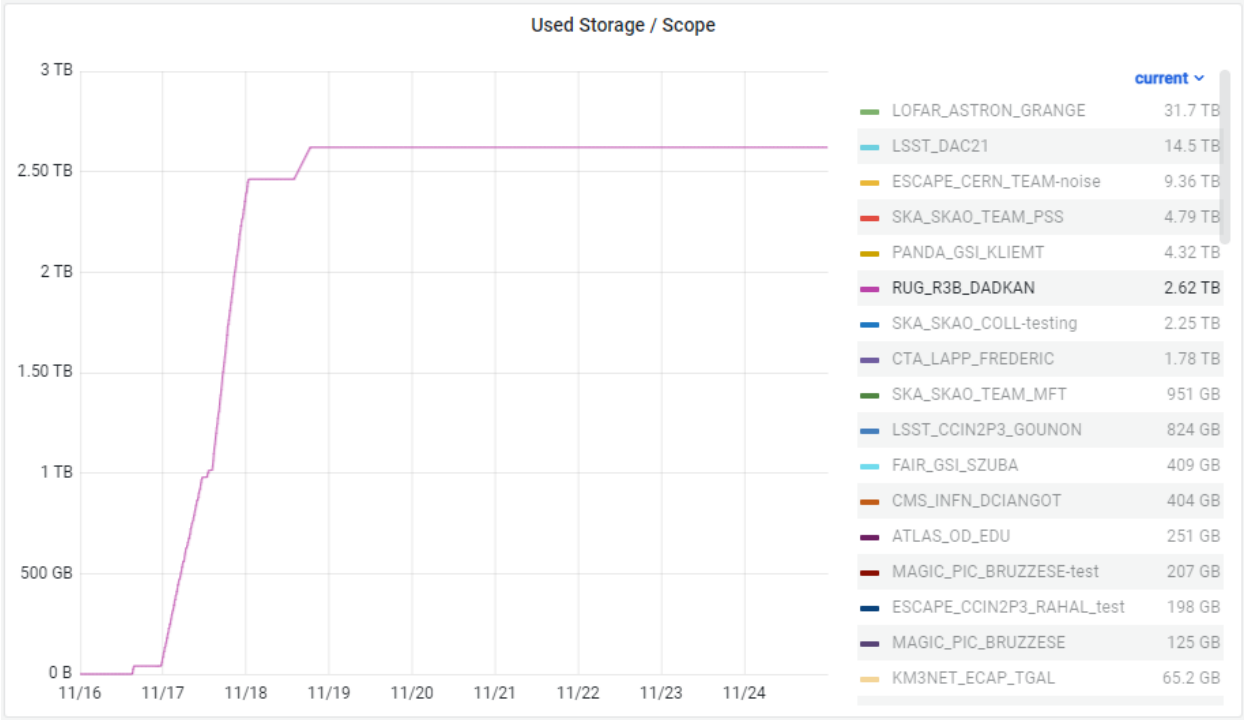

Storage
Storage
Service
Storage
Storage
User

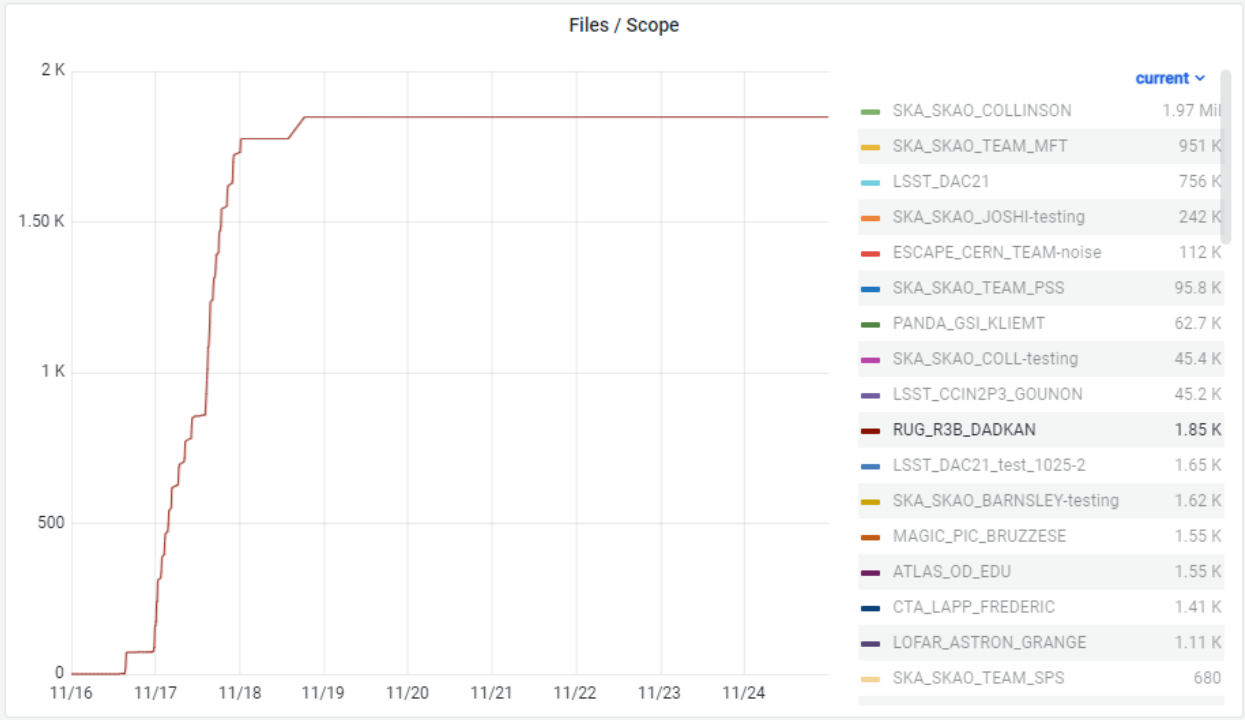
Data Analysis Challenge 2021 (DAC21)



Funded by the European Union's Horizon 2020 - Grant N° 824064

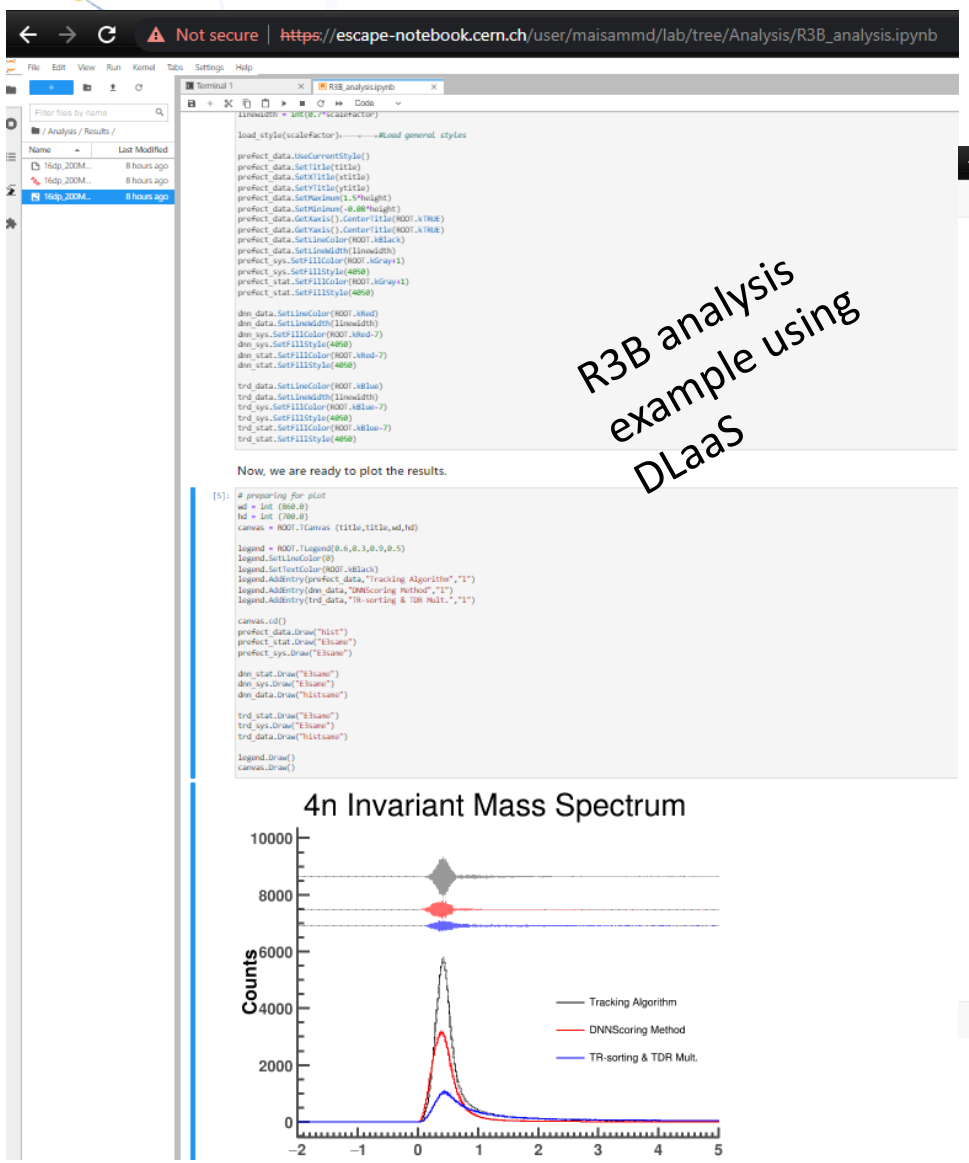
Data Analysis Challenge 2021 (DAC21)

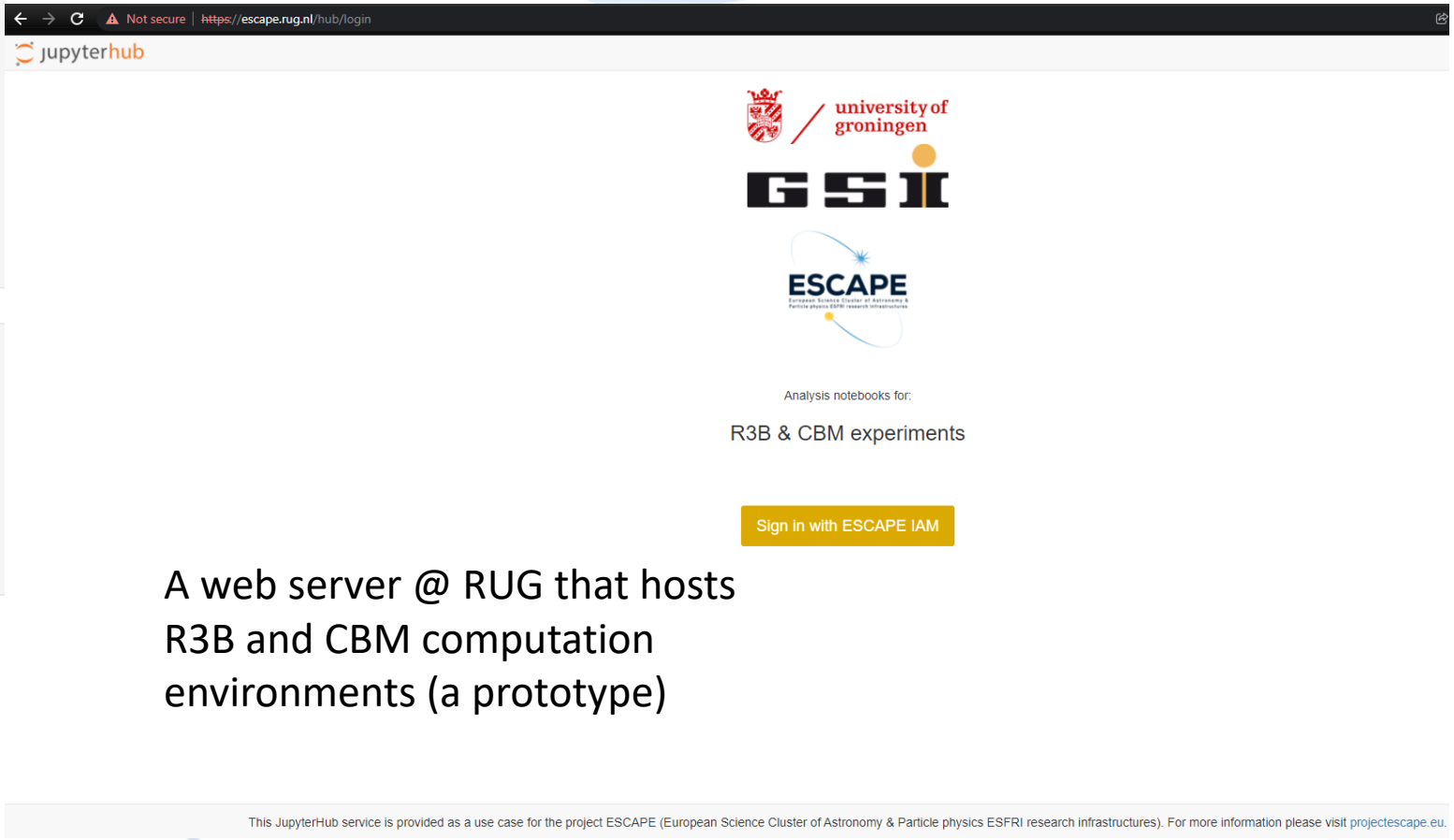




Funded by the European Union's Horizon 2020 - Grant N° 824064

Summary





A web server @ RUG that hosts R3B and CBM computation environments (a prototype)

Funded by the European Union's Horizon 2020 - Grant N° 824064

Summary

- Issues with the portability & reusability of R3BRoot software stack were resolved using containers
- R3BRoot was Inetegrated with an Interactive analysis interface (Jupyter Lab)
- Deployed as a web application to ensure findability & accessibility of the software and the data
- The ESCAPE Data lake proved that it can be a solution to tackle big data challenges
- Still a lot of effort needs to be made to make data and software FAIR
- Making a sutainable and federated ecosystem for research data under EOSC is still a big challenge but moving forward.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Thank you!
&
Questions!

'Research is born free and everywhere is in chains'

(apologies to Rousseau)

Funded by the European Union's Horizon 2020 - Grant N° 824064

Future of ESCAPE & EOSC

- **ESCAPE intends to remain as a collaboration after the end of the current project**

→ Sustained thematic consortium for the sustainability of EOSC

→ Additional thematic RI's are interested in participating

- **Many potential collaborative actions**

→ Data Management at Exascale

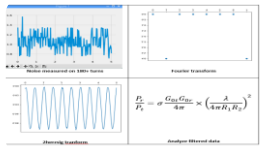
→ Common software challenges

→ Cross-fertilisation of software tools and services

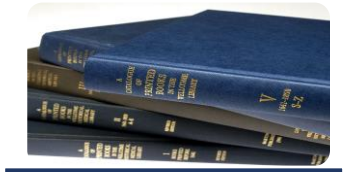
→ Building a platform for multi-messenger astronomy and cross-domain collaborations

→ Connection of data infrastructure to heterogeneous computing

Services for integration of the Exabyte scale data infrastructure supporting Open Science and FAIR data



Interoperable data-analysis platform



Development and sustainability of the services for the software catalogue and workflow

Partnership with EuroHPC/Prace/FENIX to provide the services for integration of Data-Lake with HPC resources

Collaborative scientific environment. Include new ESFRI and SME

Global integration of some of the above with global partners in USA, Australia and Africa

Funded by the European Union's Horizon 2020 - Grant N° 824064

Future of ESCAPE & EOSC

EOSC-Future

Five thematic Science Clusters (80% of ESFRI RIs)

EOSC FUTURE

INFRAEOSC-07 Projects

EUROPEAN OPEN SCIENCE CLOUD

EuroHPC Joint Undertaking

PRACE PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE

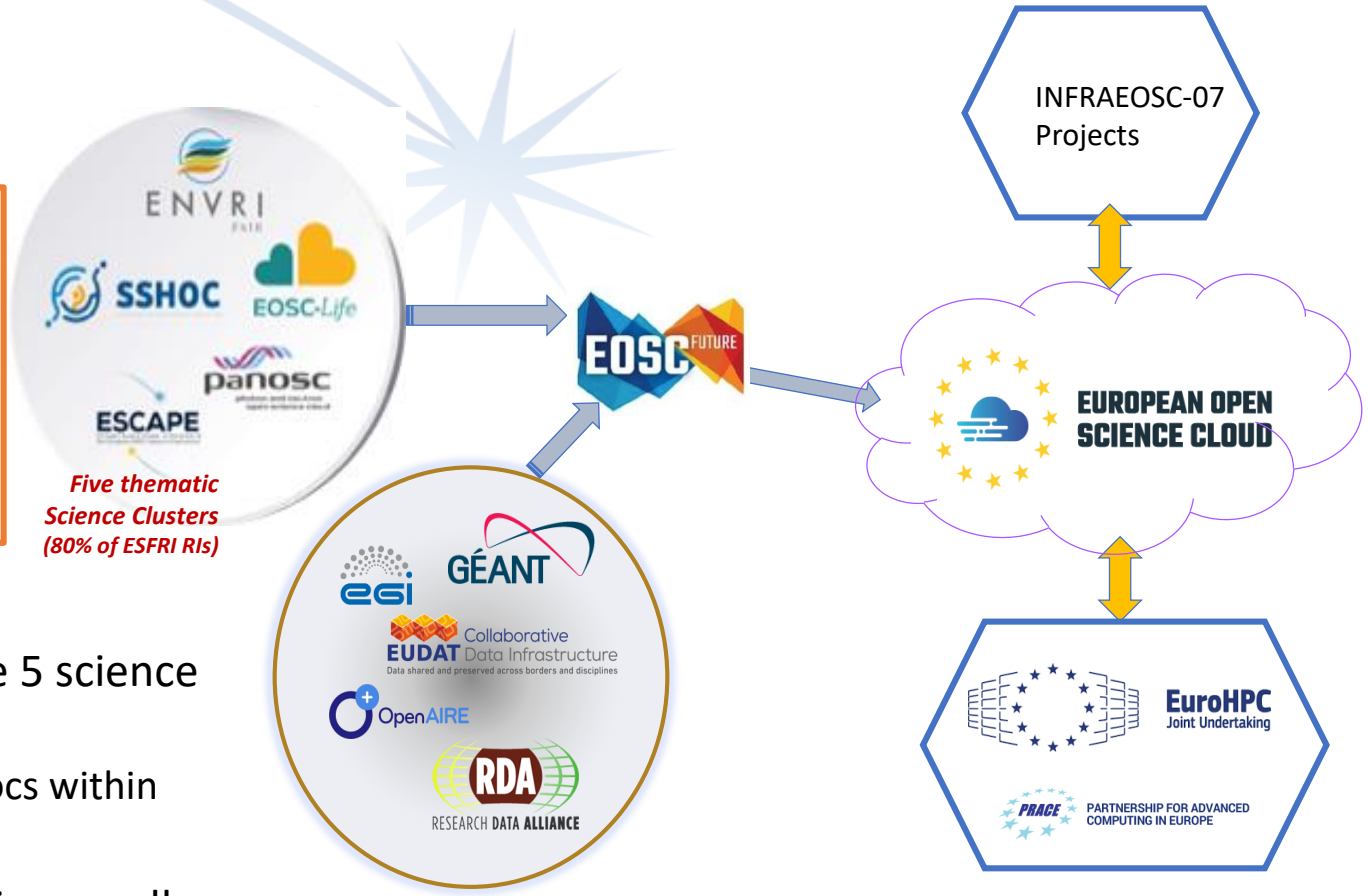
Funded by the European Union's Horizon 2020 - Grant N° 824064

Future of ESCAPE & EOSC

EOSC-Future

- Project – started 1st April;
- Responding to EU H2020 funding call, (INFRAEOSC-03-2020): 30 months, 40 M euros

➢ ***EOSC-Future is a prototype of an integrated EOSC***

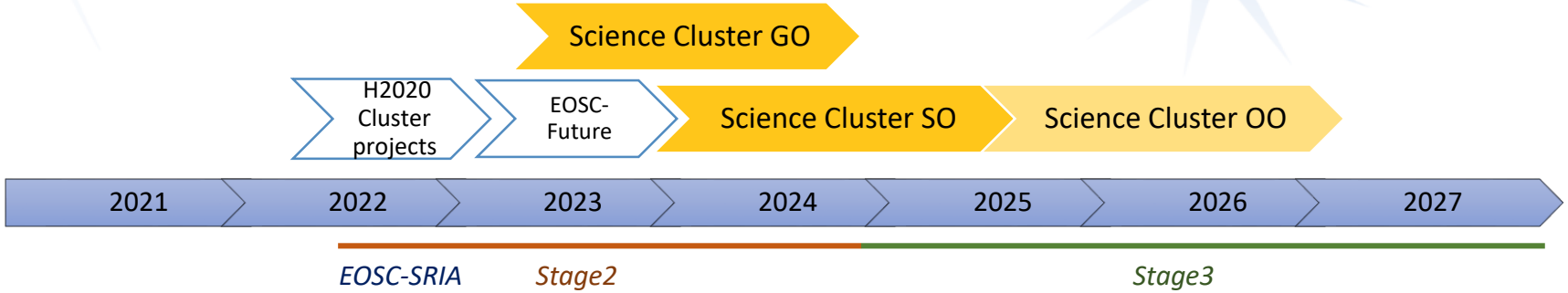


- WP6 (T6.3) includes 2 science projects from each of the 5 science clusters
  - Several partners from ESCAPE have funding for 6 post-docs within that to work on the TSP's
- ESCAPE (with other clusters & e-infras) is also involved in overall Technical Coordination of EOSC-Future
  - Specific actions around portal integration, work with HPC community, etc

Funded by the European Union's Horizon 2020 - Grant N° 824064

A work plan for the future of the Science Clusters

There is a prompt need and opportunity to support the Science Clusters further (in 2022-2025) within **Horizon Europe framework**



H2020 and potential **Horizon Europe** funded actions (*aligned with the EOSC-SRIA* Stage2 -to- Stage3*)

General Objectives (**GO**):
- consolidation of thematic data infrastructures (cluster VREs, platforms and a "few core services") as parts of a federation.

Specific Objectives (**SO**):
- relevant scientific results from clusters;
- increased number of RIs;
- enhance researchers uptake of OS and widening dimension.

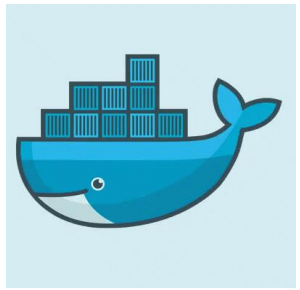
Operational Objectives (**OO**):
- sustainable operation of the deployed cluster as a "platform infrastructure";
- continuous promotion, extension and hosting of inter-domain FAIR Science Projects (new Open Science Objectives).

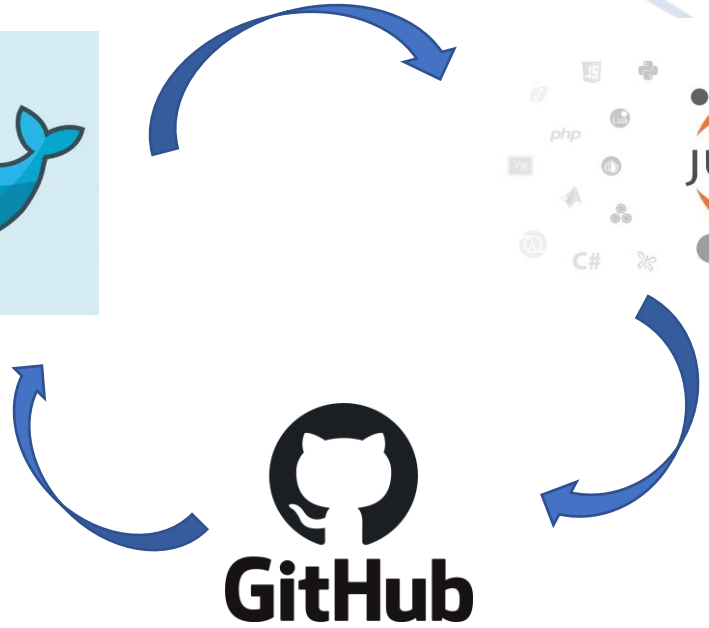
*Strategic Research and Innovation Agenda

Funded by the European Union's Horizon 2020 - Grant N° 824064

Recap

A standard OS-virtualization platform for **packing up software** and all the **dependencies** to **run applications on different environments**.



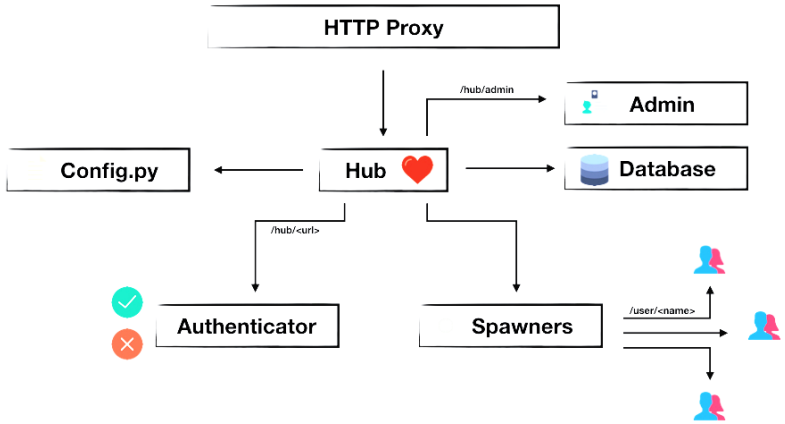


A version control system for software with a free online repository

A **Web IDE** for **live code**, equations, **visualizations** and **narrative text (Jupyter Notebook)**

Brings the power of **notebooks** to **groups of users** & handles users access to **computational environments** and ... **(Jupyter Hub)**

JupyterHub



- Reusability
- Accessibility
- Interoperability
- Portability
- Micro services & Scaling

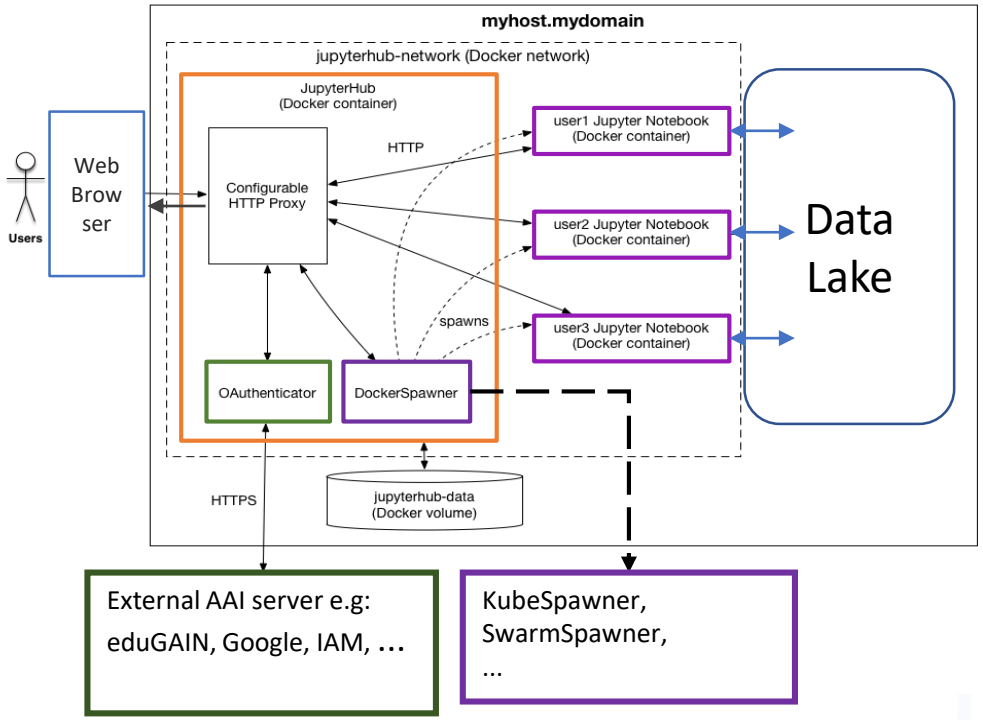
Funded by the European Union's Horizon 2020 - Grant N° 824064

Demo!

Funded by the European Union's Horizon 2020 - Grant N° 824064

The prototype of the R3B interactive data analysis (IDA)

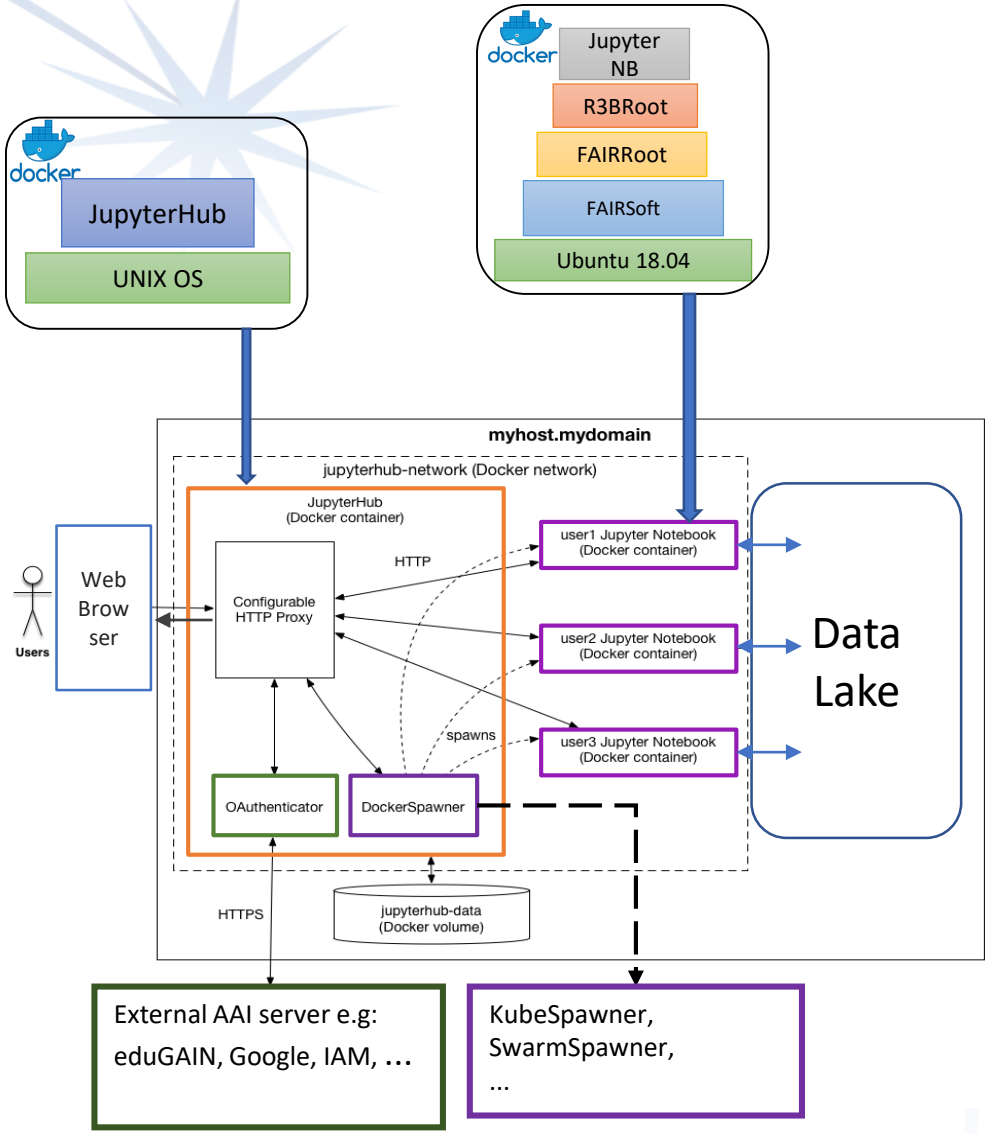
- Tested the alpha version on the cloud
- Integrated with an OIDC AuthN/Z (Github)
- Tested connecticity and R/W to the data lake (X509)



Funded by the European Union's Horizon 2020 - Grant N° 824064

The prototype of the R3B interactive data analysis (IDA)

- Tested the alpha version on the cloud
- Integrated with an OIDC AuthN/Z (Github)
- Tested connecticity and R/W to the data lake (X509)



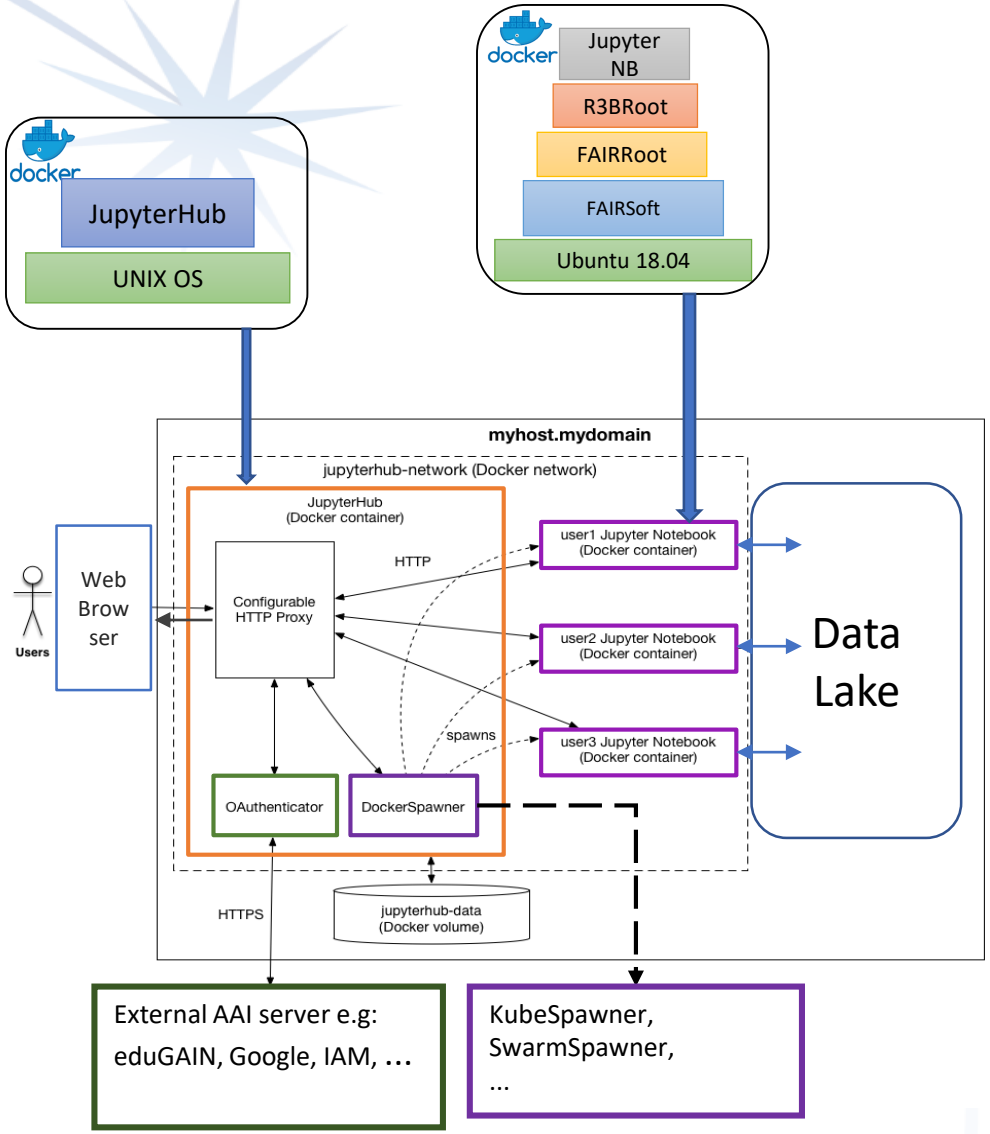
Funded by the European Union's Horizon 2020 - Grant N° 824064

The prototype of the R3B interactive data analysis (IDA)

- Tested the alpha version on the cloud
- Integrated with an OIDC AuthN/Z (Github)
- Tested connecticity and R/W to the data lake (X509)

Things that needed to be improved:

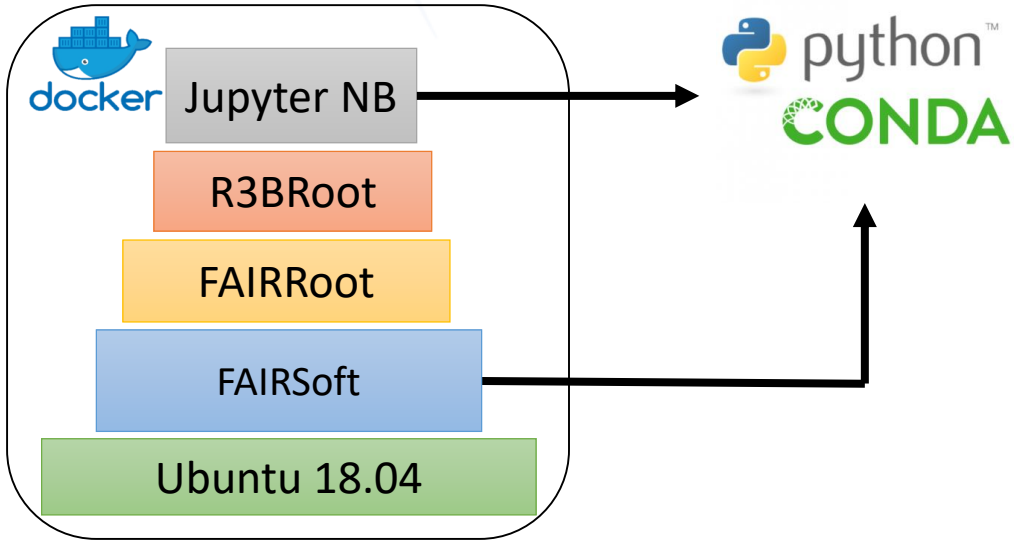
- Not in standard Python environment (CONDA)
- Not ready to be in production (scalability)
- Not integrated with ESCAPE AuthN/Z (IAM)
- Manual settings to access to the Data Lake (DL)



Funded by the European Union's Horizon 2020 - Grant N° 824064

Latest progress

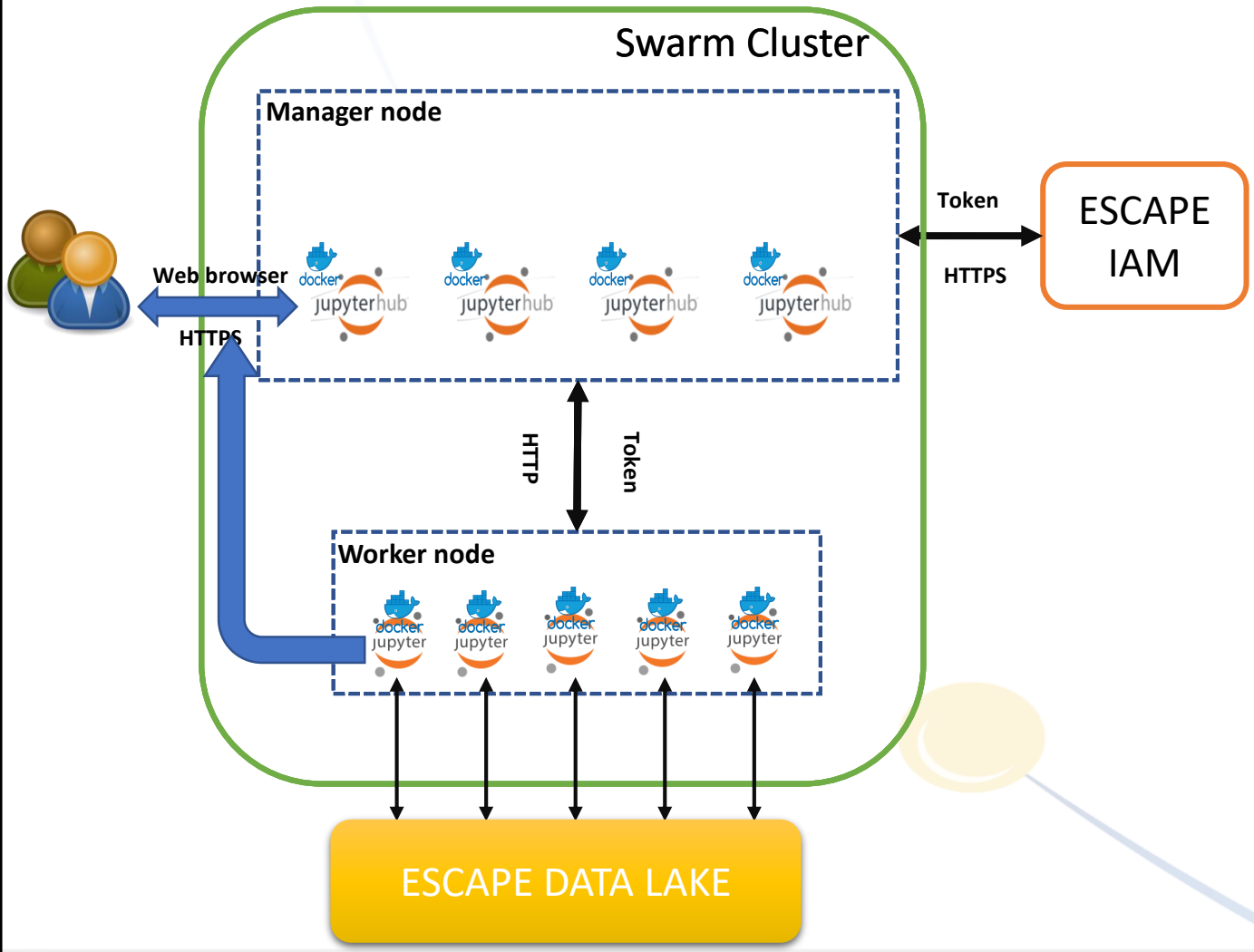
✓ **Standard Python environment**



- New FAIRSoft (some bug are fixed)
- PyRoot and Jupyter NB use the same python version (ensure compatibility)
- Docker images are made for each part (can be used in other FAIR projects)
- Some security issues were resolved (root access limited)

Funded by the European Union's Horizon 2020 - Grant N° 824064

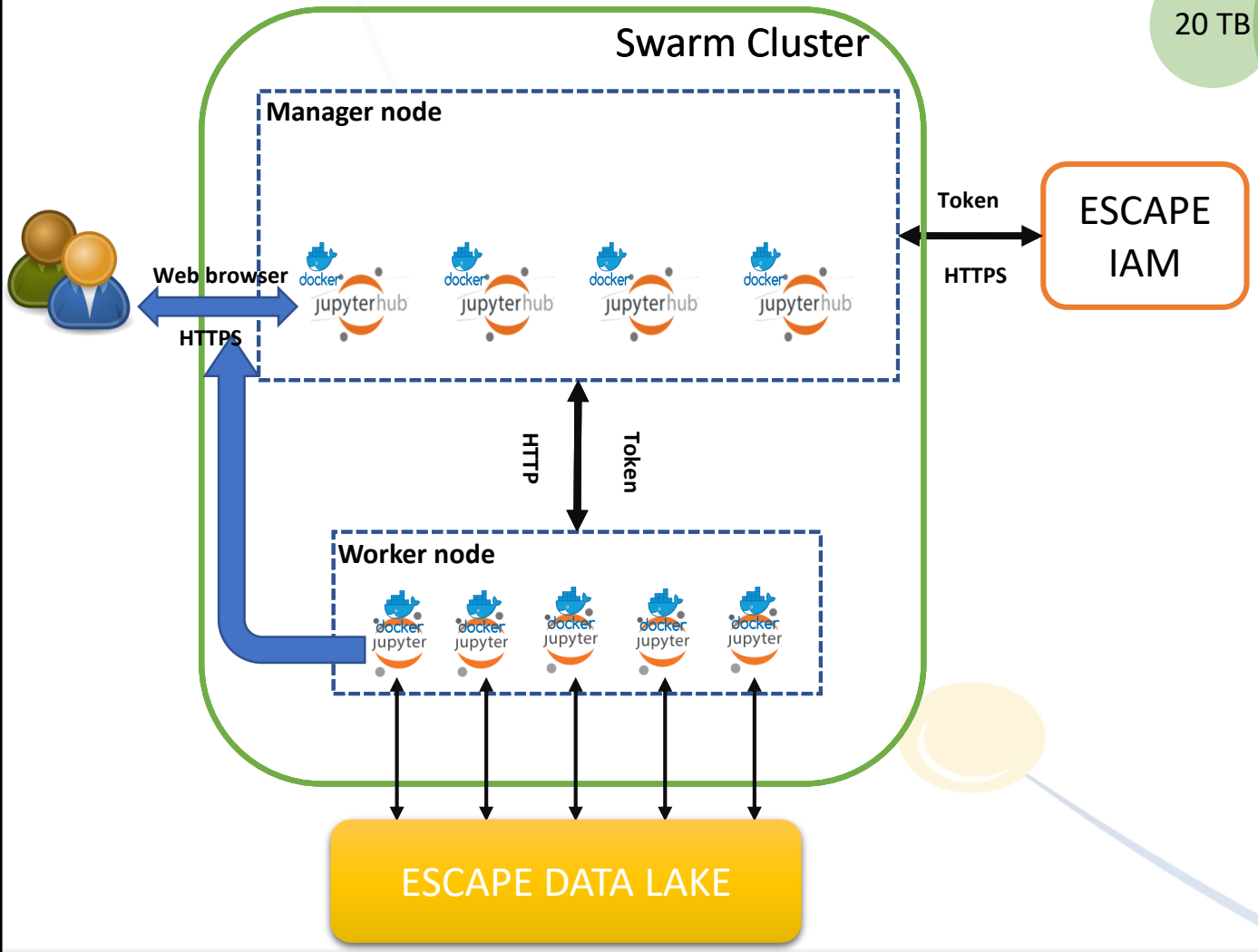
✓ **Ready to be deployed in production (scalability)**

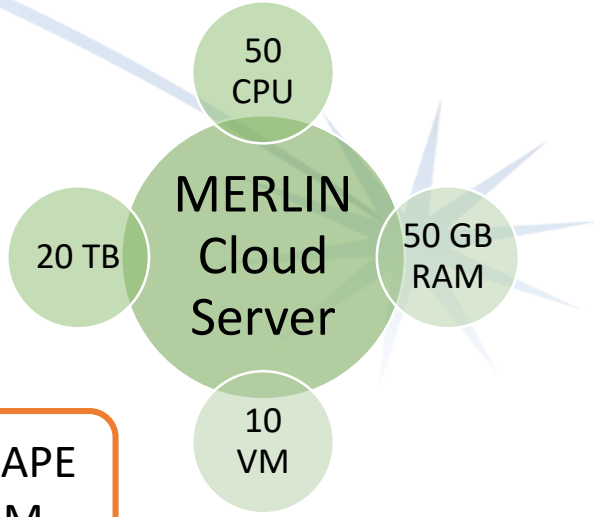


Funded by the European Union's Horizon 2020 - Grant N° 824064

Latest progress

✓ **Ready to be deployed in production (scalability)**





- The JupyterHub works in a different node (higher performance)
- It is scalable
- More security settings were added to the JupyterHub

https://escape.rug.nl

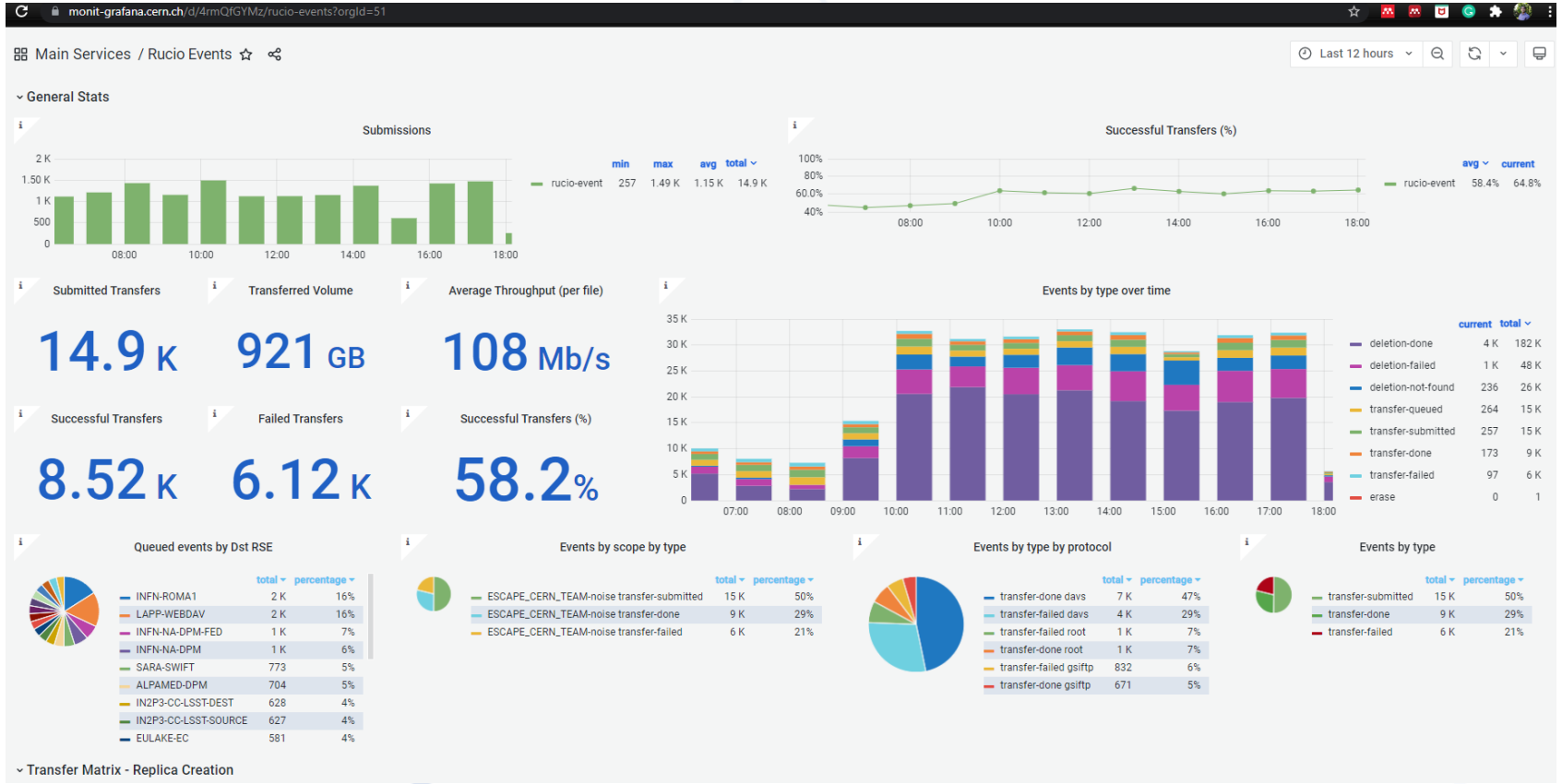
- A token propagation mechanism needs to be added
- Local storage per user needs to be assigned
- End-to-end AuthN/Z needs to be enabled for the data lake

Funded by the European Union's Horizon 2020 - Grant N° 824064

Monitoring of the Data Lake

- Track all the task related to DL
- Show all the informations of data transfering rates
- Health check of RSEs
- ....

https://monit-grafana.cern.ch/



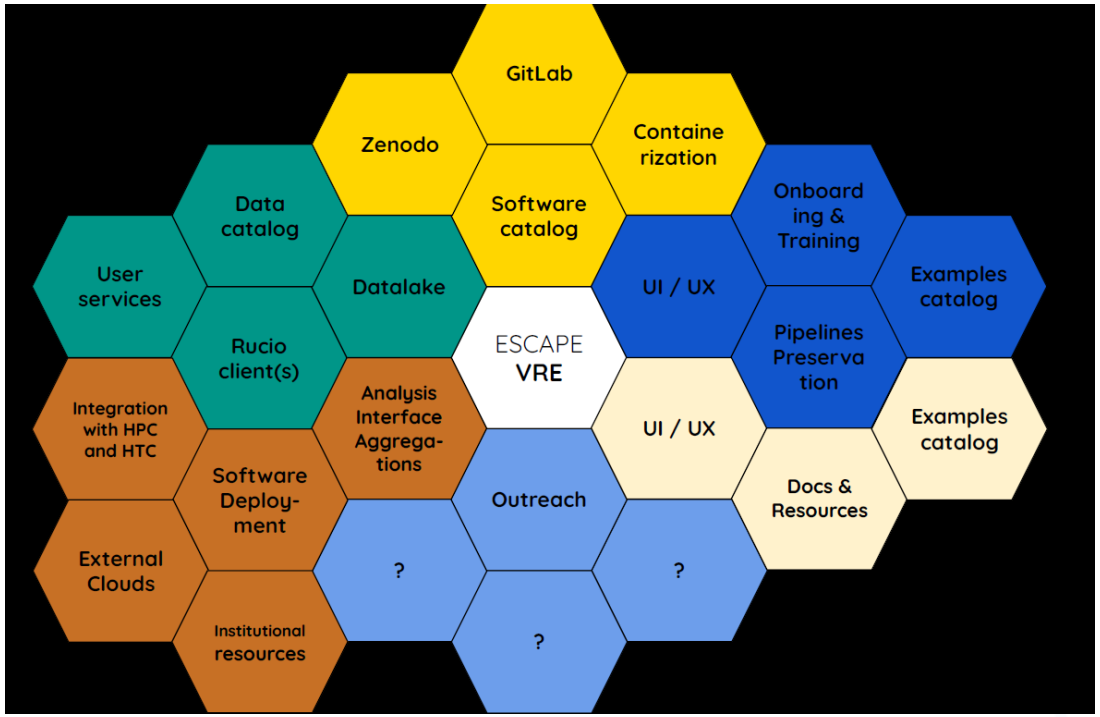
Funded by the European Union's Horizon 2020 - Grant N° 824064

Virtual Research Enviroment (VRE)

https://escape2020.pages.in2p3.fr/virtual-environment/home/

An exercise to imagine how current resources can be aggregated so as to be reachable and use from a single entry point

- Simple User Onboarding
- Good UI/UX as an essential component
- Scalability
- CI/CD by default
- Portability



Funded by the European Union's Horizon 2020 - Grant N° 824064

DAC21: Data and Analysis Challenge (November 2021):

The aim is to finish the prototype evaluation by the end of the year.

Goals:

- Run production Data Management, Processing and Analysis workloads
  - Data Management: acquisition, injection, replication, lifecycles
  - Data Processing: Production
  - Data processing: User Analysis. Use cases
- Demonstrate Data Lake orchestration layer sustainability after ESCAPE (towards EOSC)
  - Leverage, integrate and use experiment and site's dedicated installations, e.g. RUCIO and FTS.
  - Demonstrate integration and interplay of these several instances in a common Data Lake/storage infrastructure.
- End-to-End AAI
  - Assessment ranging from experiments experts, to advanced users to newcomers and to sporadic web-based access.
  - Token-based Authentication.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Future of ESCAPE & EOSC

- **ESCAPE intends to remain as a collaboration after the end of the current project**

→ Sustained thematic consortium for the sustainability of EOSC

→ Additional thematic RI's are interested in participating

- **Many potential collaborative actions**

→ Data Management at Exascale

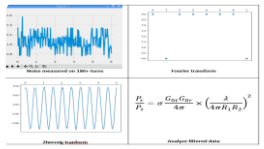
→ Common software challenges

→ Cross-fertilisation of software tools and services

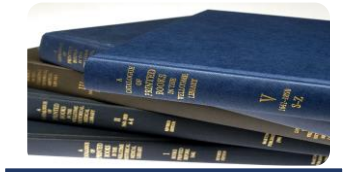
→ Building a platform for multi-messenger astronomy and cross-domain collaborations

→ Connection of data infrastructure to heterogeneous computing

Services for integration of the Exabyte scale data infrastructure supporting Open Science and FAIR data



Interoperable data-analysis platform



Development and sustainability of the services for the software catalogue and workflow

Partnership with EuroHPC/Prace/FENIX to provide the services for integration of Data-Lake with HPC resources

Collaborative scientific environment. Include new ESFRI and SME

Global integration of some of the above with global partners in USA, Australia and Africa

Funded by the European Union's Horizon 2020 - Grant N° 824064

Future of ESCAPE & EOSC

EOSC-Future

Five thematic Science Clusters (80% of ESFRI RIs)

EOSC FUTURE

INFRAEOSC-07 Projects

EUROPEAN OPEN SCIENCE CLOUD

EuroHPC Joint Undertaking

PRACE PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE

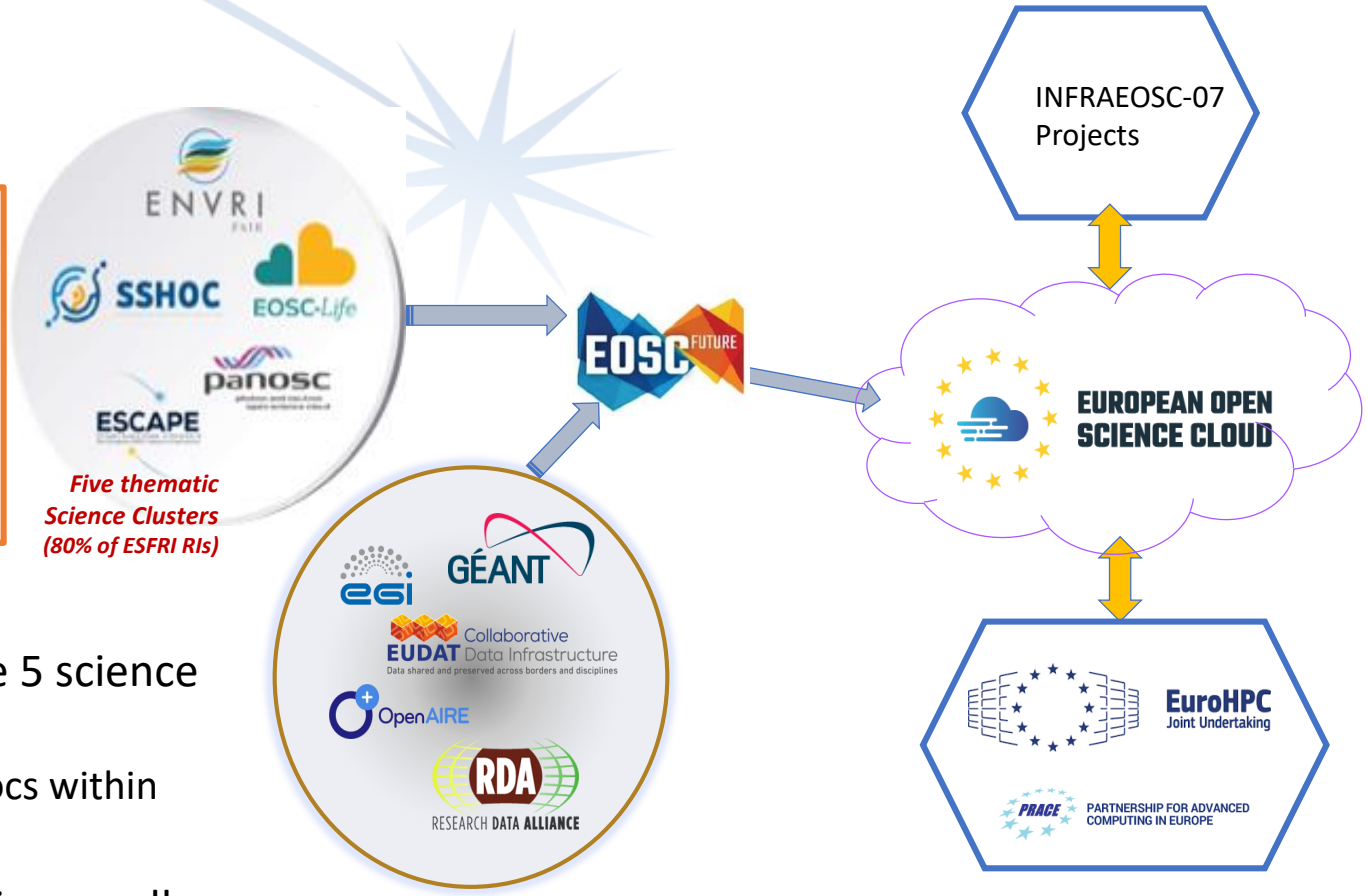
Funded by the European Union's Horizon 2020 - Grant N° 824064

Future of ESCAPE & EOSC

EOSC-Future

- Project – started 1st April;
- Responding to EU H2020 funding call, (INFRAEOSC-03-2020): 30 months, 40 M euros

➢ ***EOSC-Future is a prototype of an integrated EOSC***

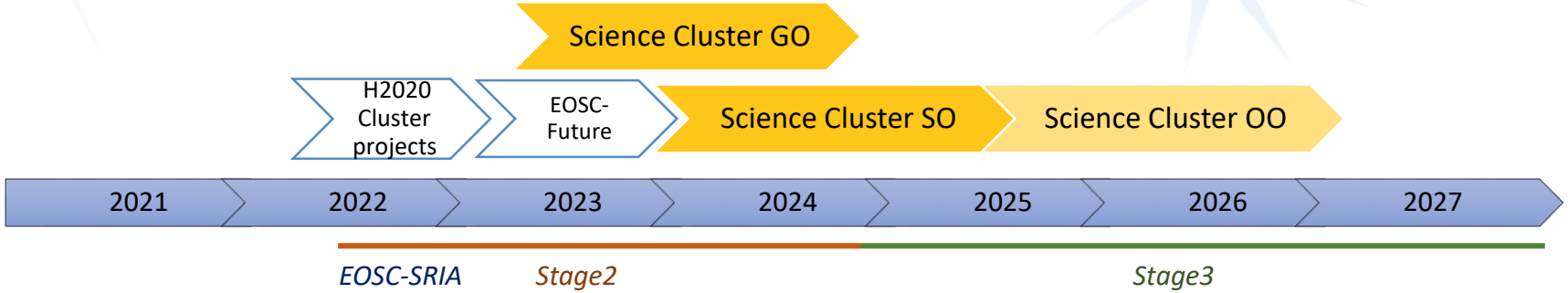


- WP6 (T6.3) includes 2 science projects from each of the 5 science clusters
  - Several partners from ESCAPE have funding for 6 post-docs within that to work on the TSP's
- ESCAPE (with other clusters & e-infras) is also involved in overall Technical Coordination of EOSC-Future
  - Specific actions around portal integration, work with HPC community, etc

Funded by the European Union's Horizon 2020 - Grant N° 824064

A work plan for the future of the Science Clusters

There is a prompt need and opportunity to support the Science Clusters further (in 2022-2025) within **Horizon Europe framework**



H2020 and potential **Horizon Europe** funded actions (*aligned with the EOSC-SRIA* Stage2 -to- Stage3*)

General Objectives (**GO**):
- consolidation of thematic data infrastructures (cluster VREs, platforms and a "few core services") as parts of a federation.

Specific Objectives (**SO**):
- relevant scientific results from clusters;
- increased number of RIs;
- enhance researchers uptake of OS and widening dimension.

Operational Objectives (**OO**):
- sustainable operation of the deployed cluster as a "platform infrastructure";
- continuous promotion, extension and hosting of inter-domain FAIR Science Projects (new Open Science Objectives).

*Strategic Research and Innovation Agenda

Funded by the European Union's Horizon 2020 - Grant N° 824064

Next steps

Priority 1

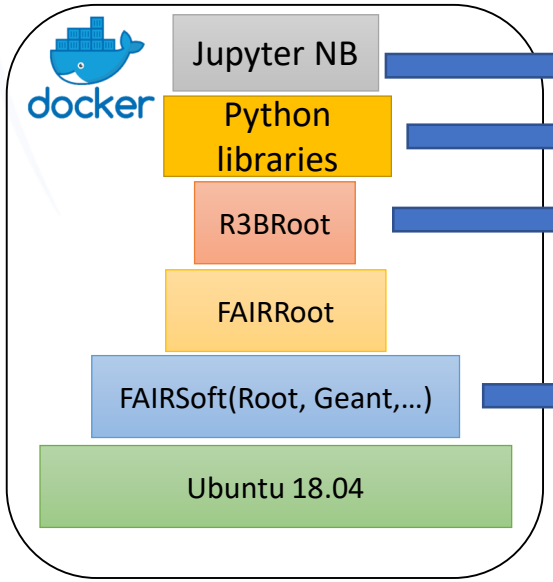
- Finalizing the configuration of DL (FAIR & CERN)
- Testing the Beta version of the JHub service
- Cooperating in the relevant task forces of DAC21 (in harmony with GSI)

Priority 2

- Integration the JHub platform with token based IAM
- Integration with ESAP (ESCAPE Science Analysis Platform)
- Put & test Release Candidate (RC) version of our service in production.

Funded by the European Union's Horizon 2020 - Grant N° 824064

R3B-NeuLAND

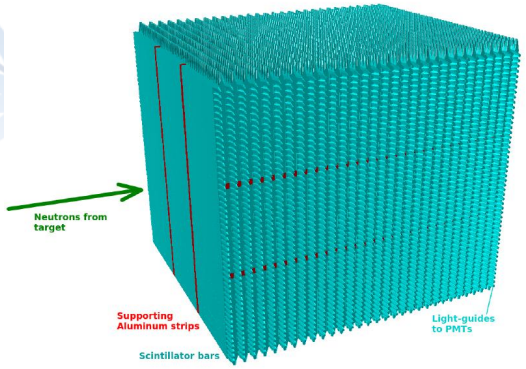


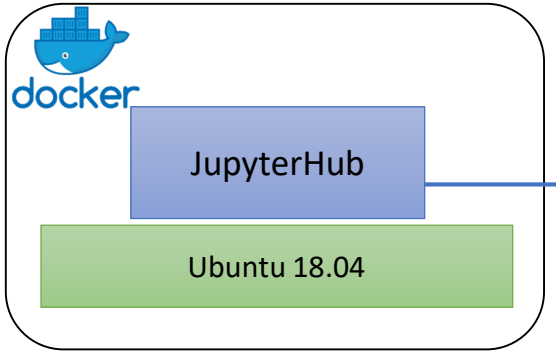
Configured JNB

Numpy, TensorFlow, Keras,...

Made compatible with Docker container needs

Python2 -> Python3





All the necessary configurations for the web server:

- AuthN & AuthZ
- Spawning
- Proxy settings
- ...

Funded by the European Union's Horizon 2020 - Grant N° 824064

RUCIO for Data management

Findable:

- Globally unique and persistent identifier
- Indexed in a searchable resource

Accessible:

- Retrievable by their identifier using a standardised communications protocol
- Support AuthN/Z if necessary

Interoperable:

- Formal, accessible, shared, and broadly applicable language

Reusable:

- Clear and accessible data usage license
- Associated with detailed provenance

SCIENTIFIC DATA MANAGEMENT

RUCIO

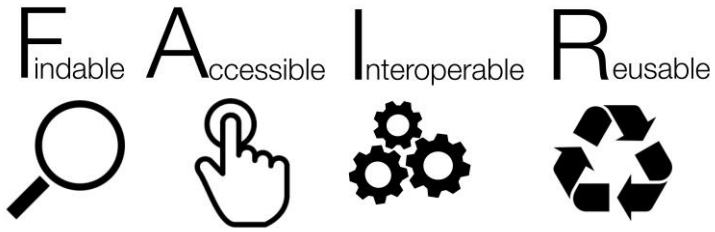
Funded by the European Union's Horizon 2020 - Grant N° 824064

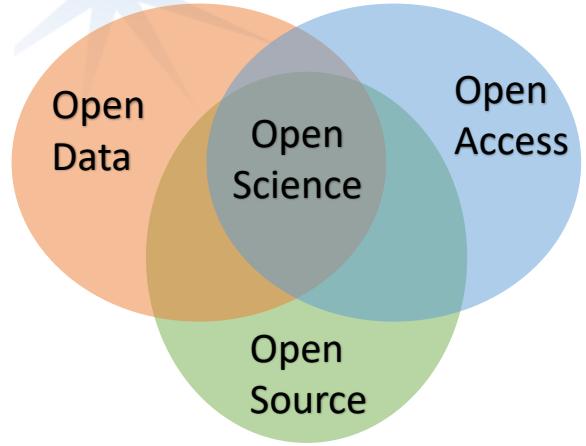
Next steps

- Finilizing configuration of the data-lake VMs on MERLIN
- Orchestrating all the VMs (Kubernetes maybe!)
- Doing a connection test for data-lake between RUG & GSI
- Configuring RUCIO on JupyterLab
- Test GUI prototype of the R3B-NeuLAND simulation software (Using JupyerLab widgets!).

Funded by the European Union's Horizon 2020 - Grant N° 824064

Introduction





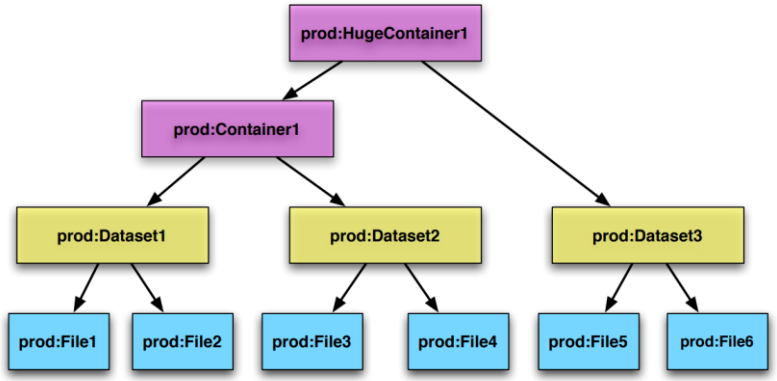
1. When scientists share knowledge and resources their results improve people's lives.

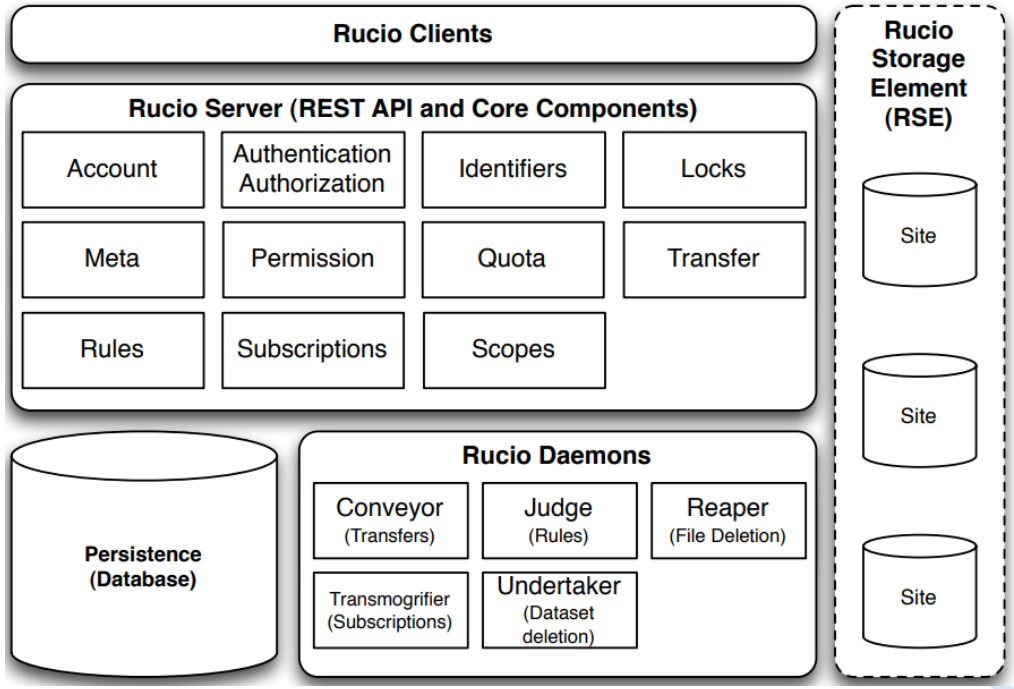
2. Scientists need easy access to state-of-the-art infrastructures to test their ideas.

3. Avoiding unnecessary duplication of efforts

Funded by the European Union's Horizon 2020 - Grant N° 824064

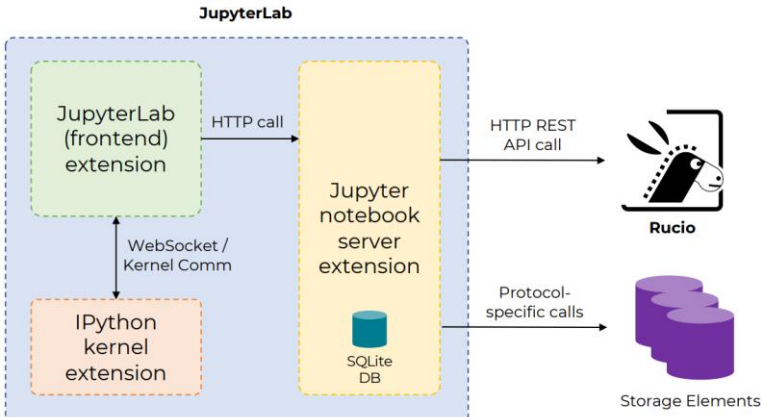
FAIR & Data management





SCIENTIFIC DATA MANAGEMENT

RUCIO



RUCIO Jupyterlab extension

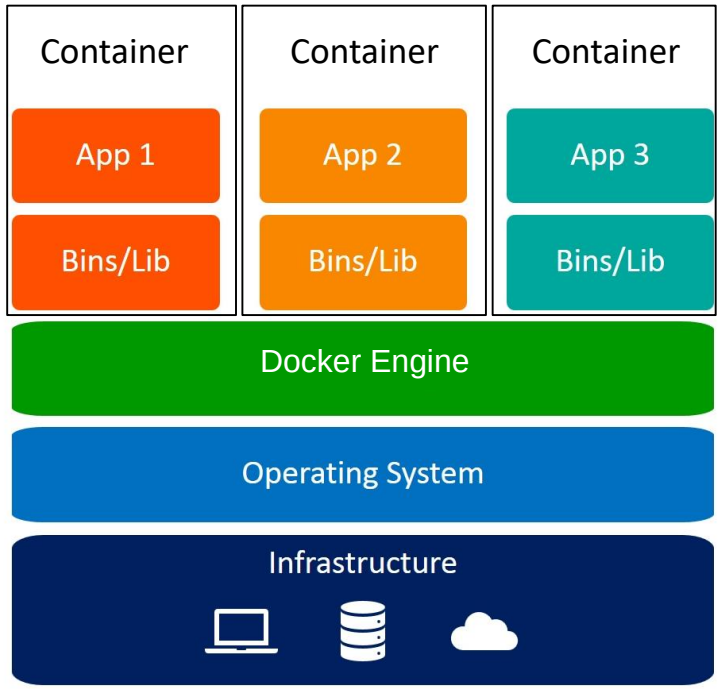
Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Science Analysis Platform (ESAP)

A standard OS-virtualization platform for **packing up software** and all the **dependencies** to **run applications on different environments**.

Advantages:

- Lighter than VM (Mb << Gb)
- Higher performance (~seconds)
- Less resource utilization
- Portability
- Micro services & Scaling



Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Science Analysis Platform (ESAP)

Jupyter Notebook (lab)

The Jupyter Notebook is an open-source **web application** that allows you to **create and share** documents that contain **live code**, **equations**, **visualizations** and **narrative text**.

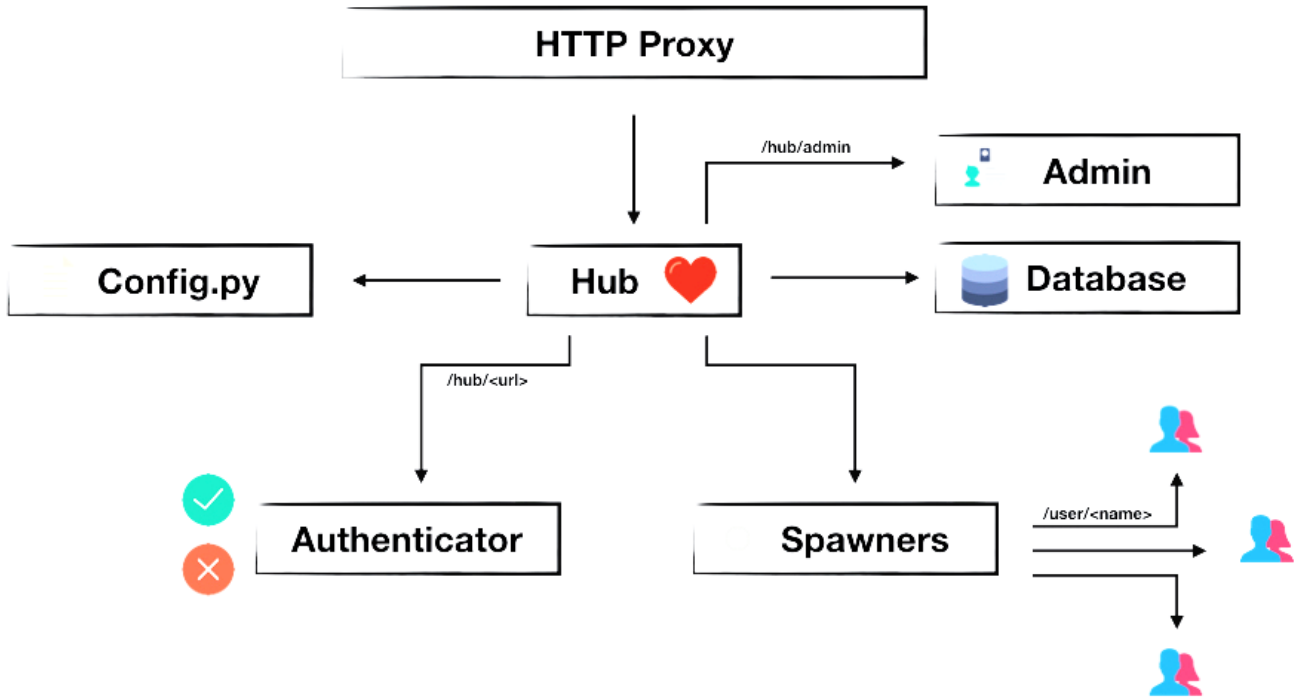
JupyterHub

JupyterHub brings the power of **notebooks** to **groups of users**. It gives users access to **computational environments** and **resources without** burdening the users with **installation and maintenance tasks.**

Funded by the European Union's Horizon 2020 - Grant N° 824064

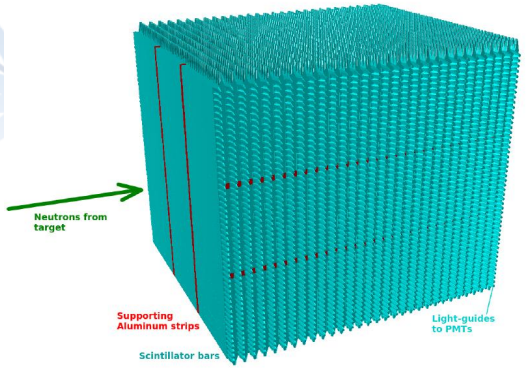
AAI and Spawning

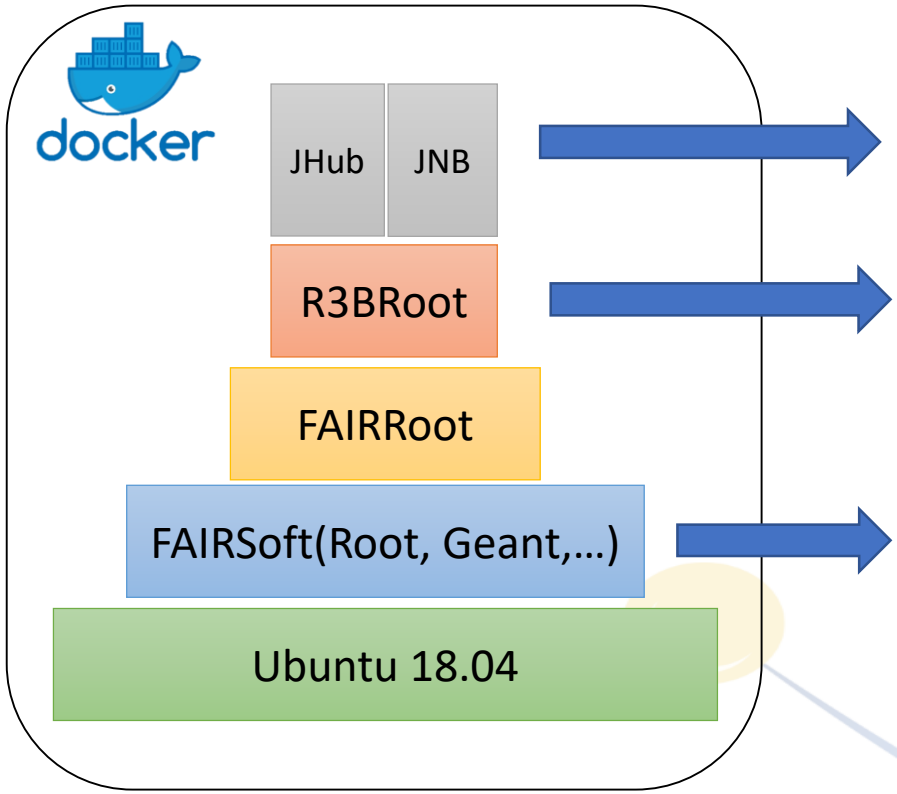
JupyterHub



Funded by the European Union's Horizon 2020 - Grant N° 824064

R3B-NeuLAND





Configured JHub and JNB

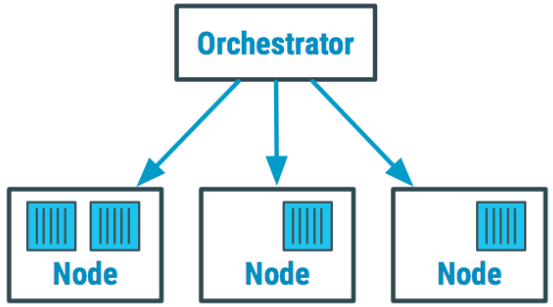
Made compatible with Docker container needs

Python2 -> Python3

Funded by the European Union's Horizon 2020 - Grant N° 824064

Spawning

Container Orchestration:



docker Compose

Orchestration on a single host

- No autho scaling
- Easy to start cluster
- Lower scalibility
- Limited API
- Suitable for small groups (<a few 100s)

docker SWARM

kubernetes

- Autho scaling
- Difficult to start cluster
- Higher scalibility
- All in All API
- Suitable for very large group

Orchestration on cluster

Funded by the European Union's Horizon 2020 - Grant N° 824064

Structure of ESCAPE

ESCAPE DIOS | Data Infrastructure for Open Science

ESCAPE OSSR | Open-source Scientific Software and Service Repository

ESCAPE VO | Virtual Observatory

ESCAPE SAP | Science Analysis Platform

ESCAPE CS | Citizen Science

- Distributed data management
- Keeps track of data locations
- Replication
- Metadata management
- Usage Logging
- Access to existing storage elements
- Definition of access control rules/systems

Funded by the European Union's Horizon 2020 - Grant N° 824064

Structure of ESCAPE

ESCAPE DIOS | Data Infrastructure for Open Science

WP2: **Design, implement, and operate a cloud of data services** for open access and open science supporting FAIR principles at the Exabyte scale.

ESCAPE OSSR | Open-source Scientific Software and Service Repository

WP3: Support for a sustainable **open-access repository** to **share scientific software and services** to the science community.

ESCAPE VO | Virtual Observatory

WP4: **Integrating** distributed infrastructures into one **single virtual astronomy** facility.

ESCAPE SAP | Science Analysis Platform

WP5: A **platform-service gateway** with the capability to **access and analyze data** from multiple collections, access ESFRIs' **software tools**, and bring their own **customized workflows**.

ESCAPE CS | Citizen Science

WP6: **Involving society** in the **scientific discoveries** of the **astronomy** and **subatomic** physics facilities in EOSC.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Structure of ESCAPE

ESCAPE DIOS | Data Infrastructure for Open Science (WP2: EDIOS)

- Task 2.1 Data Lake infrastructure and federation services.
- Task 2.2 Data Lake orchestration service.
- Task 2.3 Integration with compute services.
- Task 2.4 Networking.
- Task 2.5 Authentication and authorization.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Structure of ESCAPE

ESCAPE DIOS | Data Infrastructure for Open Science (WP2: EDIOS)

- Task 2.1 Data Lake infrastructure and federation services.
- Task 2.2 Data Lake orchestration service.
- Task 2.3 Integration with compute services.
- Task 2.4 Networking.
- Task 2.5 Authentication and authorization.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Structure of ESCAPE

ESCAPE SAP | Science Analysis Platform

(WP5: ESAP)

Funded by the European Union's Horizon 2020 - Grant N° 824064

Structure of ESCAPE

ESCAPE SAP Science Analysis Platform (WP5: ESAP)

- **Task 5.1: Discovery and Staging**
  - This task will provide users of the **science platform** with the capability to access and combine data from **multiple collections** and stage that data for **subsequent analysis**.
- **Task 5.2: Software deployment and virtualization**
  - This task will incorporate the work on the software repository described in WP3 and focus on tools and services to **support the virtualization of relevant software packages and pipelines.**
- **Task 5.3: Analysis interfaces, work flows and reproducibility**
  - The analysis interface task combines a number of elements to form the **working surface for the user of the EOSC science platform**.
- **Task 5.4: Integration with HPC and HTC infrastructures**
  - Once data for analysis has been located and staged, and workflows have been defined, the next step is to **deploy those workflows on the underlying processing infrastructure**.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Structure of ESCAPE

ESCAPE SAP Science Analysis Platform (WP5: ESAP)

- **Task 5.1: Discovery and Staging**
  - This task will provide users of the **science platform** with the capability to access and combine data from **multiple collections** and stage that data for **subsequent analysis**.
- **Task 5.2: Software deployment and virtualization**
  - This task will incorporate the work on the software repository described in WP3 and focus on tools and services to **support the virtualization of relevant software packages and pipelines.**
- **Task 5.3: Analysis interfaces, work flows and reproducibility**
  - The analysis interface task combines a number of elements to form the **working surface for the user of the EOSC science platform**.
- **Task 5.4: Integration with HPC and HTC infrastructures**
  - Once data for analysis has been located and staged, and workflows have been defined, the next step is to **deploy those workflows on the underlying processing infrastructure**.

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Science Analysis Platform (ESAP)

Software deployment & virtualization:

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Science Analysis Platform (ESAP)

Software deployment & virtualization:

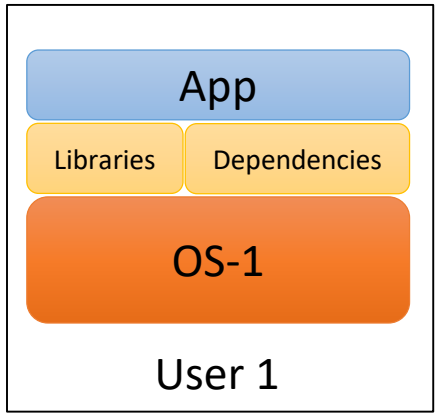
Compatibility and portability issues

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Science Analysis Platform (ESAP)

Software deployment & virtualization:

Compatibility and portability issues

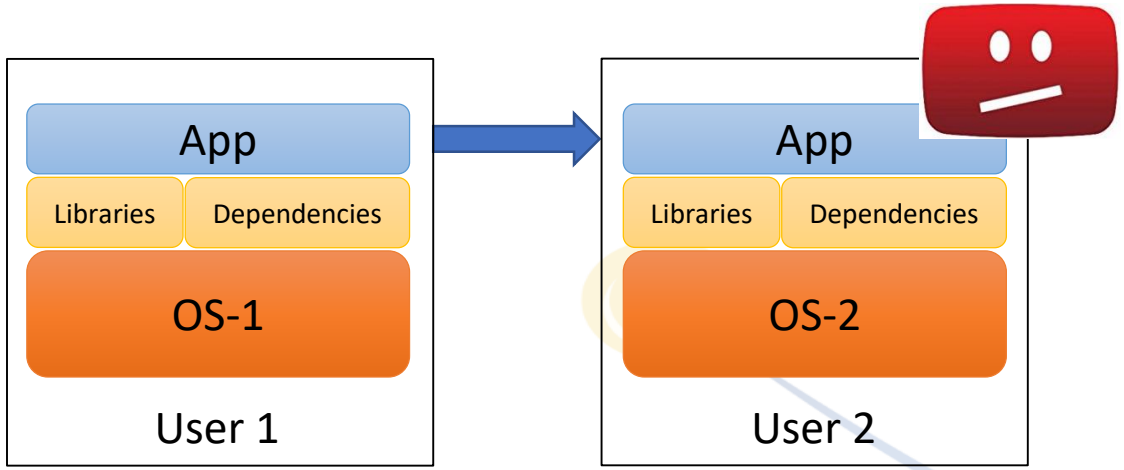


Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Science Analysis Platform (ESAP)

Software deployment & virtualization:

Compatibility and portability issues

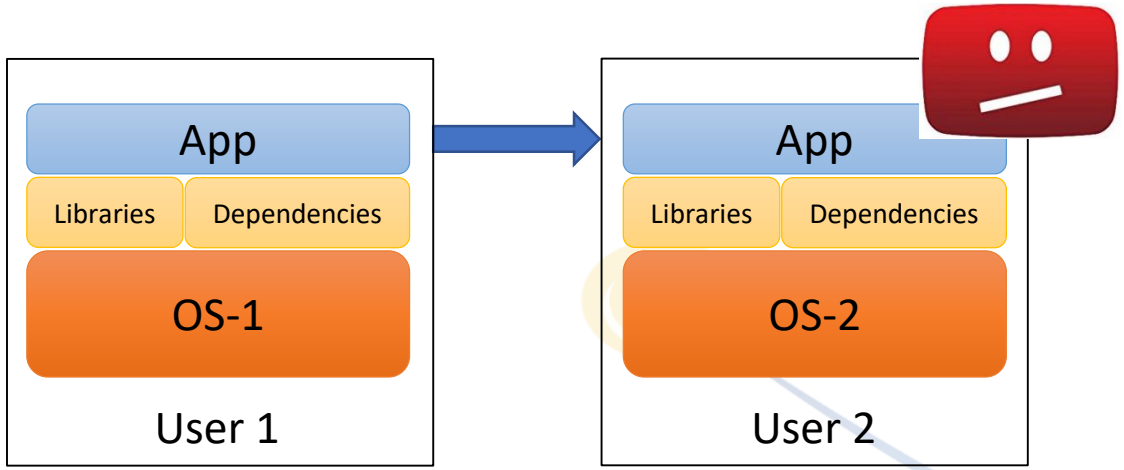


Funded by the European Union's Horizon 2020 - Grant N° 824064

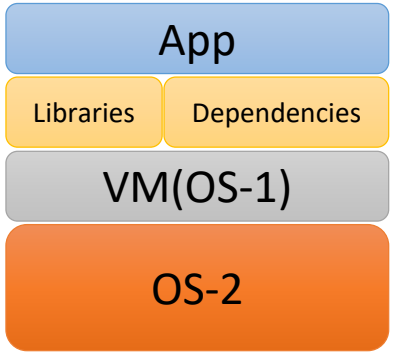
ESCAPE Science Analysis Platform (ESAP)

Software deployment & virtualization:

Compatibility and portability issues



Virtual Machine?!



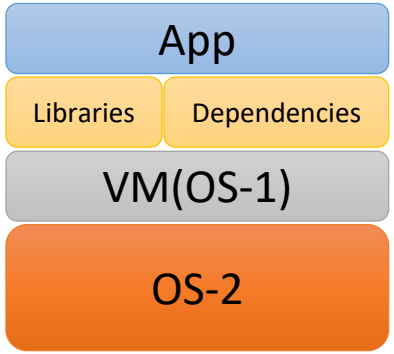
Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Science Analysis Platform (ESAP)

Software deployment & virtualization:

1. **Compatibility and portability issues**
2. **Configuration is time consuming**
3. **Scaling**
4. **Resource utilization**

Virtual Machine?!



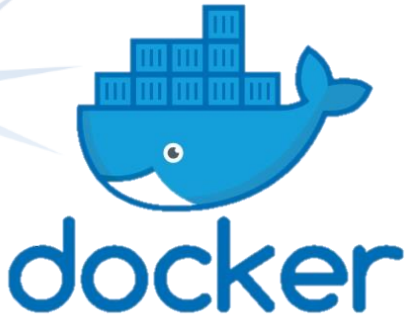
User 2

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Science Analysis Platform (ESAP)

Software deployment & virtualization:

A standard OS-virtualization platform for **packing up software** and all the **dependencies** to **run applications on different environments**.



Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Science Analysis Platform (ESAP)

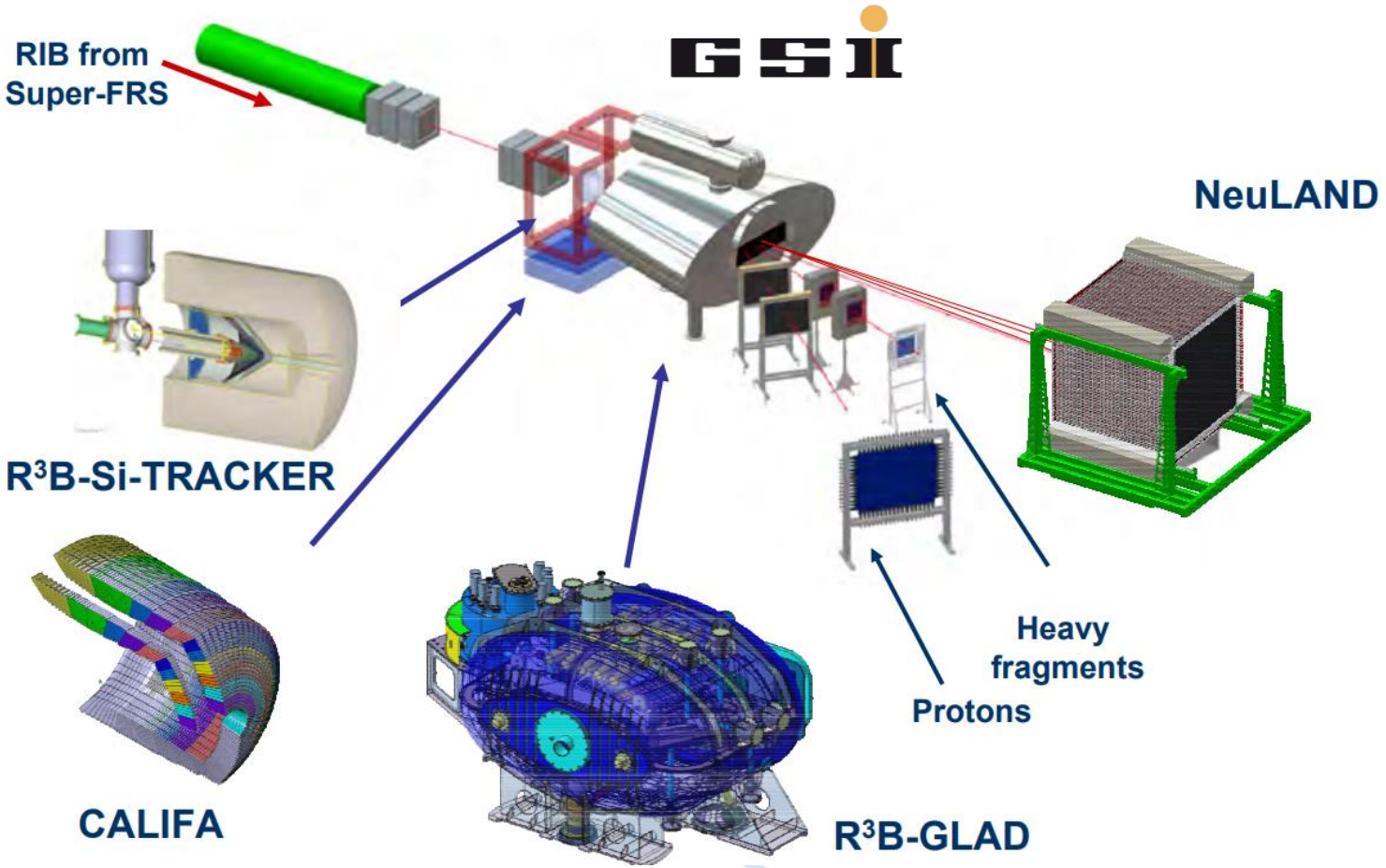
Software deployment & virtualization:

A standard OS-virtualization platform for **packing up software** and all the **dependencies** to **run applications on different environments**.

Funded by the European Union's Horizon 2020 - Grant N° 824064

R3B-NeuLAND

R3B (**R**eactions with **R**elativistic **R**adioactive **B**eams)

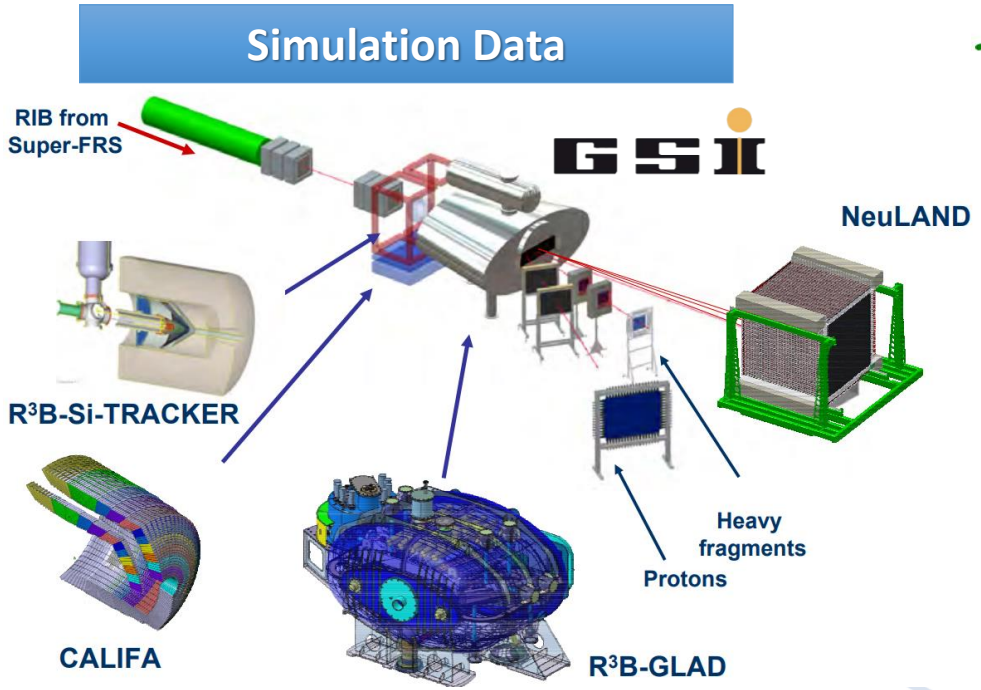


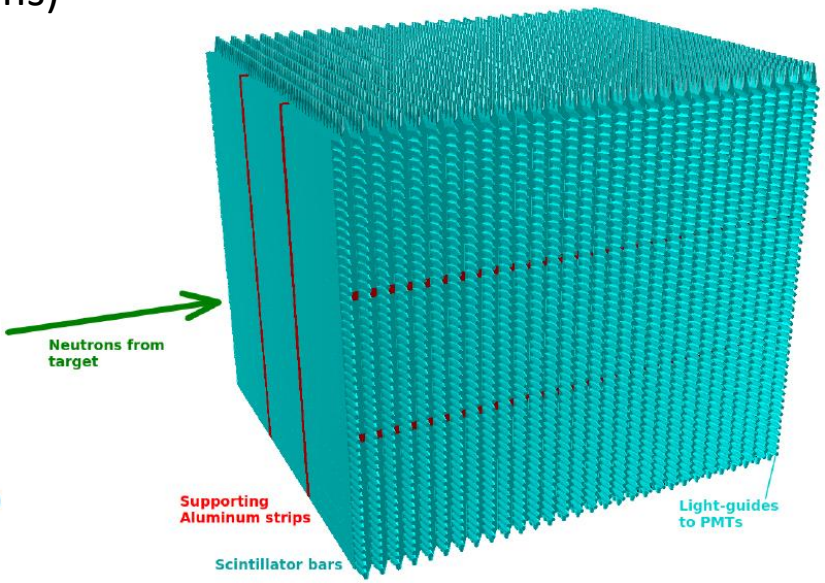
Funded by the European Union's
Horizon 2020 - Grant N° 824064

R3B-NeuLAND

R3B (**R**eactions with **R**elativistic **R**adioactive **B**eams)

- Multiplicity determination
- Shower head determination

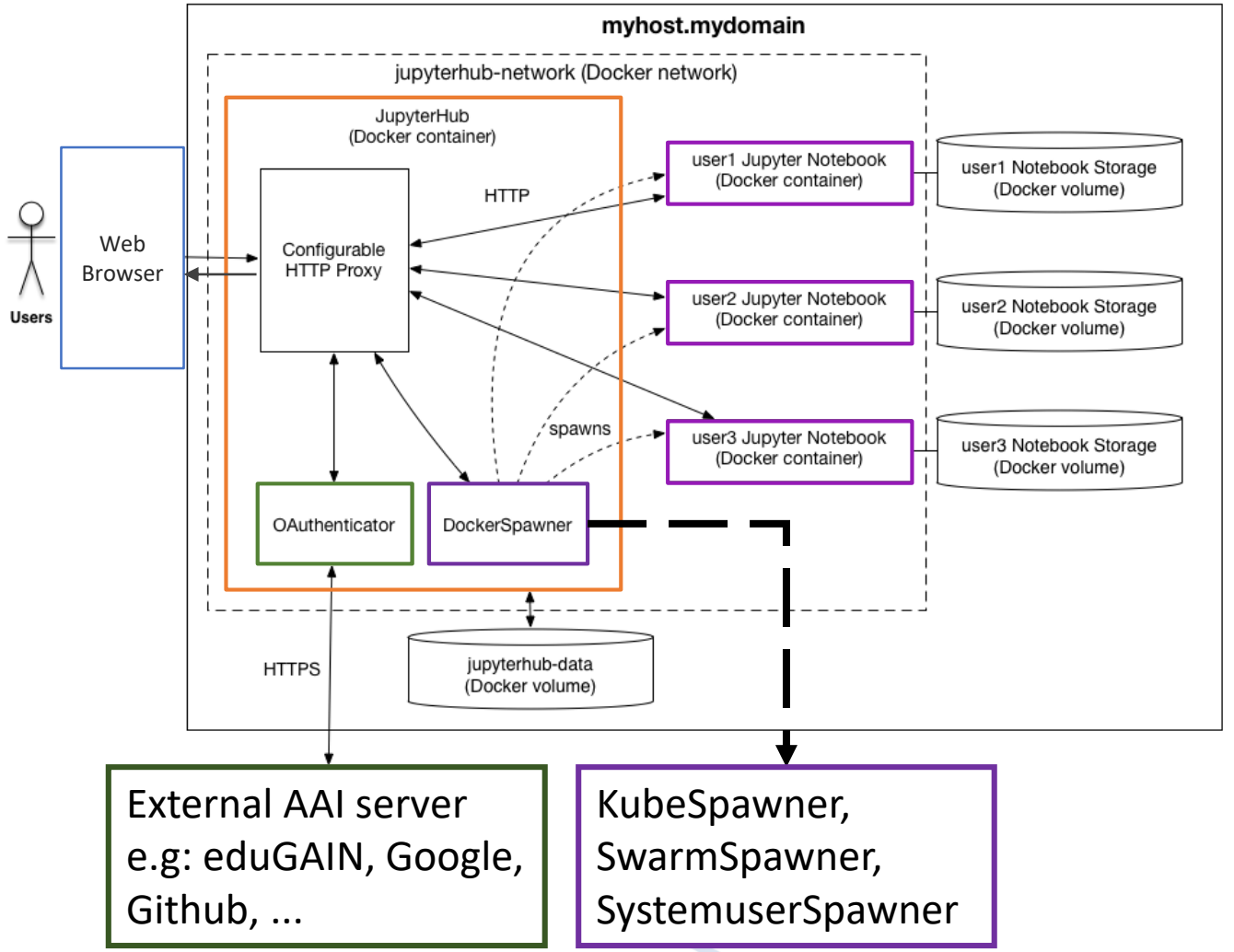




- 30 doubleplane (dp)
- 100 V-H scintillators/dp
- 2.5*2.5 m^2

Funded by the European Union's Horizon 2020 - Grant N° 824064

...Recap so far...



Funded by the European Union's
Horizon 2020 - Grant N° 824064