

The Virtual Research Environment

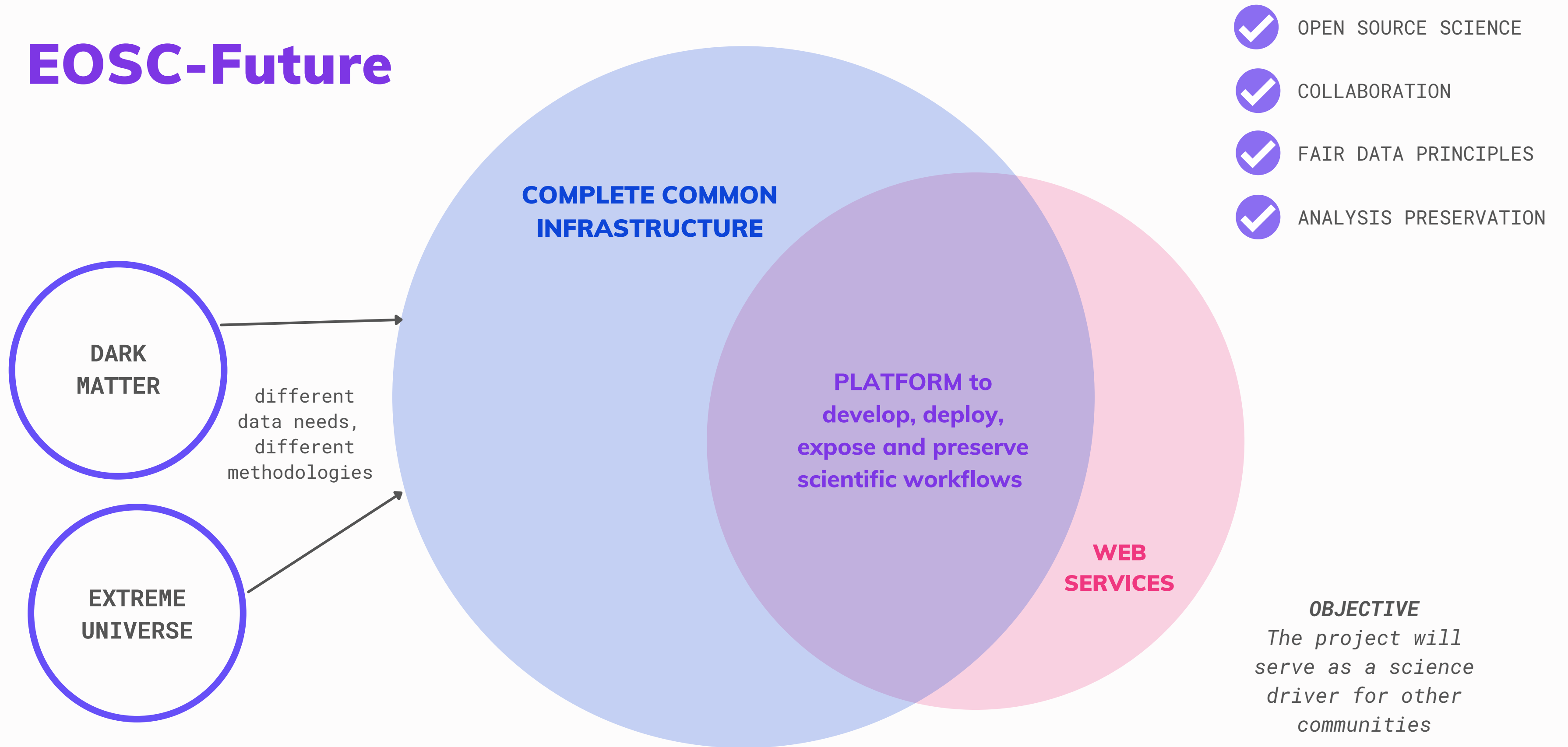
Developing an online collaborative platform to serve astro-particle physics communities

Elena Gazzarrini
Alba Vendrell Moya
Enrique Garcia

EOSC Future

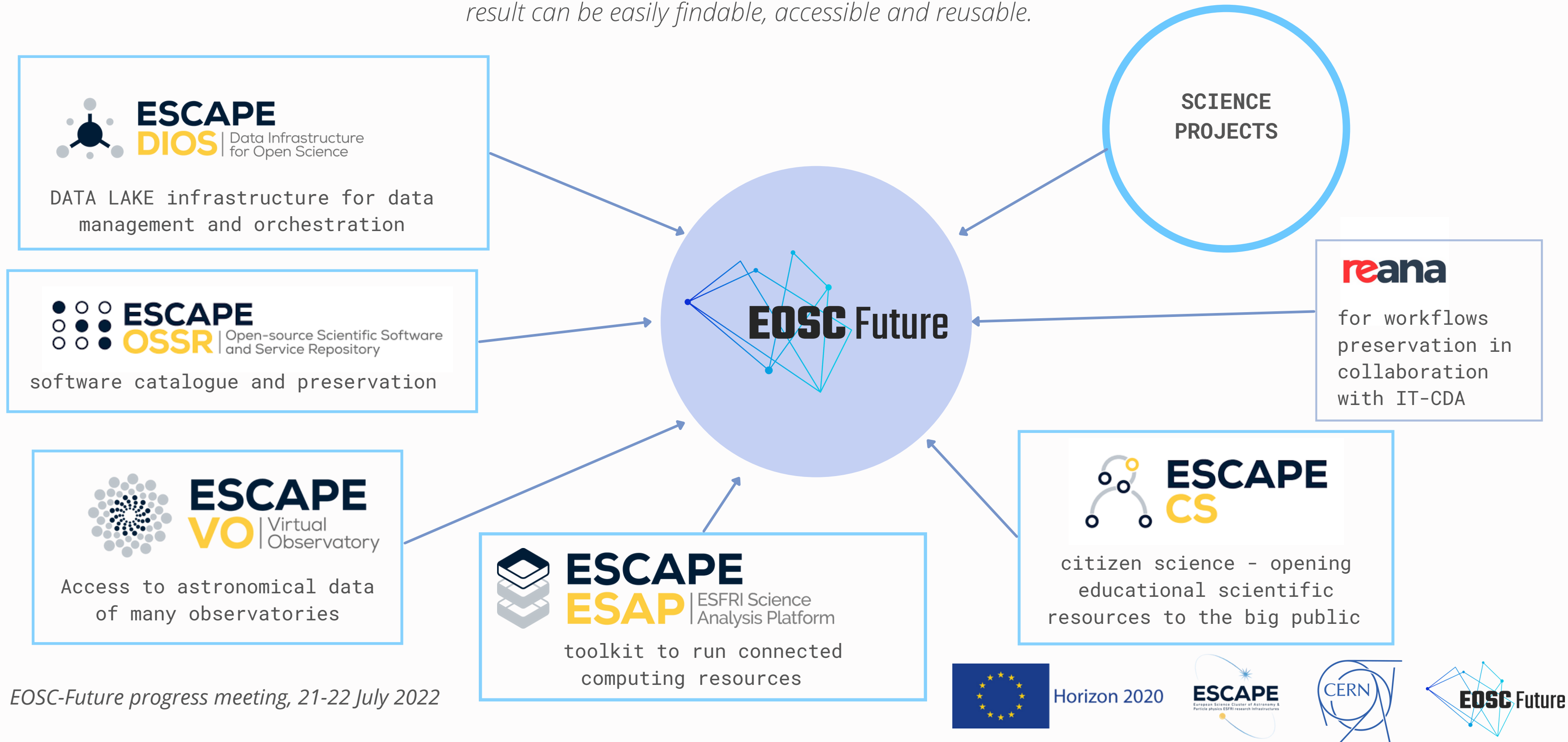


EOSC-Future

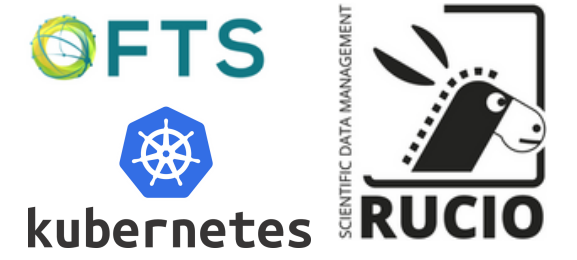
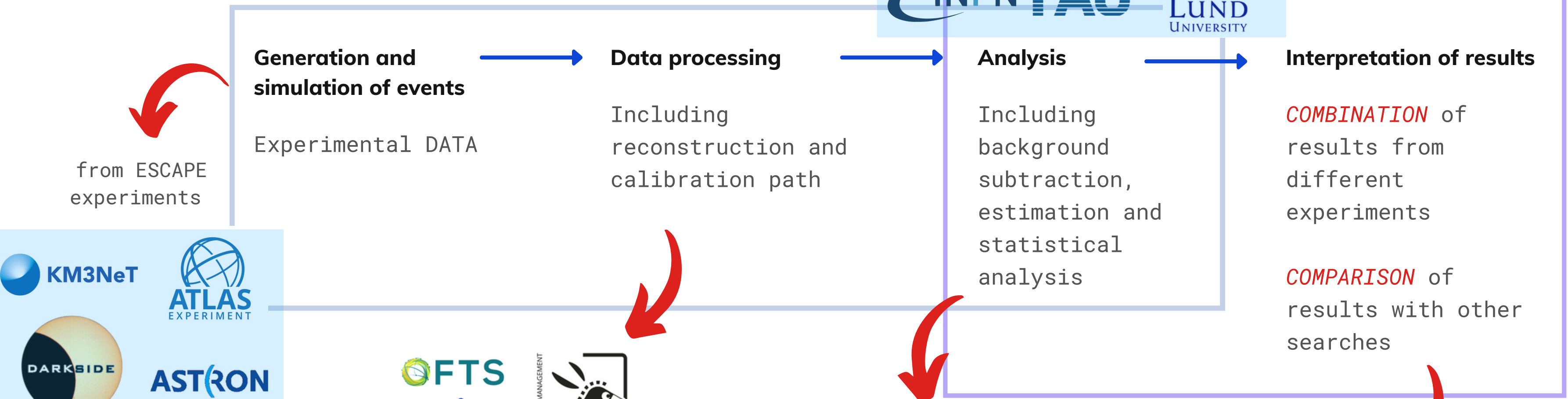
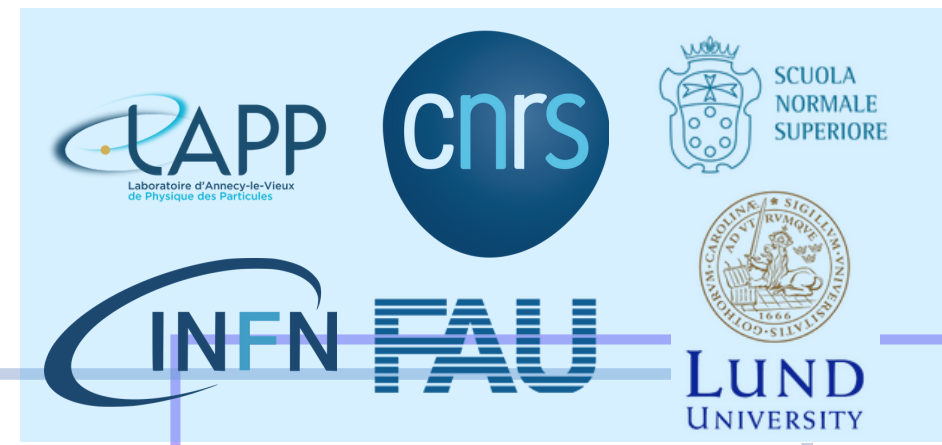


The VRE

A collaborative **online platform** where interested researchers members will be able develop their **Science Projects**. It will host and implement the services provided by the **ESCAPE WPs 2, 3, 4, 5 and 6** to provide an environment where the digital content of a scientific result can be easily findable, accessible and reusable.



VRE concept



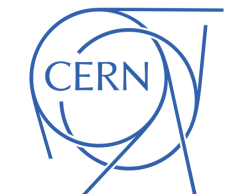
ESCAPE's Data Lake running on a K8s cluster:
storage orchestration
+
data management



WebUI:
DataLake-as-a-Service
(credits to Riccardo Di Maria, Muhammad Aditya Hilmy)



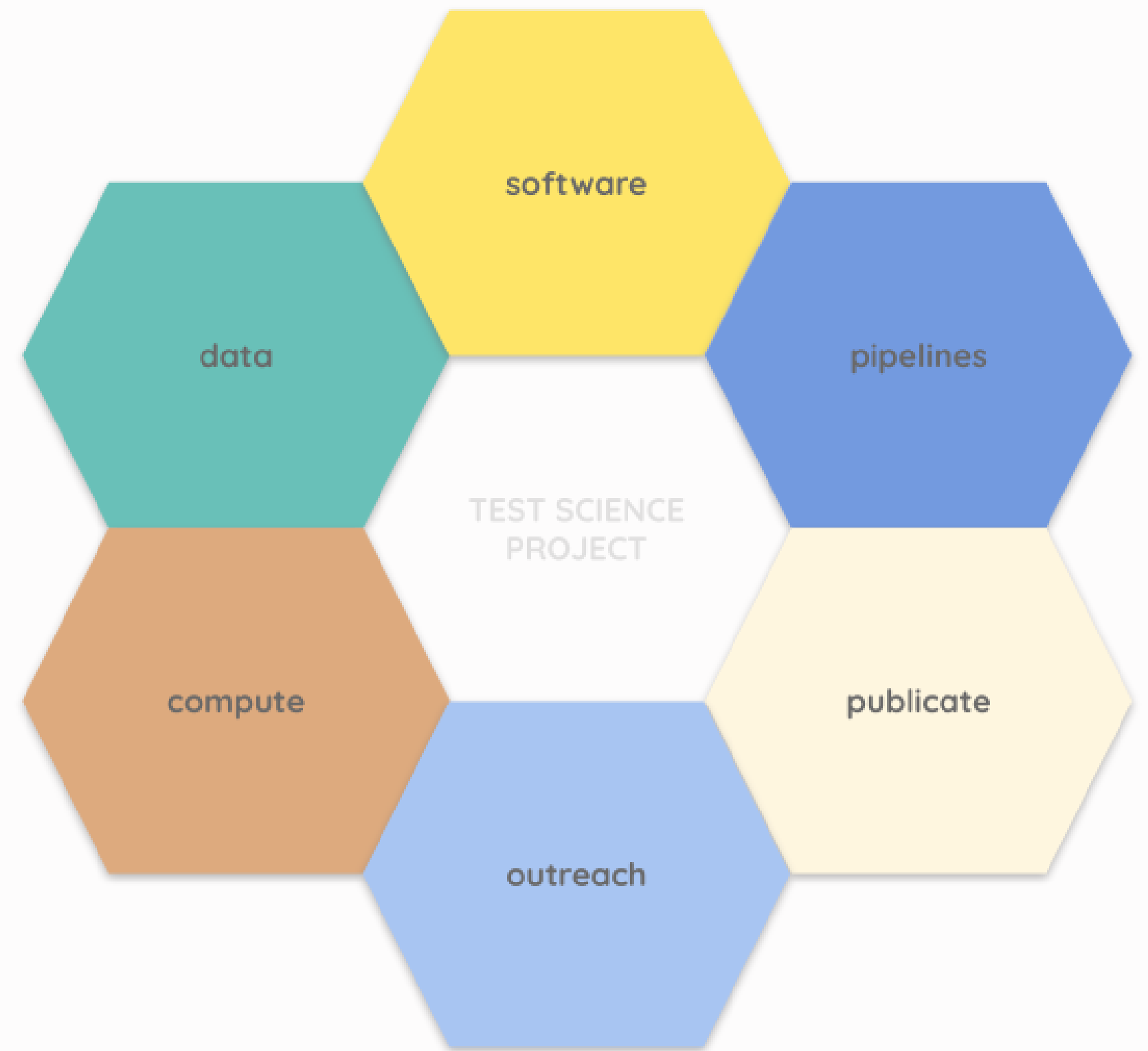
a platform to reproduce workflows running on various computing backends



VRE principles

WebUI
interface

To facilitate
onboarding.
Good
documentation
is key.



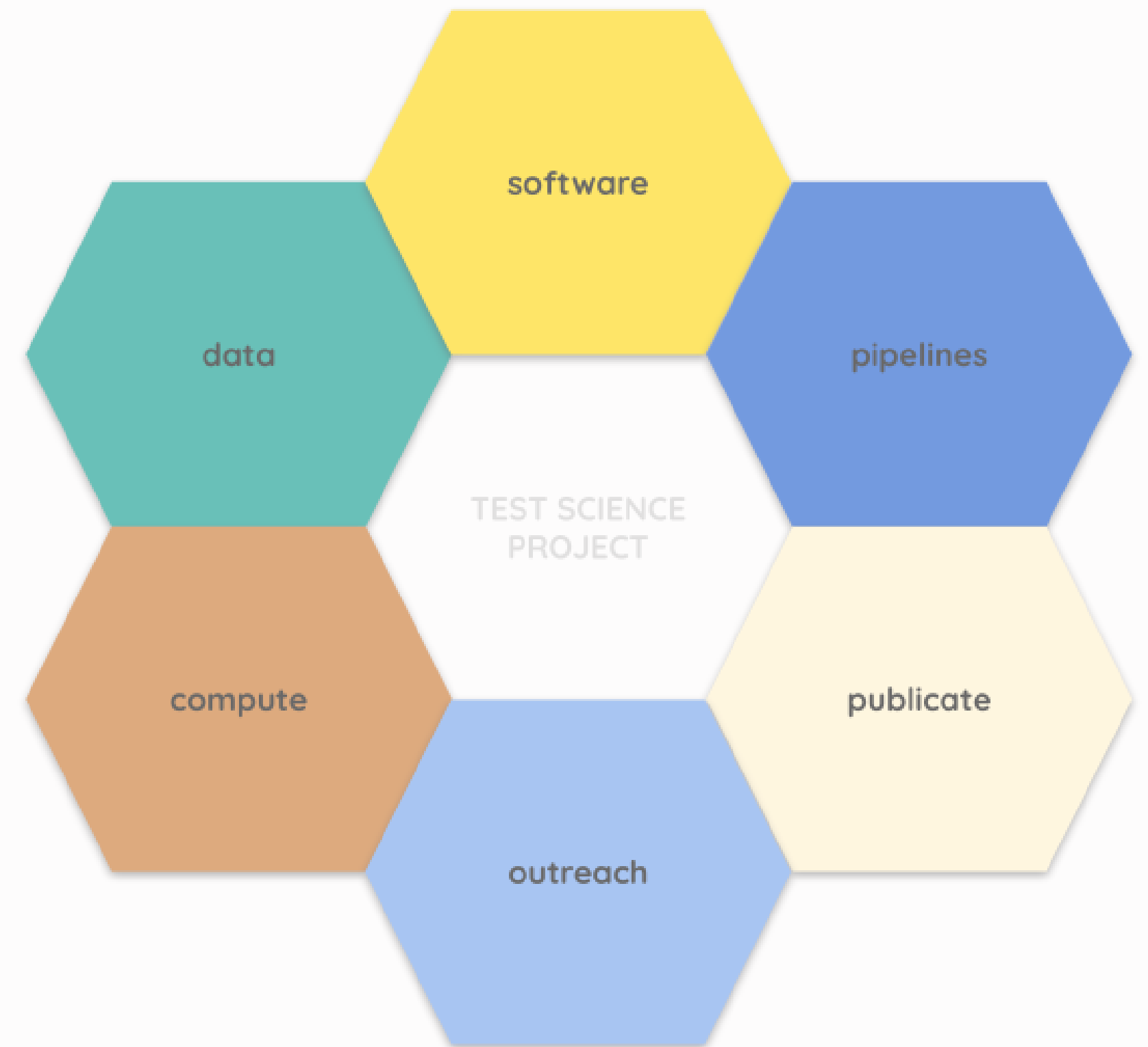
VRE principles

WebUI interface

To facilitate onboarding.
Good documentation is key.

Scalability

Possibility of running jobs at various scales.
Collaboration to make it easy to re-deploy the service on external clusters.



VRE principles

WebUI interface

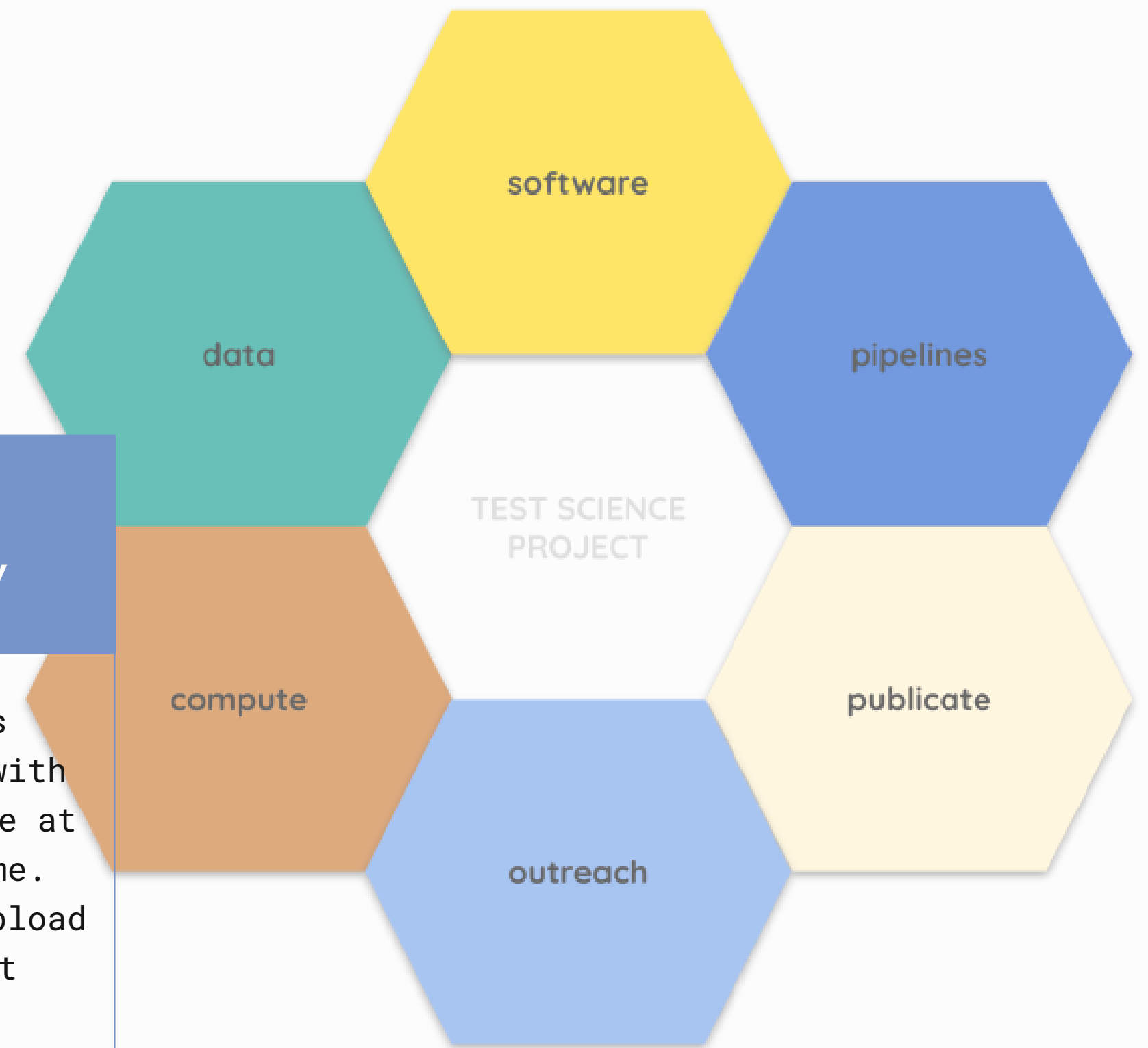
To facilitate onboarding.
Good documentation is key.

Scalability

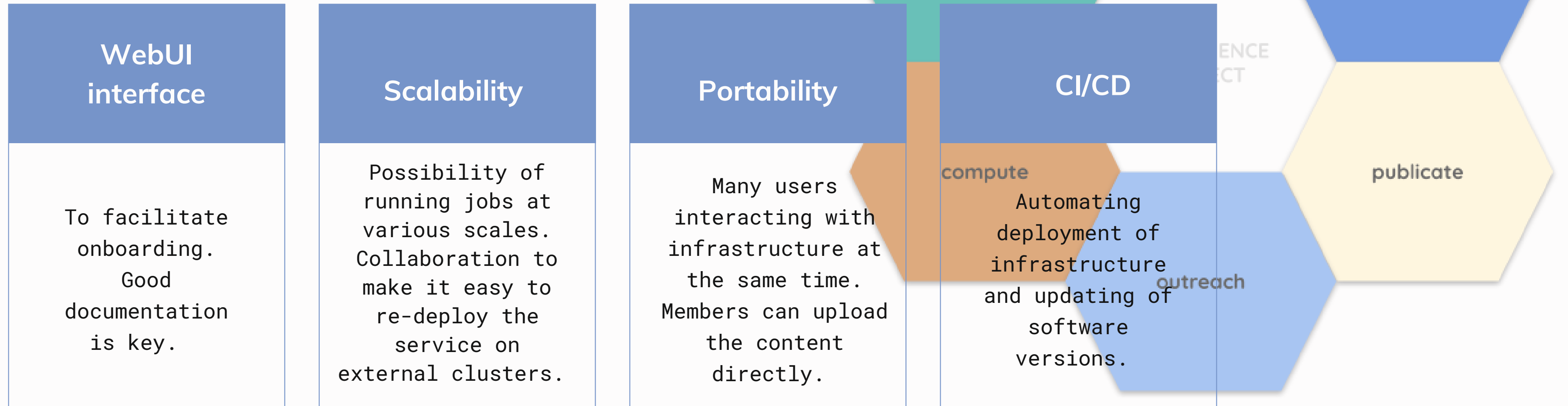
Possibility of running jobs at various scales. Collaboration to make it easy to re-deploy the service on external clusters.

Portability

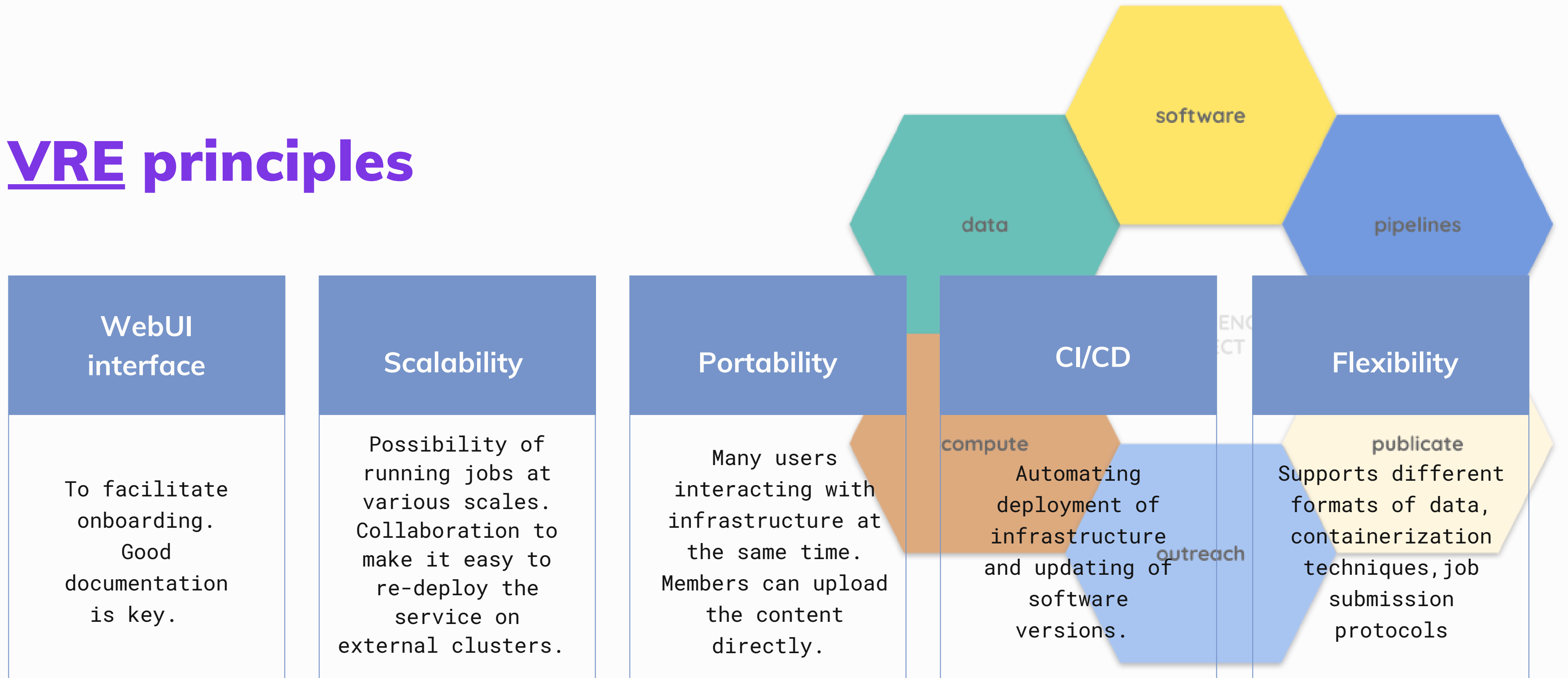
Many users interacting with infrastructure at the same time. Members can upload the content directly.



VRE principles



VRE principles



VRE landing page

ESCAPE VRE

Virtual Research Environment

AN AGGREGATION PORTAL PROPOSAL

GET STARTED →

RESOURCES

- HOME
- RESEARCH TOOLS
- TEST SCIENCE PROJECT - HIGGS
- TEST SCIENCE PROJECT - ASTRO
- DOCS
- ESCAPE RESEARCHER
- EXTERNAL RESEARCHER

Test Science Project: the Higgs @ATLAS

AN EXERCISE OF HOW A PARTICLE PHYSICS TSP CAN LOOK LIKE

The Physics

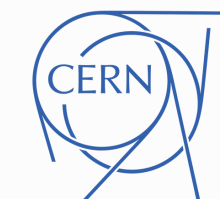
Take a look at the fundamental physics that support and guide the experimental data analysis searches -and discoveries- relative to Higgs candidates, predicted by the Standard Model of Particle Physics.

LEARN MORE

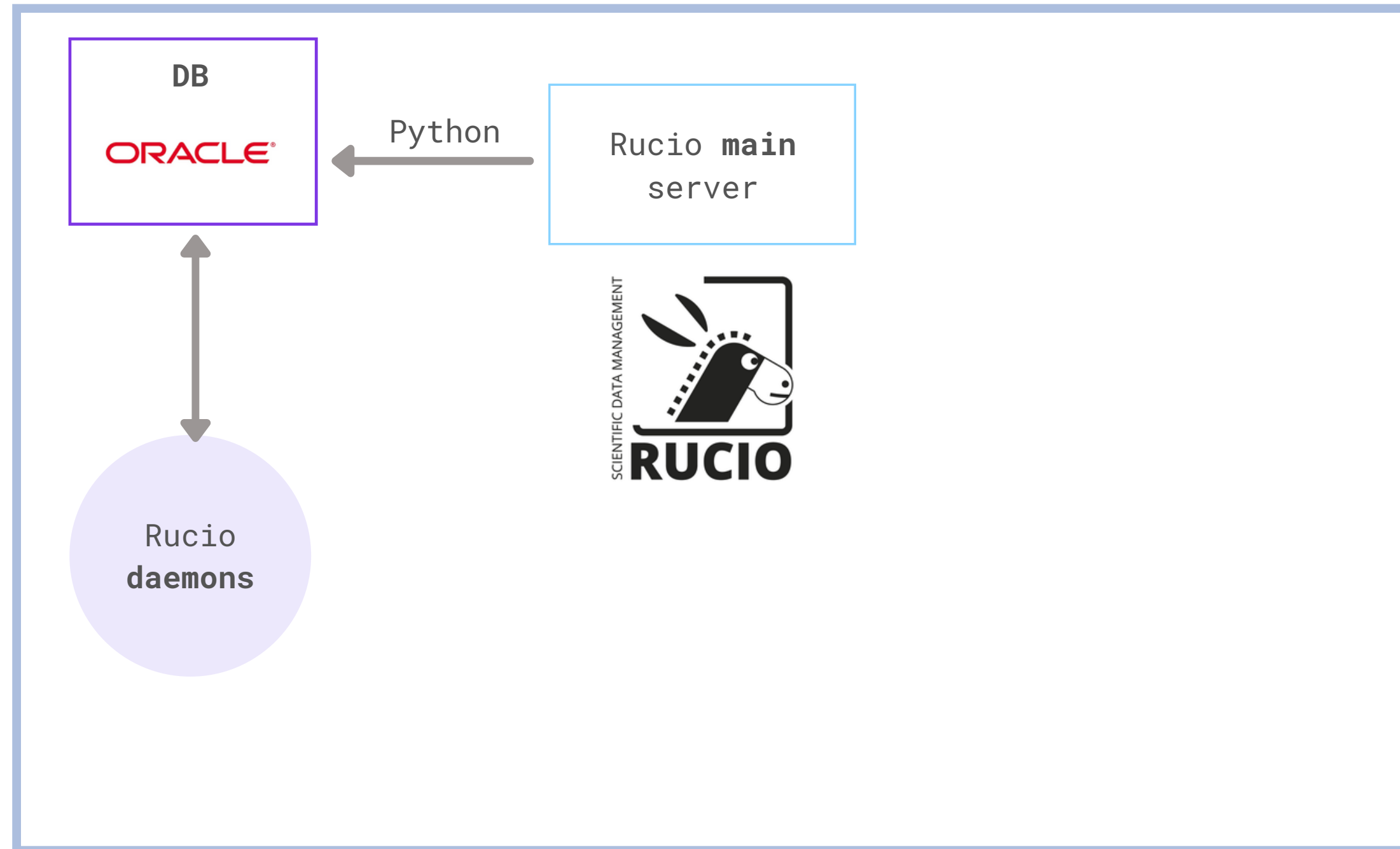
The Experiments

The ATLAS experiment at the LHC is a many-layered instrument designed to detect some of the tiniest yet most energetic particles ever created on earth. It consists of six different detecting subsystems wrapped concentrically in layers around the collision point to record the trajectory, momentum, and energy of particles.

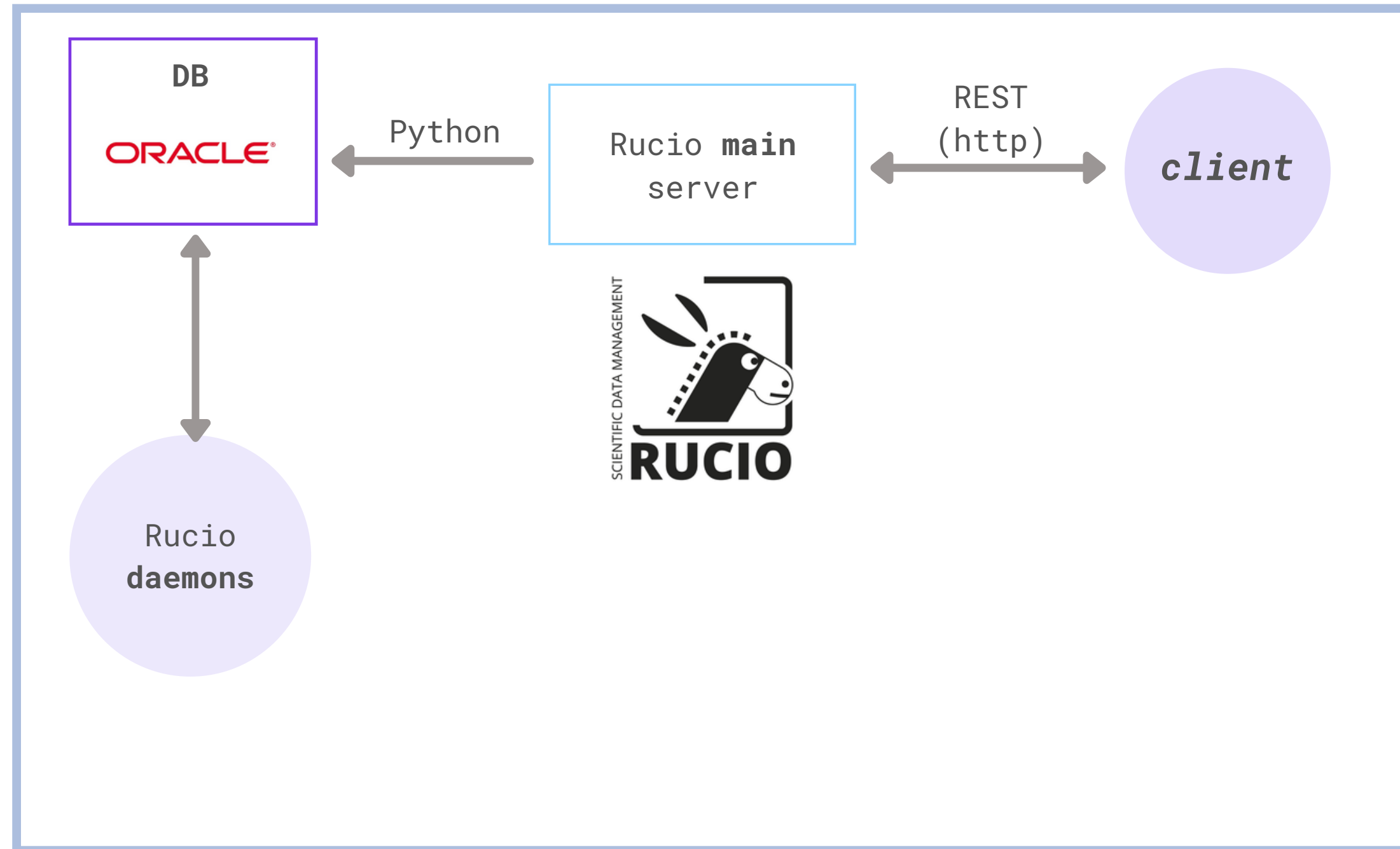
credits to Arturo Sanchez Pineda



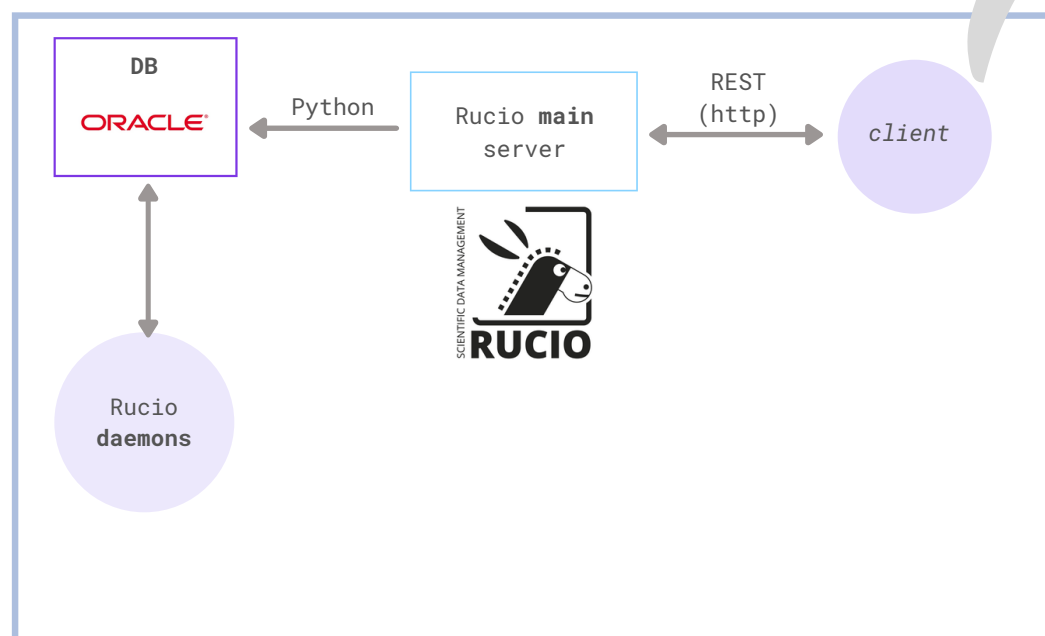
Infrastructure overview



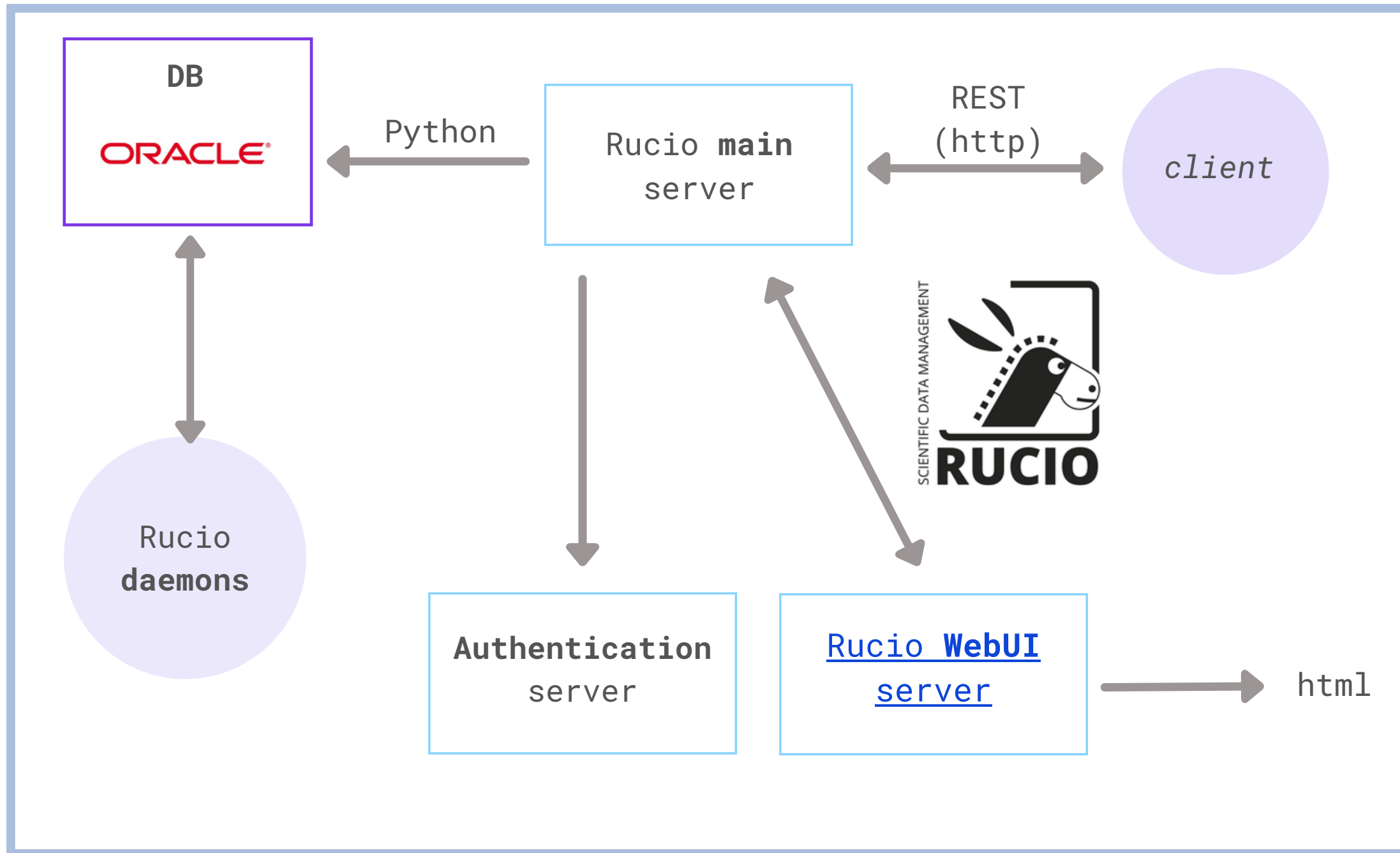
Infrastructure overview



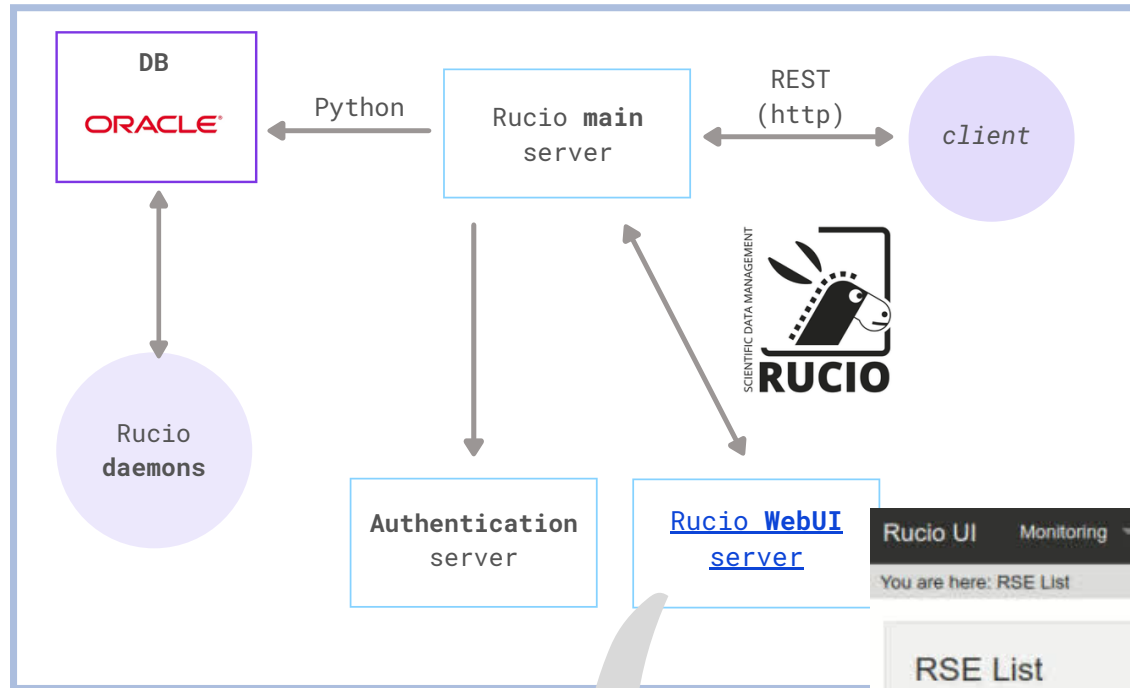
Infrastructure overview



Infrastructure overview



Infrastructure overview



Rucio UI Monitoring Data Transfers (R2D2) Admin pattern OR name OR rule id Search Using account: egazzarr

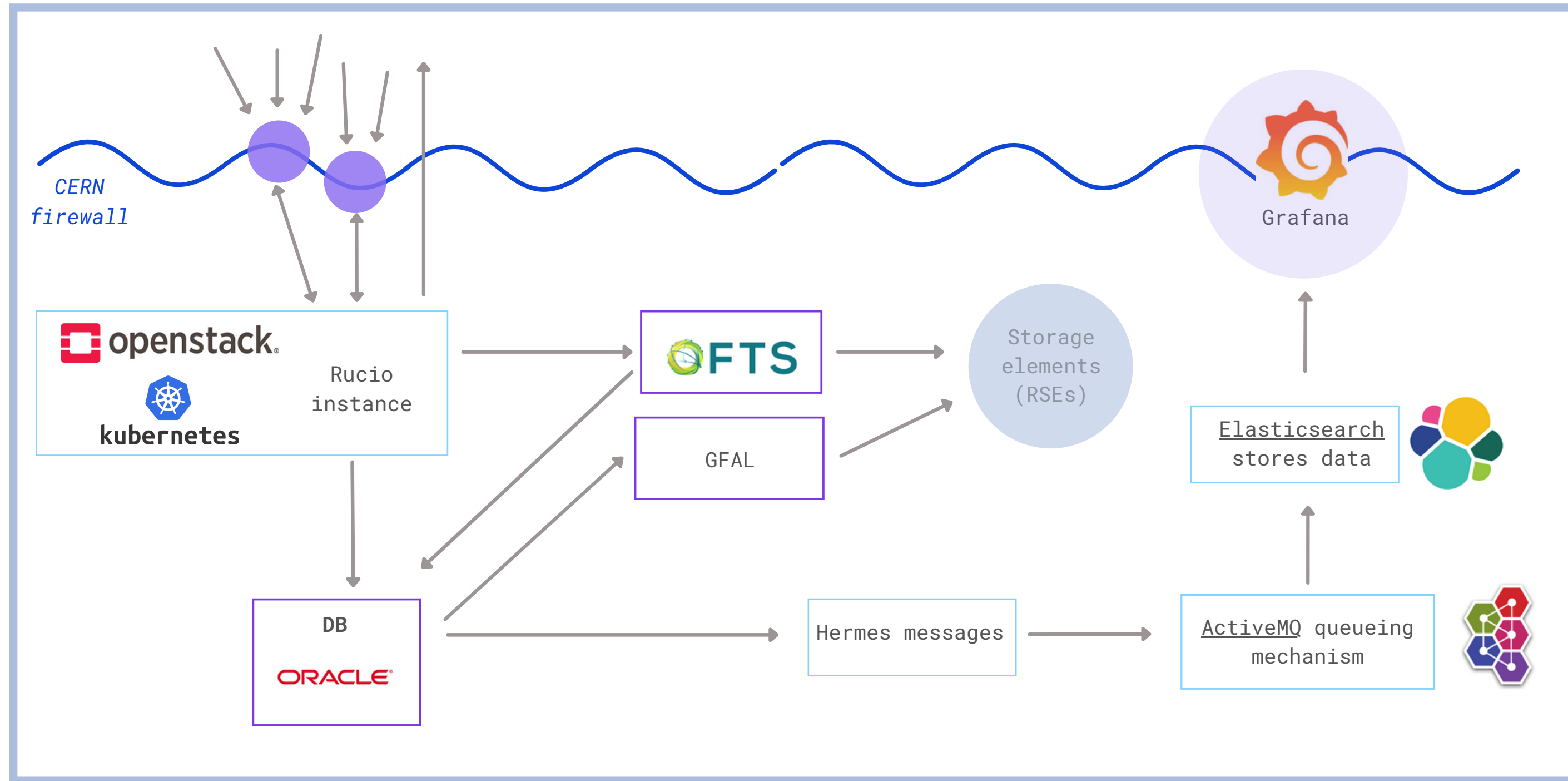
You are here: RSE List Rucio Version (WebUI / Server): 1.25.7 / 1.28.7

RSE List

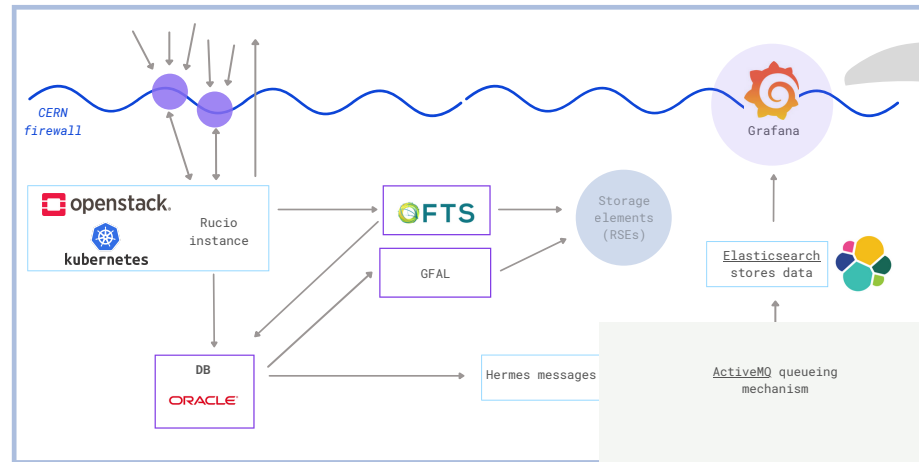
Show 100 Search:

| RSE | Type | City | Region | Country | ISP | Time Zone | Volatile | Deterministic | Staging Area |
|----------------------|------|-------------------|--------|---------|-----|-----------|----------|---------------|--------------|
| ALPAMED-DPM | DISK | Missing from CRIC | | NULL | | | true | false | false |
| AWS_WEBDAV | DISK | Missing from CRIC | | NULL | | | true | false | false |
| CERNBOX-CS3 | DISK | Missing from CRIC | | NULL | | | true | false | false |
| CNAF-STORM | DISK | Missing from CRIC | | NULL | | | true | false | false |
| CNAF-STORM-TAPE | TAPE | Missing from CRIC | | NULL | | | true | false | false |
| CNAF_CMS_TEMP | DISK | Missing from CRIC | | NULL | | | true | false | false |
| DESY-DCACHE | DISK | Missing from CRIC | | NULL | | | true | false | false |
| DESY-DCACHE-NDR | DISK | Missing from CRIC | | NULL | | | false | false | false |
| DESY-DCACHE-TAPE | TAPE | Missing from CRIC | | NULL | | | true | false | false |
| EULAKE-1 | DISK | Missing from CRIC | | NULL | | | true | false | false |
| EULAKE-EC | DISK | Missing from CRIC | | NULL | | | true | false | false |
| FAIR-ROOT | DISK | Missing from CRIC | | NULL | | | true | false | false |
| GSI-ROOT | DISK | Missing from CRIC | | NULL | | | true | false | false |
| IN2P3-CC-DCACHE | DISK | Missing from CRIC | | NULL | | | true | false | false |
| IN2P3-CC-LSST-DEST | DISK | Missing from CRIC | | NULL | | | true | false | false |
| IN2P3-CC-LSST-SOURCE | DISK | Missing from CRIC | | NULL | | | true | false | false |
| INFN-NA-DPM | DISK | Missing from CRIC | | NULL | | | true | false | false |
| INFN-NA-DPM-FED | DISK | Missing from CRIC | | NULL | | | true | false | false |
| INFN-ROMA1 | DISK | Missing from CRIC | | NULL | | | true | false | false |
| JUPYTER-SCRATCH | DISK | Missing from CRIC | | NULL | | | false | false | false |
| LAPP-DCACHE | DISK | Missing from CRIC | | NULL | | | true | false | false |

Networking + monitoring overview



Networking + monitoring overview

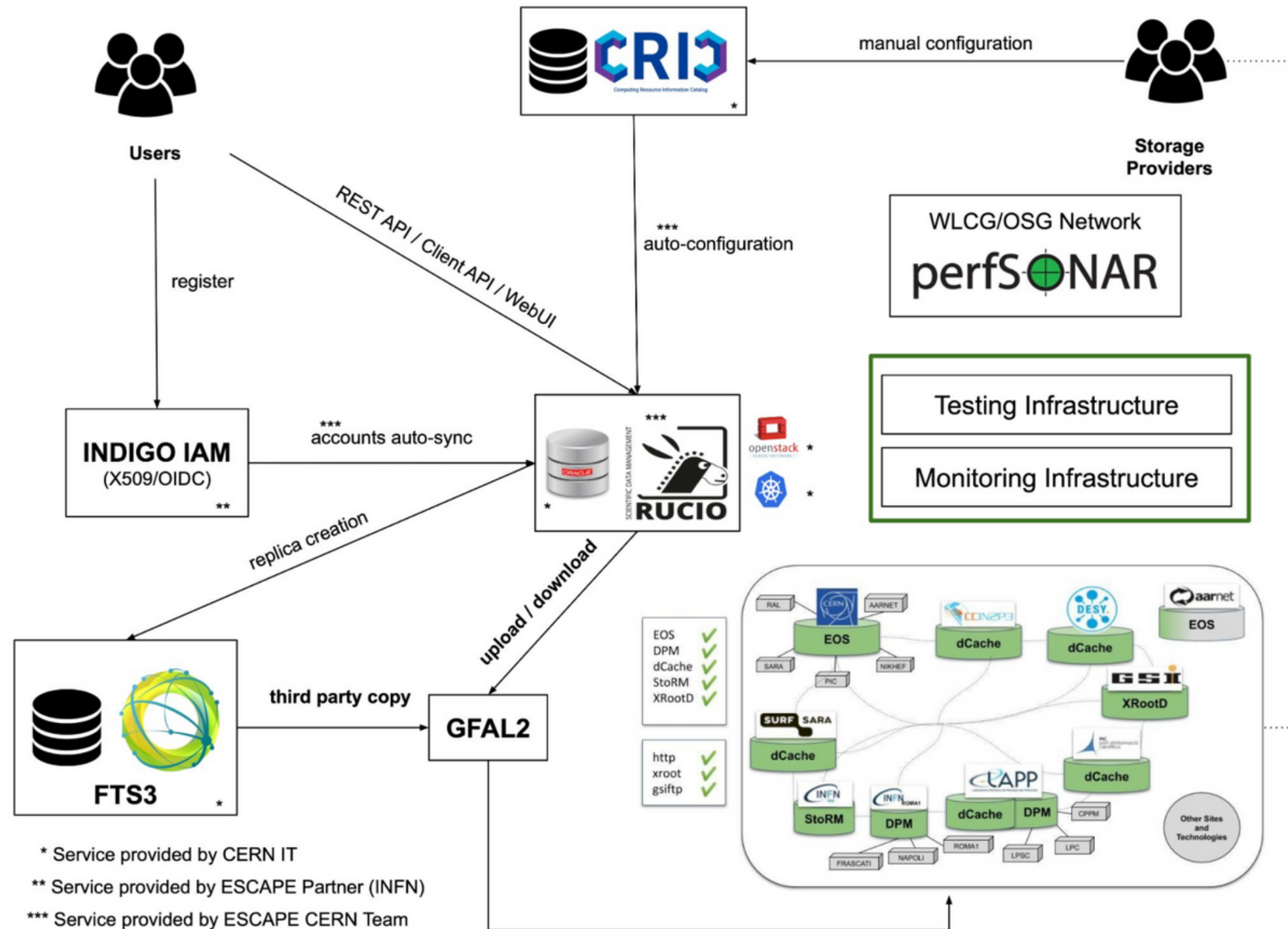


folder:current

Sort (Default A-Z) Filter by tag

- Main Services
 - FTS Transfers
 - Main Services
 - escape fts
 - Gfal Testing
 - Main Services
 - gfal
 - Rucio Events
 - Main Services
 - rucio
 - Rucio Stats (DIDs)
 - Main Services
 - Rucio Stats (Replicas)
 - Main Services

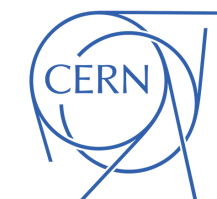
The Data Lake architecture



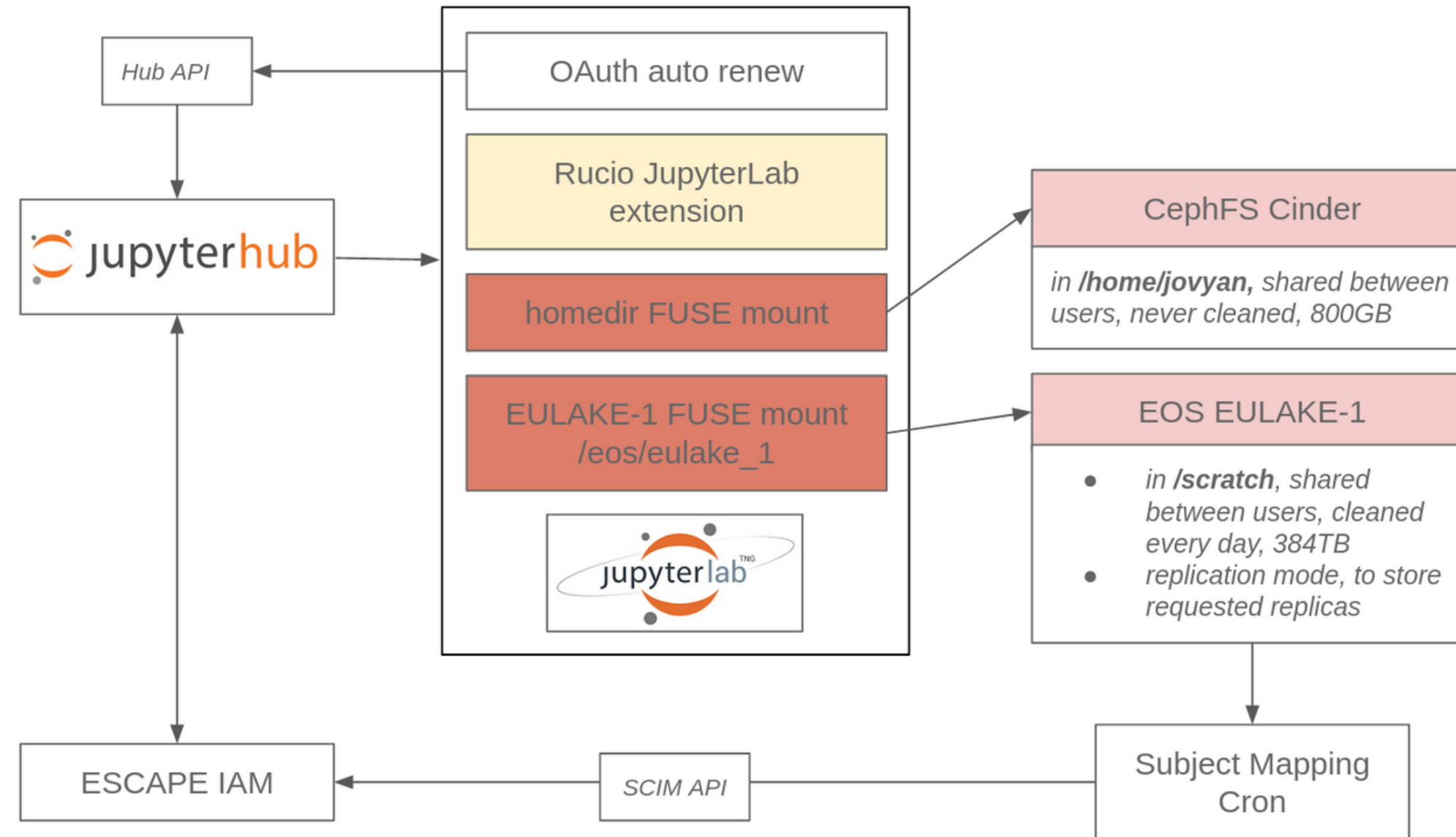
A federated data infrastructure

- **storage** services (RSEs) across Europe
- **data orchestration** with [Rucio](#), policy-driven data management
- **file transfer** and with [FTS3](#), multiple protocols (davs, xrootd, gsiftp)
- grid file access library with [GFAL2](#)
- Identity and Access Management (**IAM**) deployed at INFN/CNAF - *new Rucio version supporting token-based auth*
- [CRIC](#) Global Information Catalogue

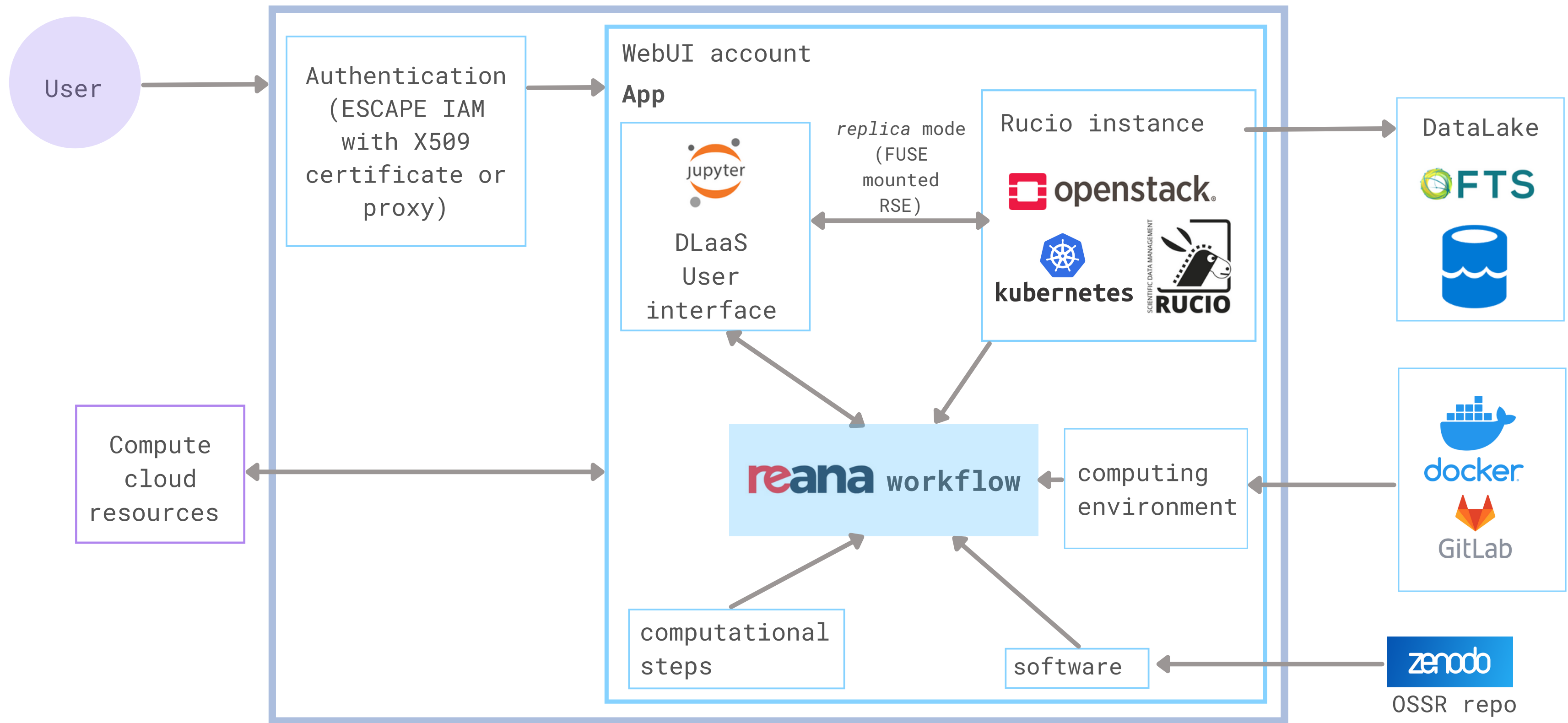
* Service provided by CERN IT
 ** Service provided by ESCAPE Partner (INFN)
 *** Service provided by ESCAPE CERN Team



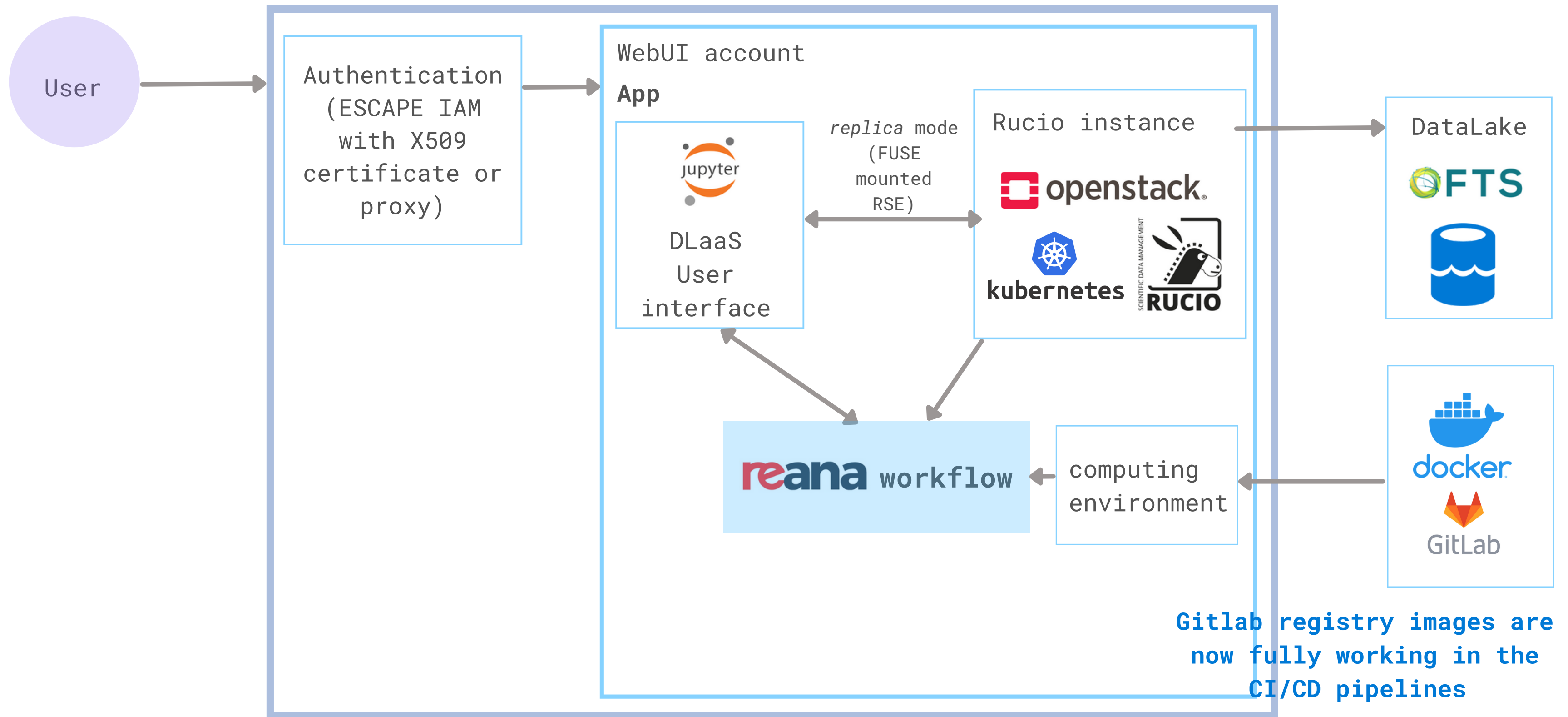
A step further - Data Lake as a Service (DLaaS)



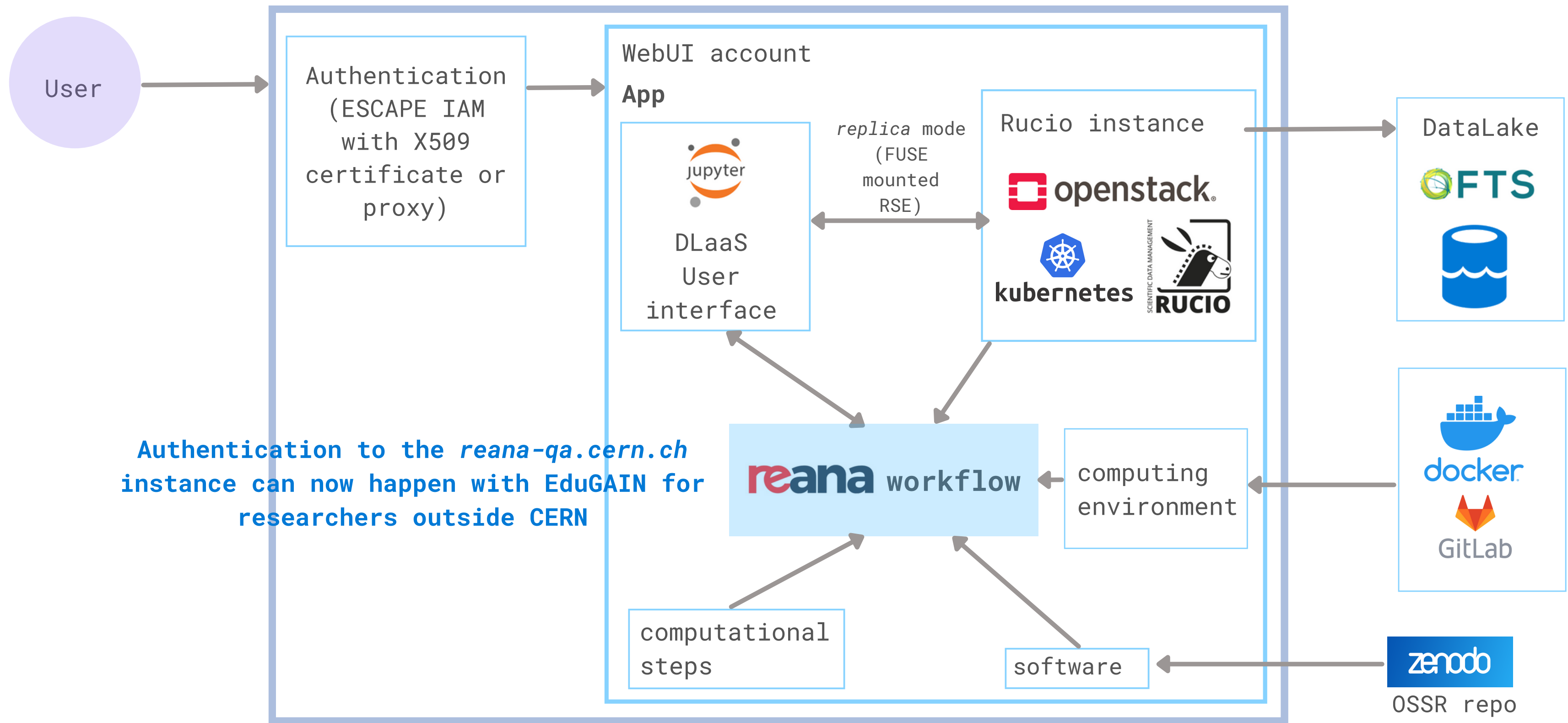
The VRE server



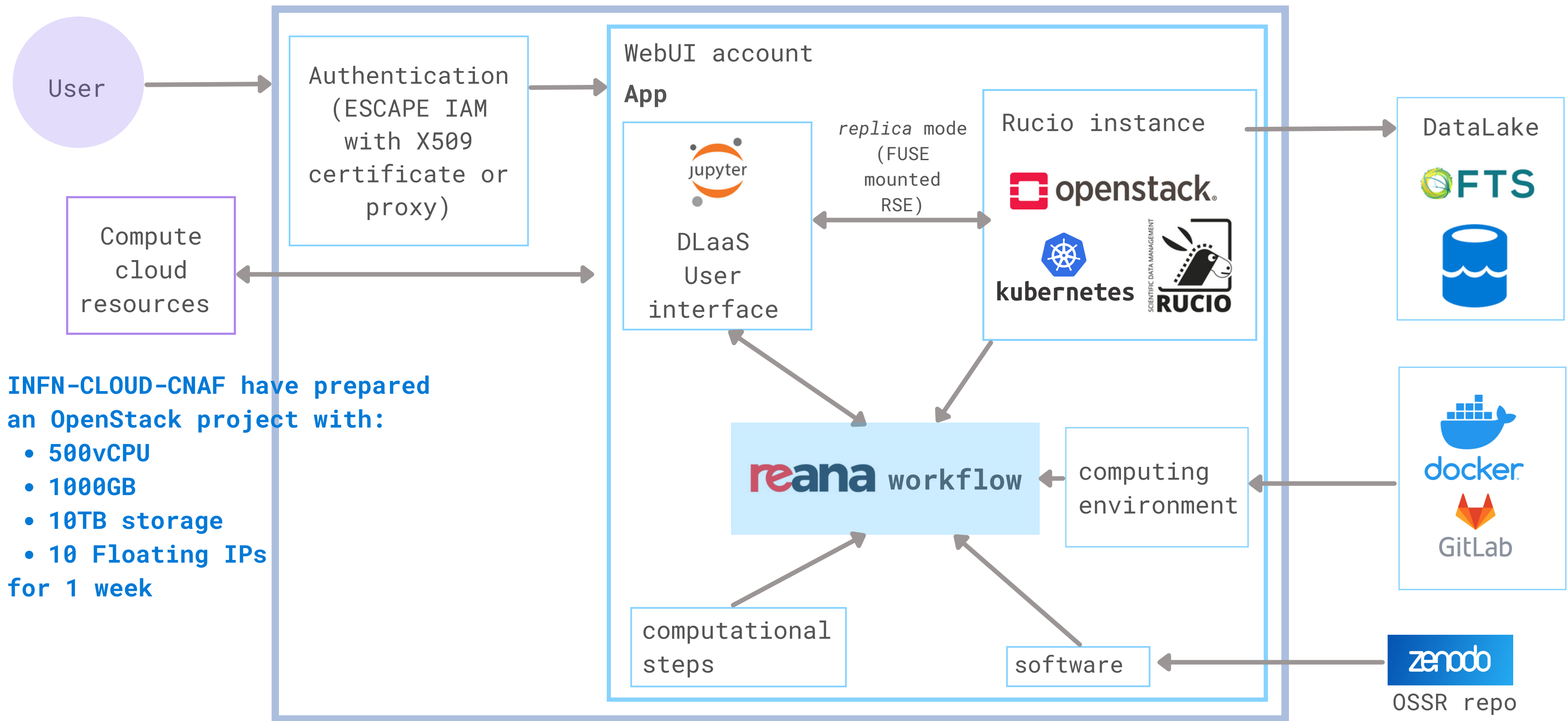
The VRE server



The VRE server



The VRE server

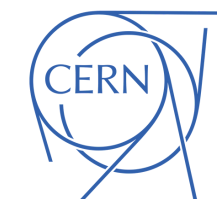


INFN-CLOUD-CNAF have prepared an OpenStack project with:

- 500vCPU
- 1000GB
- 10TB storage
- 10 Floating IPs for 1 week

AN EXAMPLE

Starting to connect the dots



ATLAS Dark Matter Reinterpretation - Dilepton Resonance

2a. Output generation

1. Import files as variables into notebook

environment is already spawned

```

[1]: axial_ee, axial_mumu, limit_intepol

[1]: (/eos/eulake_1/ATLAS_LAPP_SP/9d/f2/DMCrossSectionGraphs_axial_ee.root,
/eos/eulake_1/ATLAS_LAPP_SP/58/50/DMCrossSectionGraphs_axial_mumu.root,
/eos/eulake_1/ATLAS_LAPP_SP/23/c7/LimitInterpolator_CL95_14TeV.root)

[9]: import ROOT
import gfal2

[10]: type(axial_ee)

[13]: rucio_jupyterlab.kernels.ipython.types.SingleItemDID

[11]: def GetInteg(histo):
return histo.Integral()

def getDMCrossSection(medType):

outfilename = "DMCrossSectionGraphs_" + medType
outfile = ROOT.TFile("./output/"+outfilename+".root","recreate")
# outtuple = ROOT.TNtupleD("xsecTuple", "xsecTuple", "mass:width:massDM:xsec_truth:xsec_acc")

mgAcc = ROOT.TMultiGraph()
mgXsec = ROOT.TMultiGraph()
mgFidXsec = ROOT.TMultiGraph()
    
```

```

if finalState == "ee": leg.AddEntry(explimit, "#font[42]{Expected e^+e^-} limit")
else: leg.AddEntry(explimit, "#font[42]{Expected #mu^+#mu^-} limit", "l")
leg.AddEntry(fidXsec, "#font[42]{Vector Z'_{DM} (m_{chi}="+mDM+" TeV)}", "l")
leg.Draw()
ROOT.gPad.RedrawAxis()

fOutput.cd()
if mDM == "0.50": explimit.Write()
fidXsec.Write()
fOutput.Write()
c.SaveAs("dilepton_jared/output/Crossing_DM"+massDM+"_fs"+finalState+".pdf")

return explimit, fidXsec

def DrawAllCrossing(fOutput, finalState):

massDM = ['0p50', '1p00', '1p50', '2p00']

for mDM in massDM:
    MakeCrossing(fOutput, finalState, mDM)

if __name__ == "__main__":

ROOT.gROOT.SetBatch(True)
ROOT.gStyle.SetOptStat(False)
ROOT.gROOT.SetStyle("ATLAS")

# fOutput = ROOT.TFile(limit_intepol, "Update")
fOutput = ROOT.TFile(file_path+"LimitInterpolator_CL95_14TeV.root", "Update")

#MakeLimit(fOutput, "ee")
#MakeCrossing(fOutput, "ee", "0p50")
#MakeCrossing(fOutput, "ee", "1p00")
#MakeCrossing(fOutput, "ee", "1p50")
#DrawAllCrossing(fOutput, "ee")
#DrawAllCrossing(fOutput, "mumu")
DrawAllCrossing(fOutput, "ll")

Info in <TCanvas::Print>: pdf file dilepton_jared/output/Crossing_DM0p50_fsll.pdf ha
Info in <TCanvas::Print>: pdf file dilepton_jared/output/Crossing_DM1p00_fsll.pdf ha
Info in <TCanvas::Print>: pdf file dilepton_jared/output/Crossing_DM1p50_fsll.pdf ha
Info in <TCanvas::Print>: pdf file dilepton_jared/output/Crossing_DM2p00_fsll.pdf ha
    
```

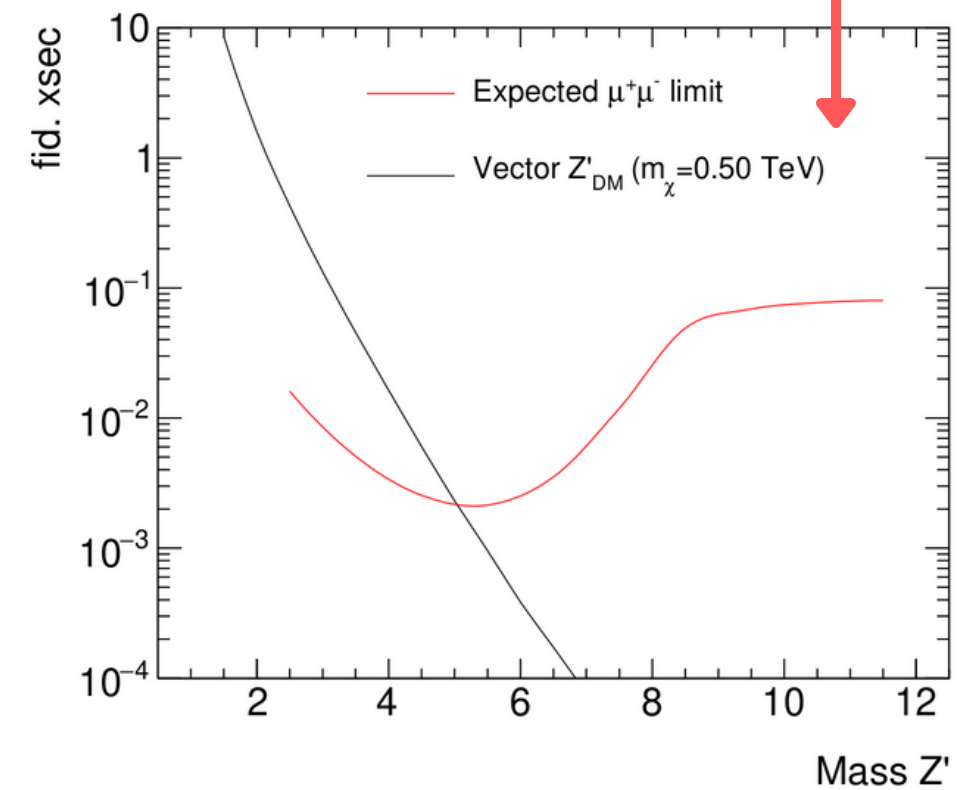
2b. REANA automates workflow execution

```

1 version: 0.8.1
2 inputs:
3   directories:
4     - python/
5     - data/
6   files:
7     - python/MakeLimit.py
8     - python/Summary.py
9     - data/DMCrossSectionGraphs_axial_massmass.root
10    - data/LimitInterpolator_CL95_14TeV.root
11 workflow:
12   type: serial
13   specification:
14     steps:
15     - name: SetLimits
16       environment: 'reanahub/reana-env-root6:6.18.04'
17       compute_backend: kubernetes
18       kubernetes_memory_limit: '9Gi'
19       commands:
20         - mkdir plots
21         - python python/MakeLimit.py
22   outputs:
23     directories:
24     - plots/
    
```

```

notebooks python README.md reana.yaml runReana.sh
jovyan@jupyter-egazzarr:~/dilepton_jared/atlas-dm-reinterpretatio
==> Verifying REANA specification file... /home/jovyan/dilepton_j
-> SUCCESS: Valid REANA specification file.
==> Verifying REANA specification parameters...
-> SUCCESS: REANA specification parameters appear valid.
==> Verifying workflow parameters and commands...
-> SUCCESS: Workflow parameters and commands appear valid.
==> Verifying dangerous workflow operations...
-> SUCCESS: Workflow operations appear valid.
==> Verifying compute backends in REANA specification file...
-> SUCCESS: Workflow compute backends appear to be valid.
SettingLimits.1
==> SUCCESS: File /python/MakeLimit.py was successfully uploaded.
==> SUCCESS: File /python/Summary.py was successfully uploaded.
    
```



Summer project

Welcome Agisilaos Kounelis !

AIM: be able to use the Data Lake data from the Reana cluster, without the restriction of having data locally.

The image shows a file explorer on the left and a terminal window on the right. The file explorer displays a directory structure with files like 'reana.yaml' and 'runReana.sh'. The terminal window shows the content of 'reana.yaml' with a red box highlighting the file paths: '- python/Summary.py' and '- data/DMCrossSectionGraphs_axial_massmass.root'. An arrow points from this box to the text 'substitute with scope:data (from Rucio DID logic)'.

```
1 version: 0.8.1
2 inputs:
3   directories:
4     - python/
5     - data/
6   files:
7     - python/Summary.py
8     - data/DMCrossSectionGraphs_axial_massmass.root
9     - data/DMCrossSectionGraphs_axial_massmass.root
10
11 workflow:
12   type: serial
13   specification:
14     steps:
15       - name: SetLimits
16         environment: 'reanahub/reana-env-root6:6.18.04'
17         compute_backend: kubernetes
18         kubernetes_memory_limit: '9Gi'
19     commands:
20       - mkdir plots
21       - python python/MakeLimit.py
22 outputs:
23   directories:
24     - plots/
```

Final remarks and next steps

1. To do:

- a. connect with IVOA (WP4) and ESAP (WP5)
- b. summer project for data fetcher/wrapper
- c. summer project for Zenodo plugin for environment import (similar to Rucio one)

1. As we want to run on EOSC-provided resources, are we in need of re-deploying the Data Lake cluster and another REANA cluster?

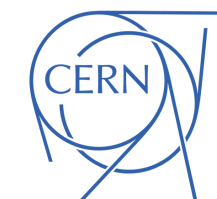
1. Federated authentication for all users and for both DL and Reana, ongoing discussion on accessing reana CERN resources

1. How to deal with scalability?



References

- [Rucio](#)
- [Rucio Escape WebUI](#)
- [K8s cluster CI/CD Helm + Flux configurations](#)
- [Cluster testing](#)
- [Grafana monitoring](#)
- [VRE webpage](#) (in progress)
- [VRE documentation](#)
- [VRE onboarding](#)
- [VRE scientific analyses + docker images for notebooks](#)





Thank you! Questions?