



Contribution ID: 15

Type: Oral

## The Virtual Research Environment: towards a complexive analysis platform

Wednesday 26 October 2022 11:30 (20 minutes)

One of the objectives of the EOSC (European Open Science Cloud) Future Project is to integrate diverse analysis workflows from Cosmology, Astrophysics and High Energy Physics in a common framework. The project's development relies on the implementation of the Virtual Research Environment (VRE), a prototype platform supporting the goals of Dark Matter and Extreme Universe Science Projects in the respect of FAIR data policies, making use of a common AAI system, and leveraging experiments data via a reliable and scalable distributed storage infrastructure for multi-science: the Data Lake. The entry point of such a platform is a jupyterhub instance sitting on top of a complex K8s infrastructure, which provides an interactive GUI interface for researchers to access and share data, as well as to run notebooks. The data access and browsability is enabled through API calls to the high level data management and storage orchestration software (Rucio).

The cluster's functionality, currently allowing data injection replication, storage and deletion, is being expanded to include a software repository plug-in enabling researchers to directly select computational environments from Docker images and to host a re-analysis platform (REANA) supporting various distributed computing backends (K8s, HTCondor, Slurm), which allows scientists to spawn and interact with complete re-analysis workflows.

The goal of the VRE project, bringing together data and software access, workflow reproducibility and enhanced user interface, is to facilitate scientific collaboration, ultimately accelerating research in various fields.

### Significance

The VRE will first and foremost provide an easy-to use prototype analysis platform based on some of the most commonly used DevOps technologies (K8s, Helm, Flux, GitLab, DB on-demand), with the nuance of hosting workflows spanning from the field of particle physics to astrophysics.

The novelty of the infrastructure will be its common AAI framework to authenticate with both federated storage services and computing infrastructure, gaining access to the data management software, the software repository and the computational environment necessary for analysis reproduction at the same time.

While similar work has been ongoing in single isolated institutes –at CERN, for example –, the VRE aims at being completely open source, easily reproducible on different clusters, and easily accessible by anyone having an account; the target audience are not only HEP sciences, but also smaller experiments who would hugely benefit from the provisioning of shared computing resources.

### References

<https://indico.in2p3.fr/event/26454/>

<https://indico.cern.ch/event/1151054/>

### Experiment context, if any

Various experiments are currently using the VRE platform and providing feedback to develop it further – making it easier to use, enhancing the documentation, improving its deployment –and others are continuously onboarding the project. The postdocs testing the infrastructure are involved in experiments such as ATLAS, Km3Net, Fermi-LAT, EGO and LOFAR.

**Primary authors:** GAZZARRINI, Elena (CERN); VENDRELL MOYA, Alba (CERN); DI MARIA, Riccardo (CERN); DONA, Rizart (CERN); ESPINAL, Xavier (CERN); GARCIA GARCIA, Enrique

**Presenter:** GAZZARRINI, Elena (CERN)

**Session Classification:** Track 1: Computing Technology for Physics Research

**Track Classification:** Track 1: Computing Technology for Physics Research