Contribution ID: **22** Type: **Presentation (15' + 5' for questions)**

The interTwin Digital Twin Engine Data Lake

Tuesday, 29 October 2024 14:40 (20 minutes)

The Horizon Europe interTwin project is developing a highly generic Digital Twin Engine (DTE) to support interdisciplinary Digital Twins(DT). The project brings together infrastructure providers, technology providers and DT use cases from High Energy and AstroParticle Physics, Radio Astronomy, Climate Research and Environmental Monitoring. This group of experts enables the co-design of both the DTE Blueprint Architecture and the prototype platform; not only benefiting end users like scientists and policymakers but also DT developers. It achieves this by significantly simplifying the process of creating and managing complex Digital Twins workflows.
In our presentation, we'll focus on the design and implementation of the interTwin DataLake, which is based on the ESCAPE Data Lake concept and the extensions and integrations done and ongoing in the project to accommodate from one side the heterogeneous resource providers ranging from HTC, HPC to Cloud and the Requirements from the User communities.
The DataLake services have been integrated with EGI Check-in, including Rucio, FTS and storage technologies deployed at the sites already part of the testbed (VEGA EuroHPC, EODC, DESY, INFN), and under integration (PSNC, Julich, KBFI)
New developments have been performed to ease the integration at sites. One of the new developments is Teapot, an application that provides multi-tenant WebDAV support. Teapot is built on StoRM-WebDAV and includes a manager that accepts requests, authenticates users, maps them to local usernames, and launches a dedicated StoRM-WebDAV server where the manager then forwards the request to.
And finally the DataLake is also available for exploitation via JupyterHub thanks to the Rucio jupyterLab Plugin developed by CERN in the ESCAPE project and further enhanced in interTwin.
Details of integration by some of the DTs and upper architecture layers will also be presented.

Primary author: VRBANEC, Dijana (DESY)

Co-authors: MANZI, Andrea (EGI Foundation); FILIPCIC, Andrej (JSI); CIANGOTTINI, Diego (INFN); GARCIA GARCIA, Enrique (CERN); MILLAR, Paul (DESY); PRICA, Teo (IZUM); SPIGA, daniele (INFN)

Presenter: VRBANEC, Dijana (DESY)

Session Classification: IBERGRID

Track Classification: Development, implementation and operation of Data Lakes