Contribution ID: **101**   Type: **Short Talk**

Enabling interoperable data and application services in a federated ScienceMesh

Tuesday, 18 May 2021 11:16 (13 minutes)

In recent years, cloud sync & share storage services, provided by academic and research institutions, have become a daily workplace environment for many local user groups in the High Energy Physics (HEP) community. These, however, are primarily disconnected and deployed in isolation from one another, even though new technologies have been developed and integrated to further increase the value of data. The EU-funded CS3MESH4EOSC project is connecting locally and individually provided sync and share services, and scaling them up to the European level and beyond. It aims to deliver the ScienceMesh service, an interoperable platform to easily sync and share data across institutions and extend functionalities by connecting to other research services using streamlined sets of interoperable protocols, APIs and deployment methodologies. This supports multiple distributed application workflows: data science environments, collaborative editing and data transfer services.

In this paper, we present the architecture of ScienceMesh and the technical design of its reference implementation, a platform that allows organizations to join the federated service infrastructure easily and to access application services out-of-the-box. We discuss the challenges faced during the process, which include diversity of sync & share platforms (Nextcloud, Owncloud, Seafile and others), absence of global user identities and user discovery, lack of interoperable protocols and APIs, and access control and protection of data endpoints. We present the rationale for the design decisions adopted to tackle these challenges and describe our deployment architecture based on Kubernetes, which enabled us to utilize monitoring and tracing functionalities. We conclude by reporting on the early user experience with ScienceMesh.

Primary authors: ARORA, Ishank (CERN); ALFAGEME SAINZ, Samuel (CERN); GONZALEZ LABRADOR, Hugo (CERN); MOSCICKI, Jakub (CERN)

Co-author: FERREIRA, Pedro (CERN)

Presenter: ARORA, Ishank (CERN)

Session Classification: Storage

Track Classification: Distributed Computing, Data Management and Facilities