

The OSG Data Federation

Thursday, 27 August 2020 11:30 (20 minutes)

The Open Science Grid operates a global Data Federation for all of Open Science. As of Fall 2019, this includes 6 data origins for a variety of scientific domains and communities, and 12 caches. The caches are deployed inside the network backbone, and at or near various endpoints, i.e. compute clusters. A combination of technologies, including CVMFS for file metadata and authentication, and XRootd for data access of large data volumes, provide scientific communities like the International Gravitational-Wave Observatory Network, DUNE, DES, Minerva, as well as individual bioinformatics researchers a platform akin to a content delivery network that looks like a read-only global file system to the end user scientist. In this talk, we describe this global infrastructure, how it is used, and how we use service containers and K8S as container orchestration system to support this global platform.

Primary author: Prof. WUERTHWEIN, Frank (UCSD/SDSC)

Presenter: Prof. WUERTHWEIN, Frank (UCSD/SDSC)

Session Classification: Data Management Session

Track Classification: Data Management & Big Data