Contribution ID: **257**                                                   Type: **Oral**

# SWAN: a Service for Web-Based Data Analysis in the Cloud

*Wednesday 12 October 2016 11:30 (15 minutes)*

SWAN is a novel service to perform interactive data analysis in the cloud. SWAN allows users to write and run their data analyses with only a web browser, leveraging the widely-adopted Jupyter notebook interface. The user code, executions and data live entirely in the cloud. SWAN makes it easier to produce and share results and scientific code, access scientific software, produce tutorials and demonstrations as well as preserve analyses. Furthermore, it is also a powerful tool for non-scientific data analytics.

The SWAN backend combines state-of-the-art software technologies, like Docker containers, with a set of existing IT services such as user authentication, virtual computing infrastructure, mass storage, file synchronisation and sharing, specialised clusters and batch systems. In this contribution, the architecture of the service and its integration with the aforementioned CERN services is described. SWAN acts as a "federator of services" and the reasons why this feature boosts the existing CERN IT infrastructure are reviewed.

Furthermore, the main characteristics of SWAN are compared to similar products offered by commercial and free providers. Use-cases extracted from workflows at CERN are outlined. Finally, the experience and feedback acquired during the first months of its operation are discussed.

## Tertiary Keyword (Optional)

Data processing workflows and frameworks/pipelines

## Secondary Keyword (Optional)

Analysis tools and techniques

## Primary Keyword (Mandatory)

Cloud technologies

**Primary authors:**   PIPARO, Danilo (CERN);  TEJEDOR SAAVEDRA, Enric (CERN)

**Co-authors:**   MOSCICKI, Jakub (CERN);  MASCETTI, Luca (CERN);  LAMANNA, Massimo (CERN);  MATO VILA, Pere (CERN)

**Presenter:**   TEJEDOR SAAVEDRA, Enric (CERN)

**Session Classification:**  Track 6: Infrastructures

**Track Classification:**  Track 6: Infrastructures