Contribution ID: **282**   Type: **Talk**

Data Movement Model for the Vera C. Rubin Observatory

Wednesday 23 October 2024 14:42 (18 minutes)

The set of sky images recorded nightly by the camera mounted on the telescope of the Vera C. Rubin Observatory will be processed in facilities located on three continents. Data acquisition will happen in Cerro Pachón in the Andes mountains in Chile where the observatory is located. A first copy of the raw image data set is stored at the summit site of the observatory and immediately transferred through dedicated network links to the archive site and US Data Facility hosted at SLAC National Laboratory in California, USA. After an embargo period of a few days, the full image set is copied to the UK and French Data Facilities where a third copy is located.

During its 10 years in operation starting late 2025, annual processing campaigns across all images taken to date will be jointly performed by the three facilities, involving sophisticated algorithms to extract the physical properties of the celestial objects and producing science-ready images and catalogs. Data products resulting from the processing campaigns at each facility will be sent to SLAC and combined to create a consistent Data Release which is served to the scientific community for its science studies via Data Access Centers in the US and Chile and Independent Data Access Centers elsewhere.

In this contribution we present an overall view of how we leverage the tools selected for managing the movement of data among the Rubin processing and serving facilities, including Rucio and FTS3. We will also present the tools we developed to integrate Rucio's data model and Rubin's Data Butler, the software abstraction layer that mediates all access to storage by the pipeline tasks which implement the science algorithms.

Primary authors: HANUSHEVSKY, Andrew (Stanford University/SLAC); HERNANDEZ, Fabio (IN2P3 / CNRS computing centre); BECKETT, George; LIM, Kian-Tat (SLAC National Accelerator Laboratory, USA); LOVE, Peter (Lancaster University (GB)); PIETROWICZ, Stephen R. (National Center for Supercomputing Applications, USA); JENNESS, Tim (Vera C. Rubin Observatory, USA); NOBLE, Timothy John (Science and Technology Facilities Council STFC (GB)); YANG, Wei (SLAC National Accelerator Laboratory (US))

Presenter: HERNANDEZ, Fabio (IN2P3 / CNRS computing centre)

Session Classification: Parallel (Track 1)

Track Classification: Track 1 - Data and Metadata Organization, Management and Access