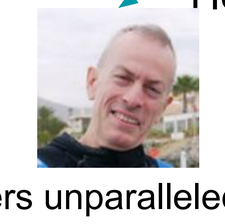
Rubin's Hybrid On Premises-Cloud Data Access Center

William O'Mullane[1], Yusra AlSayyad[2], James Chiang[3], Richard Dubois[4], Frossie Economou[1], Fabio Hernandez[5], Flora Huang[6], Tim Jenness[1], Kian-Tat Lim[4], Yee-Ting Li[4], Fritz Mueller[4], Dan Speck[7], Stephen Pietrowicz[8], and Wei Yang[4]

1. Vera C. Rubin Observatory 2. Princeton University, 3. Kavli Institute for Particle Astrophysics and Cosmology
4. SLAC National Accelerator Laboratory 5. CNRS CC-IN2P3, 6.Google LLC, 7.Burwood Group, 8. NCSA

Poster:13101-86



Cloud computing offers unparalleled flexibility, a constantly increasing set of "Infrastructure As A Service" capabilities, resource elasticity and security isolation. One of the most significant barriers in astronomy to wholesale adoption of cloud infrastructures is the cost for hot storage of large datasets - particularly for Rubin, a Big Data project sized at 0.5 Exabytes (500 Petabytes) over the duration of its 10-year mission. We are planning to reconcile this with a "hybrid" model where user-facing services are deployed on Google Cloud with the majority of data holdings residing in our on-premises Data Facility at SLAC. We discuss the opportunities, status, risks, and technical challenges of this approach.

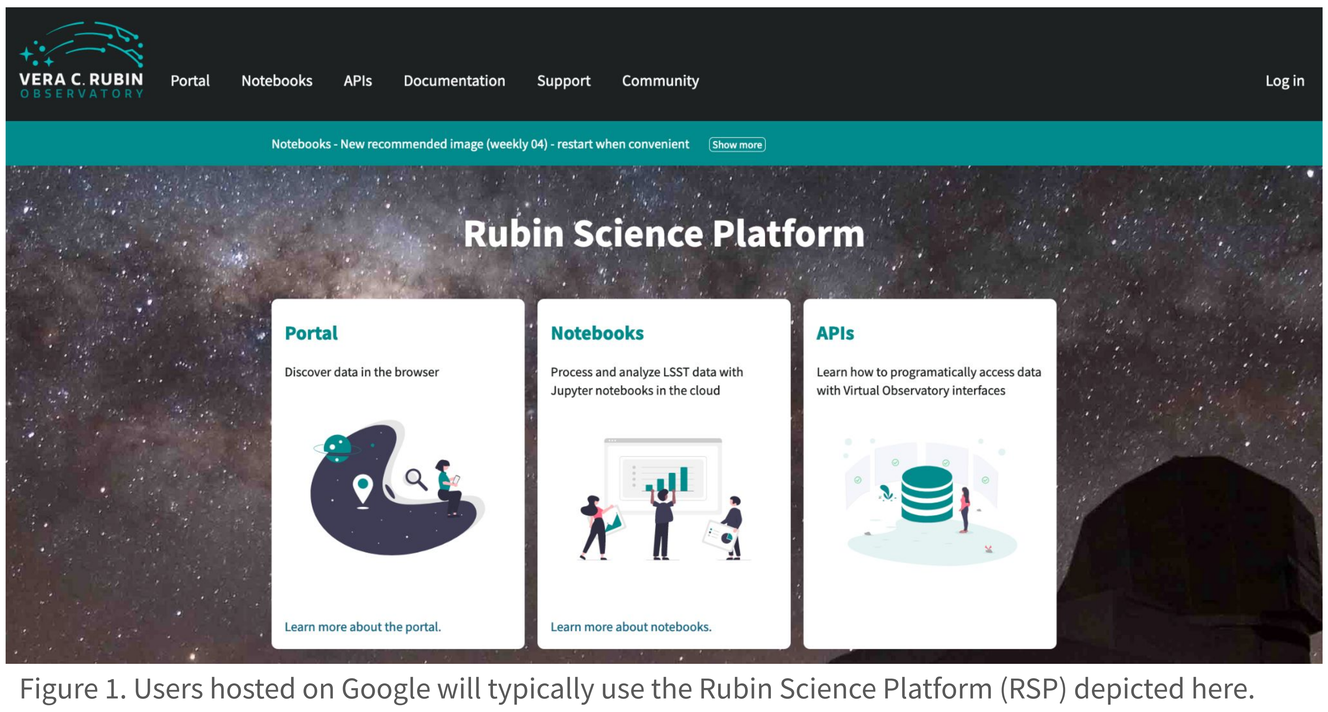


Figure 1. Users hosted on Google will typically use the Rubin Science Platform (RSP) depicted here.

Google Architecture

We built cloud ready (see below) - users typically see the RSP (Figure 1 above)

Data is mainly at USDF accessed via services like TAP and S3 (see poster 13101-129)

Deployment across Rubin facilities

In house developed Safir and Phalanx (Internal Developer Platform) atop

- Helm to describe Kubernetes applications
- ArcoCD to continuously deploy the application to one or more envs (inc. Google)
- Kubernetes to provide similar interface to hardware in multiple locations

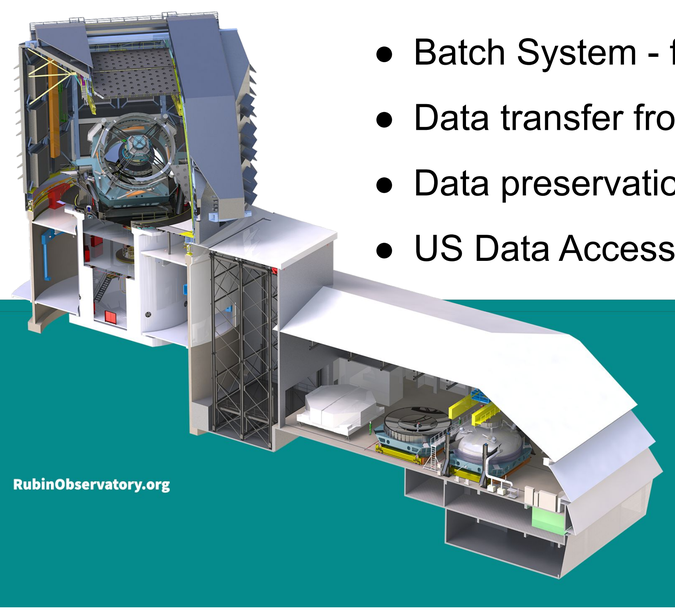
Continuous integration with GitHub actions - Phalanx can deploy specific tag

Science Pipelines nightly build with Jenkins

- Weekly and stable releases packaged as a conda* environment as well as an Apptainer for CERN's CernVM-FS Hosted by the French data facility

System Requirements

- Networking 100Gbit/s, path redundant, capacity from Chile
- Prompt processing near-real-time , Alert Production (<120s)
  - prompt data products
  - alerts corresponding to changes in the sky
  - Quality control metrics for the images also need to be generated and made available to staff. Prompt products, including both images and catalogs, and alerts are stored for retrieval by science users and staff.



- Batch System - for DRP and users
- Data transfer from Chile and to Europe
- Data preservation (alternative backup in France
- US Data Access Center

USDF Architecture (see Figure 2 below)

USDF on-prem includes data production services: prompt processing, serving alerts to the community and annual Data Release Processing. It holds the archive for all data, provides Qserv object catalog, access to image data, batch cycles for cloud-based science users. It is the home for developers and staff.

- Hosted by the SLAC Shared Scientific Data Facility (S3DF) inthe Stanford Research Computing Facility (SRCF)
- Weka filesystem for high throughput - SSDs over regular disk
  - Ceph - presents posix - tape backup
- Slurm cluster, primarily (AMD) milan processors 128 core 512GB
- combined leased-line, ESNet-supported network with routing optimized via overlay from Chile
  - ESNet to UK and FR Data facilities

Non-user-facing services on Prem:

- Prompt Processing/Alerts production - 2-minute latency
  - Knative on top of Kubernetes

Databases minimally use kubernetes for deployment:

- Qserv database - large data volume, custom built
- Engineering and Facilities Database (EFD) - influx high availability
- Casandra - back end for Alert Products Database.

Prometheus for monitoring

- native support for Kubernetes metrics
- Strimzi and the Cloud Native Postgres (CNPG) provide metrics
- Loki from Grafana Labs to capture logs from Kubernetes
  - Logs from pods and from application level logs
- Grafana for creating dashboards, visual metrics
  - generating alarms for given thresholds

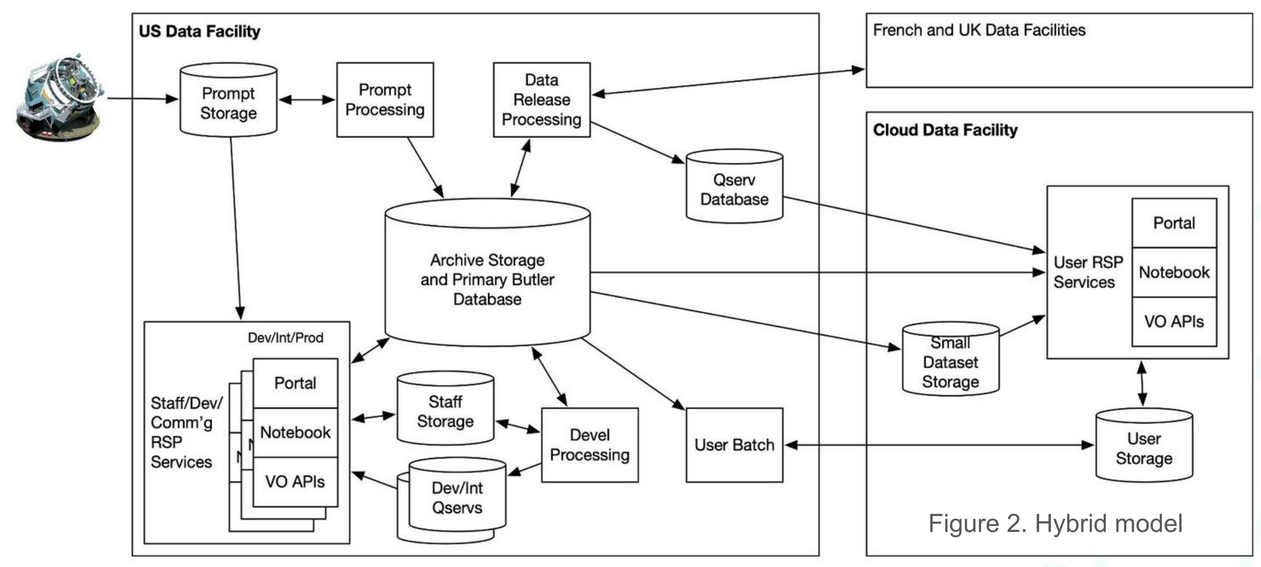


Figure 2. Hybrid model

Rubin Observatory is a joint initiative of the National Science Foundation (NSF) and the Department of Energy (DOE). Its primary mission is to carry out the Legacy Survey of Space and Time, providing an unprecedented data set for scientific research supported by both agencies. Rubin is operated jointly by NSF's NOIRLab and SLAC National Accelerator Laboratory (SLAC). NOIRLab is managed for NSF by the Association of Universities for Research in Astronomy (AURA) and SLAC is operated for DOE by Stanford University.

RubinObservatory.org

QR code for paper