VERA C. RUBIN OBSERVATORY

USDF Update

Richard Dubois

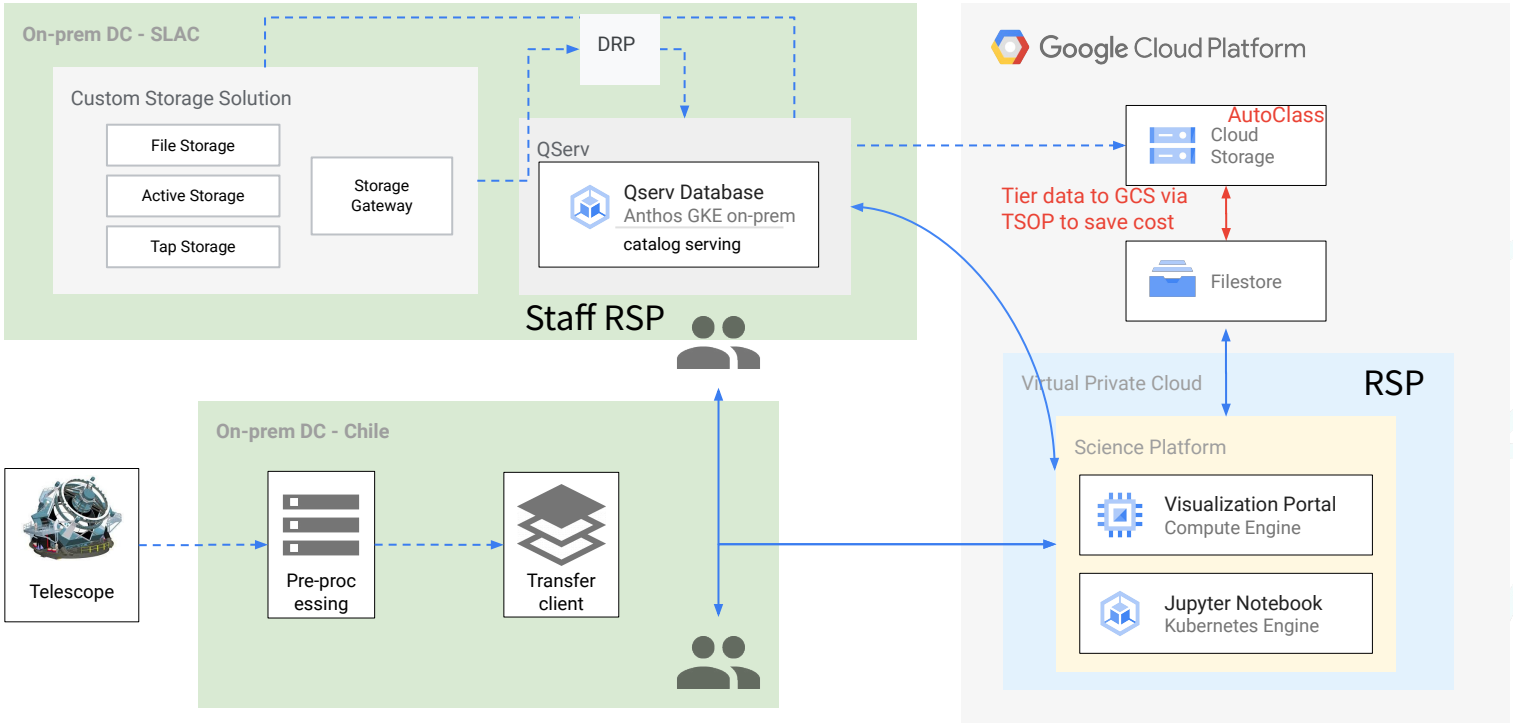
8 June 2022

NSF NOIR Lab AURA SLAC U.S. DEPARTMENT OF ENERGY

Planning

- DMTN-189 for Data Facility scope
- RTN-021 for DF transition plan (George Beckett has an update going)
- NCSA “shuts down” on ~~October 1~~ August 15 (gulp)
  - SLAC becomes primary support for ComCam commissioning this summer
  - Driven by need to build up a secure enclave for arriving data, including ComCam
- Discussed weekly (Tues 11am PT), also #ops-usdf-arch
  - Planning focused meetings on infrastructure, k8s and security in next 3 weeks
- Three workshops so far, last one was April; next planned in 2 weeks
- DOE signed off to proceed with hybrid model

Hybrid Model



Hybrid model: Data at SLAC but users on the Cloud.

Allows:

- Separation of security concerns
- Burst response
- Reduced risk

(see DMTN-209)

RSP = Rubin Science Platform

Cloud-SLAC Division of Scope

Cloud

- Science users, with access provided by the Rubin Science Platform
- Provide personal storage/CPU + cloud access to coadds
  - Effective 500 cores added each year + 2 PB → 10 PB storage

SLAC

- Prompt and DRP processing
- Qserv catalogue server
- Storage archive for all data
- Serving alerts to the community
- Home for developers and staff (and commissioners)
- What about user batch? Work submitted from the cloud and run at SLAC… ?

Hybrid Cost Mitigations

- We have worked with Google to pursue more aggressive ways to trim their costs
- Take advantage of only paying for what you use:
  - Assume that only 10% of users are concurrent - inactive data is rolled off to a cheaper tier. Very infrequent user data is rolled off to SLAC (or tossed).
  - An average user only uses a fraction of the per-person quota
  - Provide DRP Coadds in the cloud, assuming that it the most attractive image product; other products could be drawn from SLAC. Added in 10% of parquet files.
  - The user and DRP access savings were about $1.2M each by yr 5.
- There is a risk that this Hybrid model is too aggressive. Experience will tell if we to need to be less aggressive and this would have an associated cost.
- Next up: write up SoW and req for SLAC Contracts to review. They already have our sole-source justification to consider.

USDF Transition Status

- Transition from NCSA
  - Copying 4 PB of data from NCSA now; past halfway… hopefully complete in June
  - Identified active NCSA account holders needing SLAC accounts
    - Invited Pipelines, SQRE and AOS teams to get accounts so far
    - A list of in-kind commissioners from Robert/Keith is in prep (33 names on it currently)
    - Uptake has been slow - please remind your teams to sign up!
- If you already have a SLAC unix account
  - Need to be added to rubin_users group
  - If you don’t have a windows account, you can create one - currently that account is used for authentication to the cluster
    - https://sdf.slac.stanford.edu/public/doc/#/accounts-and-access

Resources Currently Available

- Dev guide being updated: https://developer.lsst.io/v/PREOPS-892/usdf/lsst-login.html
- Cluster doc: https://sdf.slac.stanford.edu/
- Questions can be posed in #ops-usdf slack channel
- Devl node: rubin-devl.slac.stanford.edu
  - Load balancer for pool - currently only one denizen - 128 core/512 GB RAM; accessible by jump from SDF login node
- Stack available from cvmfs
- Slurm batch system - we have contributed 7 Romes, and have a rubin partition
- Living on borrowed POSIX (Lustre) for the moment for the NCSA copy
  - Can use /sdf/scratch/<your_account> while proper home/group space is commissioned

S3DF status

- SLAC Shared Scientific Data Facility
- Moving to new core infrastructure - being deployed now
  - Weka filesystem for home/group space - flash
  - Object store expected to be ActiveScale - our 4 PB will be set up with it this month
    - Backup is ceph
  - k8s service will become production - we have 16 k8s nodes in place
    - Butler repo deployed on k8s
    - Permissions secret held in vault
    - PanDA and Rucio deployments via k8s in progress
    - Developer RSP in place, with authentication to SLAC accounts

Hardware being ordered

- Req working its way through SLAC
- 5000 batch cores
- 11 PB storage
- 15 Qserv nodes
- 32 k8s nodes
- 3 DTNs
- Pondering FY23 needs - given slip in survey start… (FY23 budget supports FY24 activities)

Still up in the air

- Best option for developer batch (bps)?
  - PanDA, Parsl, HTCondor?
- Shared stack
  - Currently providing weeklies, releases via cvmfs - includes singularity containers
    - SLAC is a singularity shop - will need to provide education about its use
  - No Jenkins at SLAC yet
- Getting summit data to SLAC
  - LHN overlay not in place (1 GB/s effective rate limit)
  - Final transfer mechanism not ready yet (object store posts)
- HSC reprocessing
  - Hsin-Fang has been kicking the PanDA tires at SLAC
  - New Brock-like hire in progress
- Secure rack
  - Funding not in place for routers, object store hardware
  - Presumably we'll start with a not-secure rack