VERA C. RUBIN OBSERVATORY

US Data Facility Overview

Adam Bolton, USDF Lead

Data Facilities F2F meeting
Tue 11 Feb 2025
CC-IN2P3, Lyon, France

NSF NOIRLab AURA SLAC U.S. DEPARTMENT OF ENERGY

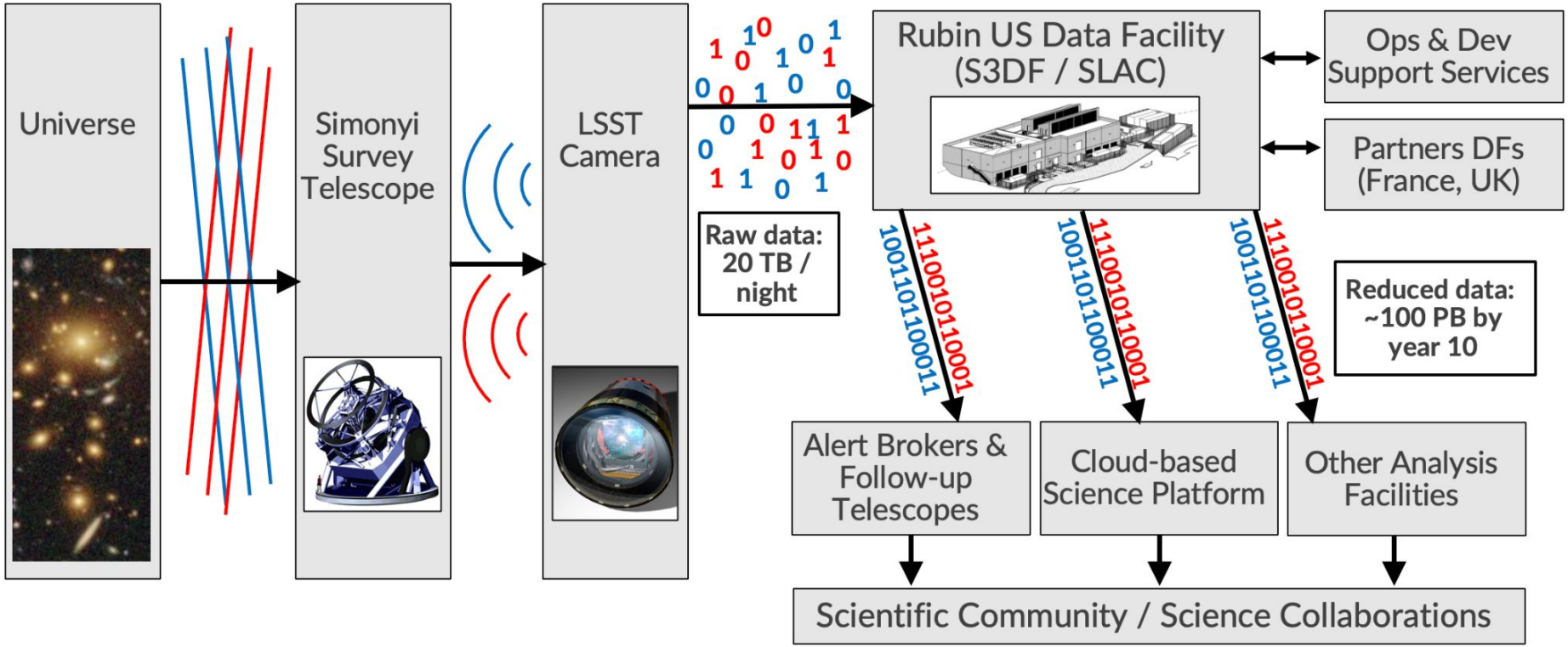
USDF meaning(s) within Rubin Observatory

- Most obviously, “USDF” denotes the USDF infrastructure itself
  - Primary hub for operation of Rubin Data Management systems & data flows
  - Hosted by SLAC National Accelerator Laboratory
  - Implemented within SLAC Shared Science Data Facility (S3DF)
- Also denotes team within the Data Management department of Rubin Operations
  - Supports Construction and Commissioning, but not part of the “capital-P” Project
  - Has collaborative interfaces with essentially all other DM teams
- Within the Rubin Operations Work Breakdown Structure (WBS), “USDF Team” encompasses all Data Facilities activities (including France and UK)
- Also denotes all US-based Data Facilities efforts
- Also denotes the USDF team at SLAC in particular

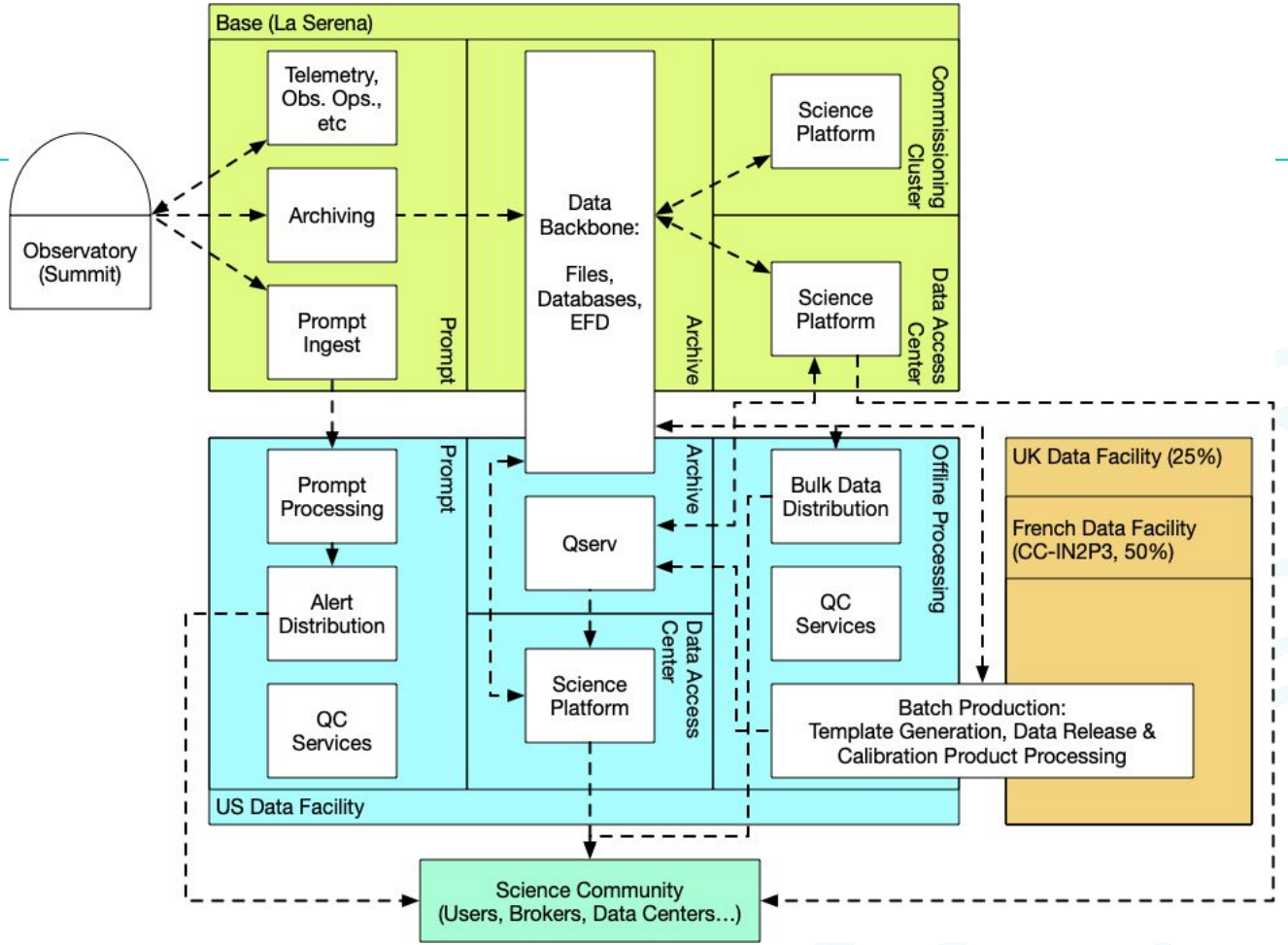
USDF within Rubin Observatory (functional)

- **Stewardship and curation** of Rubin Observatory's raw data and generated data products, including backup and disaster recovery systems
- **Computing infrastructure and environment** for prompt processing, alert generation, and data release processing, including framework for distributed multi-site processing
- **Movement and mirroring of data** between USDF and other processing and analysis sites, including FrDF, UKDF, and IDACs
- **Platform to support diverse set of Rubin scientific applications, services, and databases**, many of which are developed and operated by other Rubin teams
- **Operation of the hybrid US DAC**, in collaboration with SQuaRE, Qserv, Middleware, others
- **Coordination with the Summit facility** to ensure reliable and performant data transfer
- **Implementation of USDF aspects of Rubin Data Security policy** (DMTN-199)
- **Liaison and support for Data Facility stakeholders** within Rubin and the broader community

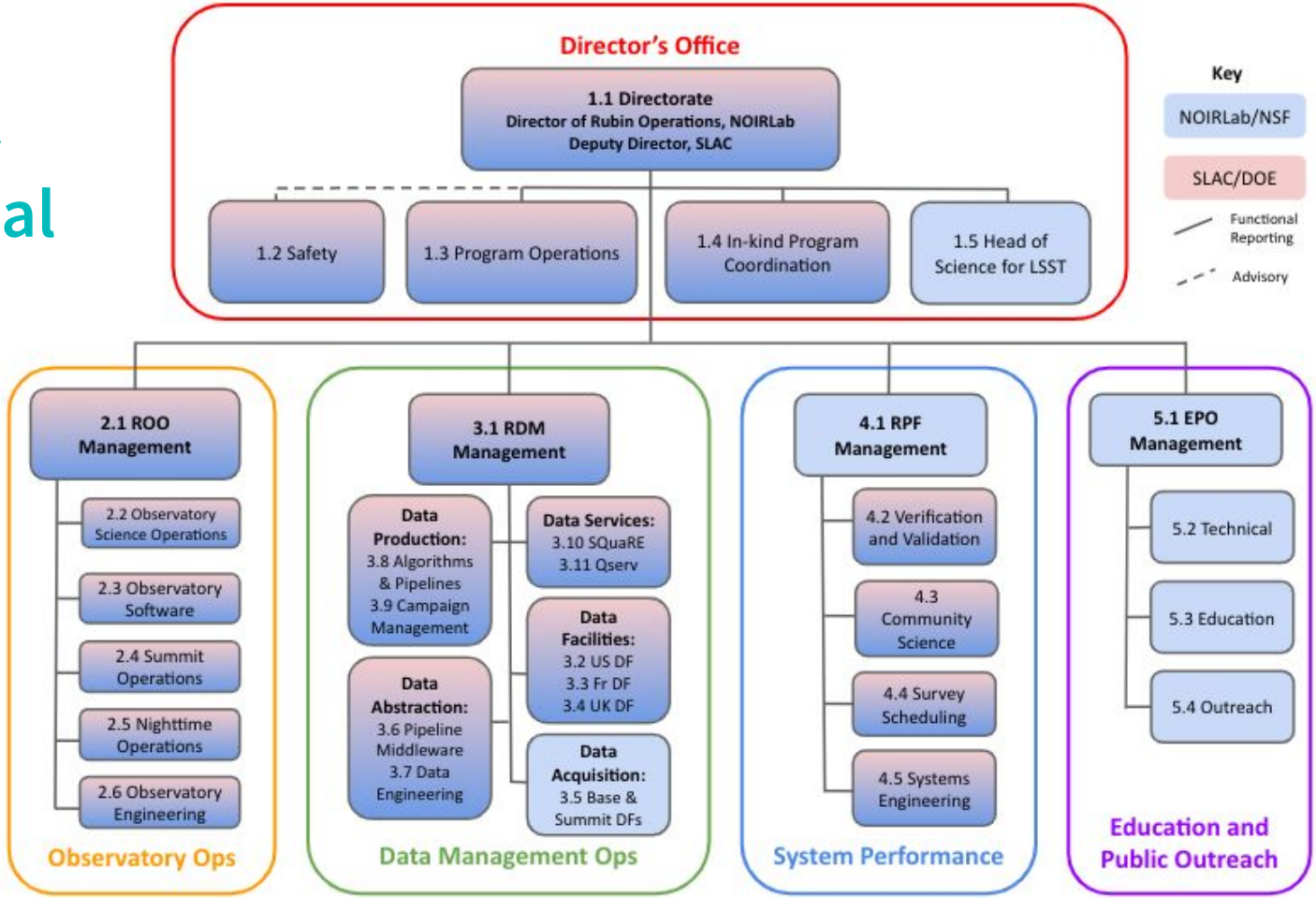
Experimental context view



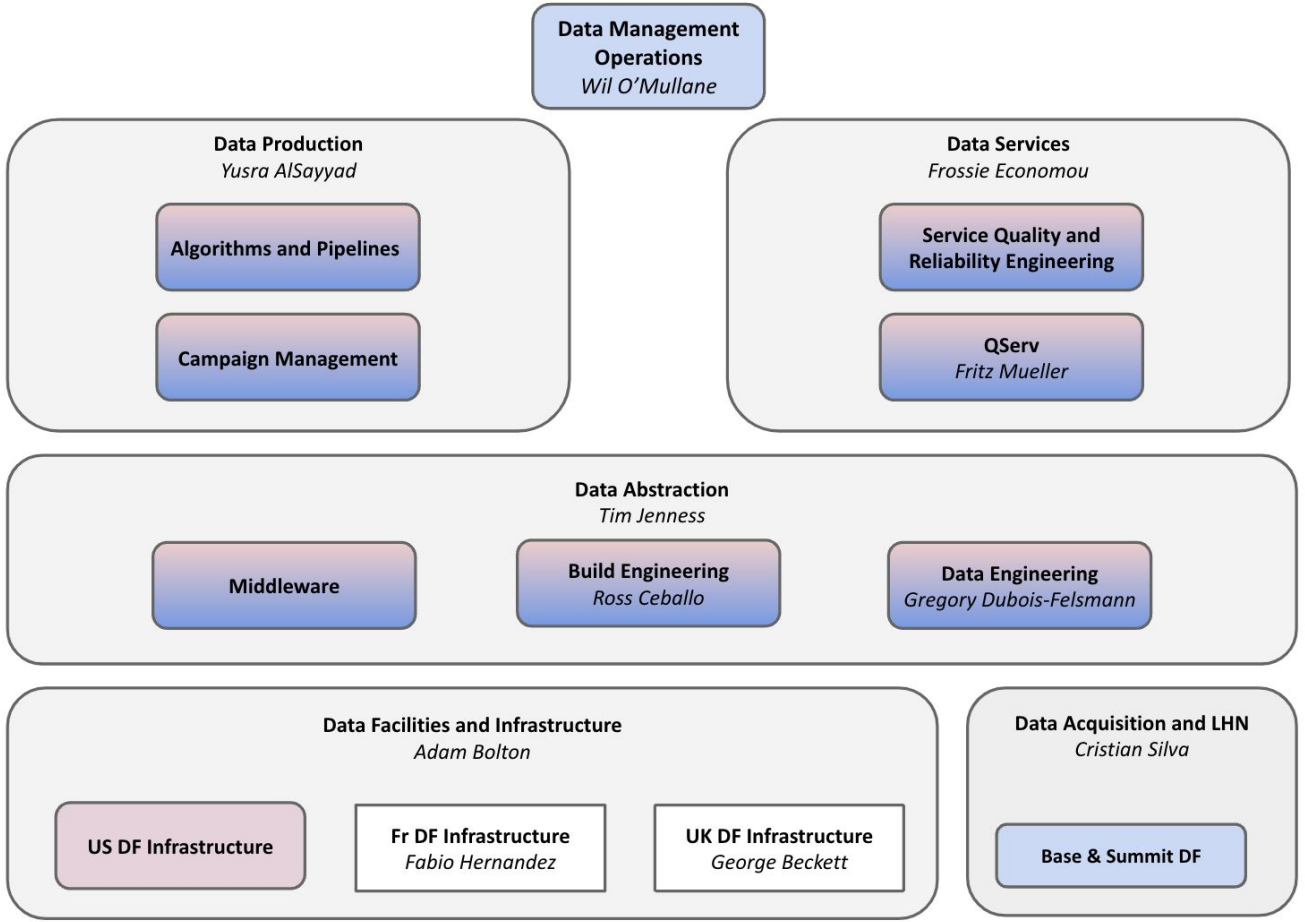
Technical view



Functional org view (Rubin)

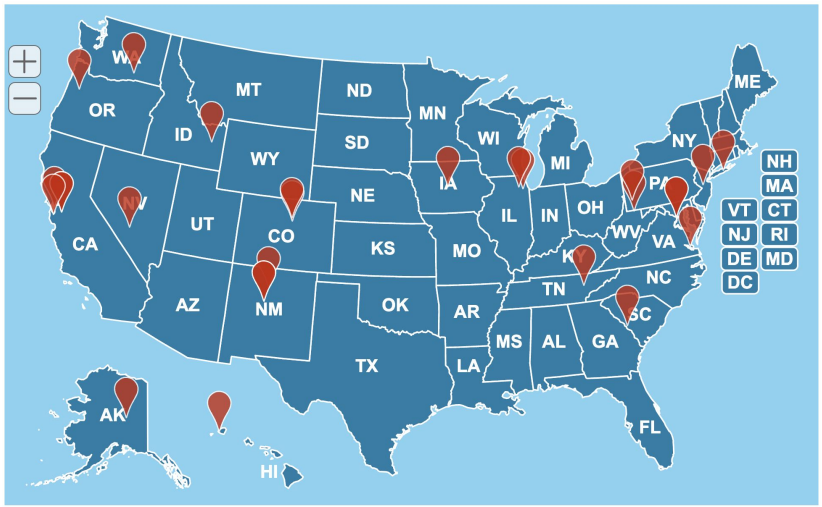


Functional org view (DM)



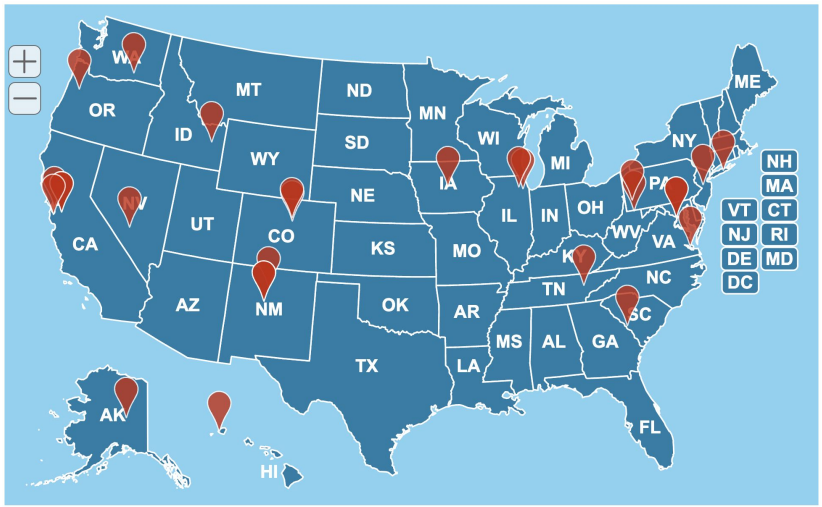
SLAC National Accelerator Laboratory

- One of 17 National Laboratories operated by the US Department of Energy (DOE)
  - More specifically, one of 10 DOE Office of Science Labs
- Operated by Stanford University under contract with DOE
- Located in Menlo Park, California, USA



SLAC National Accelerator Laboratory

- Lead DOE Lab for Rubin Observatory
  - Construction of LSSTCam, Operation of US Data Facility
  - Host lab for Dark Energy Science Collaboration (DESC)
- Construction partner with NSF-Funded / AURA-operated Rubin Observatory Project
- Operations partner with NSF NOIRLab (also AURA-operated)



USDF staff (SLAC Fundamental Physics)

- Adam Bolton, USDF Lead
  - Astronomer/physicist, past lives in NOIRLab, DESI, SDSS, University of Utah
- K-T Lim, USDF Technical Lead
  - Broad and deep USDF technical expertise; liaison to S3DF and to other DM teams
- Jim Chiang
  - USDF Deputy Lead; connection to DESC, Pipelines, Campaign Management
- Sierra Villarreal
  - Science operations, user account support, IDAC coordination
- Andy Hanushevsky
  - Data movement
- “USDF-adjacent” SLAC staff
  - Hsin-Fang Chiang, Orion Eiger (Campaign Management)
  - Fritz Mueller, Igor Gaponenko, Andy Salnikov (Qserv)

USDF staff (other institutions)

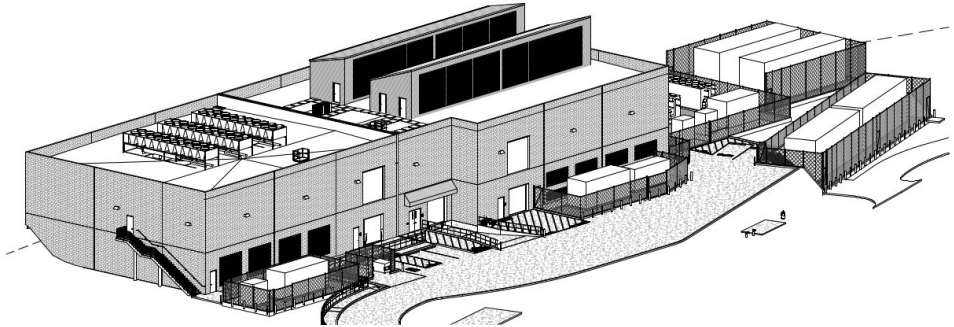
- Brandon White, Data Movement (Fermilab)
- Yuyi Guo, Data Movement (Fermilab)
- Wen Guan, Distributed Processing (Brookhaven)
- Zhaoyu Yang, Distributed Processing (Brookhaven)
- Edward Karavakis, Distributed Processing (Brookhaven)
- Greg Daues, Workflow Management Tools (NCSA)
- Dan Speck, Applications and Databases (Burwood)

USDF staff (SLAC Technology & Innovation Dir.)

- Jay Srinivasan, Scientific Computing Systems Division Director
- Yemi Adesanya, Technology Engagement Specialist
- Wei Yang, Rubin distributed data processing lead
- Yee-Ting Li, S3DF Applications and User Support Department Head
  - Heinrich Reinking, Applications and Infrastructure (AURA)
  - Peter Vaucher, Applications Support
  - Ramya Eranna, Applications Support
  - Jhonatan Amado, monitoring and databases (Fermilab)
  - DBA (TBH)
- Omar Quijano, S3DF Infrastructure & System Support Department Head
  - Lance Nakata, Storage Systems
  - Vladimir Sigunov, Storage Systems (Weka)
  - Riccardo Veraldi, Networking
  - Simon Elmir, Operating Systems
- Gregg Thayer, S3DF Data Management Department Head

SLAC Shared Science Data Facility (S3DF)

- SLAC's new shared high-throughput experimental computing infrastructure
  - Consolidating historically siloed scientific computing at SLAC
- Hosted within Stanford Research Computing Facility (SRCF)
  - Split across two adjoining facilities (SRCF-I and SRCF-II)
- Key mission area: critical, data-heavy, scientific computing workflows
  - Supports other large experiments in addition to Rubin (LCLS, UED, CryoEM, SSRL)
- Operated by Scientific Computing (SCS) Division of SLAC's Technology and Innovation Directorate (TID)



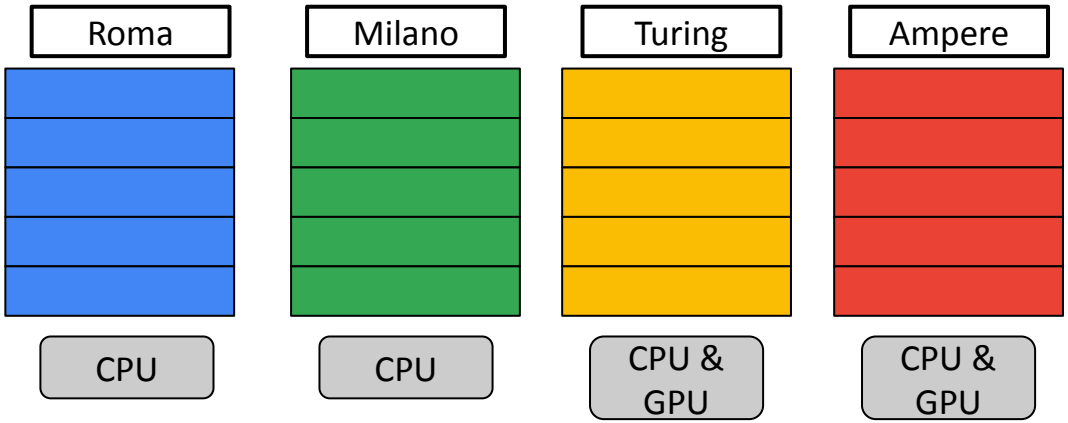
- 6MW Facility with air cooling (SLAC can use up to 2.5 MW)
- Flywheel + Generator allows for resilient power
- SLAC has over 100 racks
- 400 Gbps Networking to SLAC backbone

Slide content courtesy Jay Srinivasan

S3DF Infrastructure and Services used by Rubin

- Unix account provisioning and home directory storage
- Kubernetes provisioning & management via vCluster
- Batch compute provisioning & allocation via Slurm
- Provisioning & operation of specialized Qserv nodes
- Tiered storage (flash+disk) for large storage volumes
- Direct access to object storage-as-a-service (Ceph s3dfrgw; Weka S3 soon/now?)
- Specialized hardware setup for Embargo Rack
- Tape backup service
- Coordinated purchase of expansion storage and compute (2x per year is plan)
- *SLAC IT responsible for user account creation, external networking, cybersecurity*

S3DF Batch Compute Systems (Clusters)

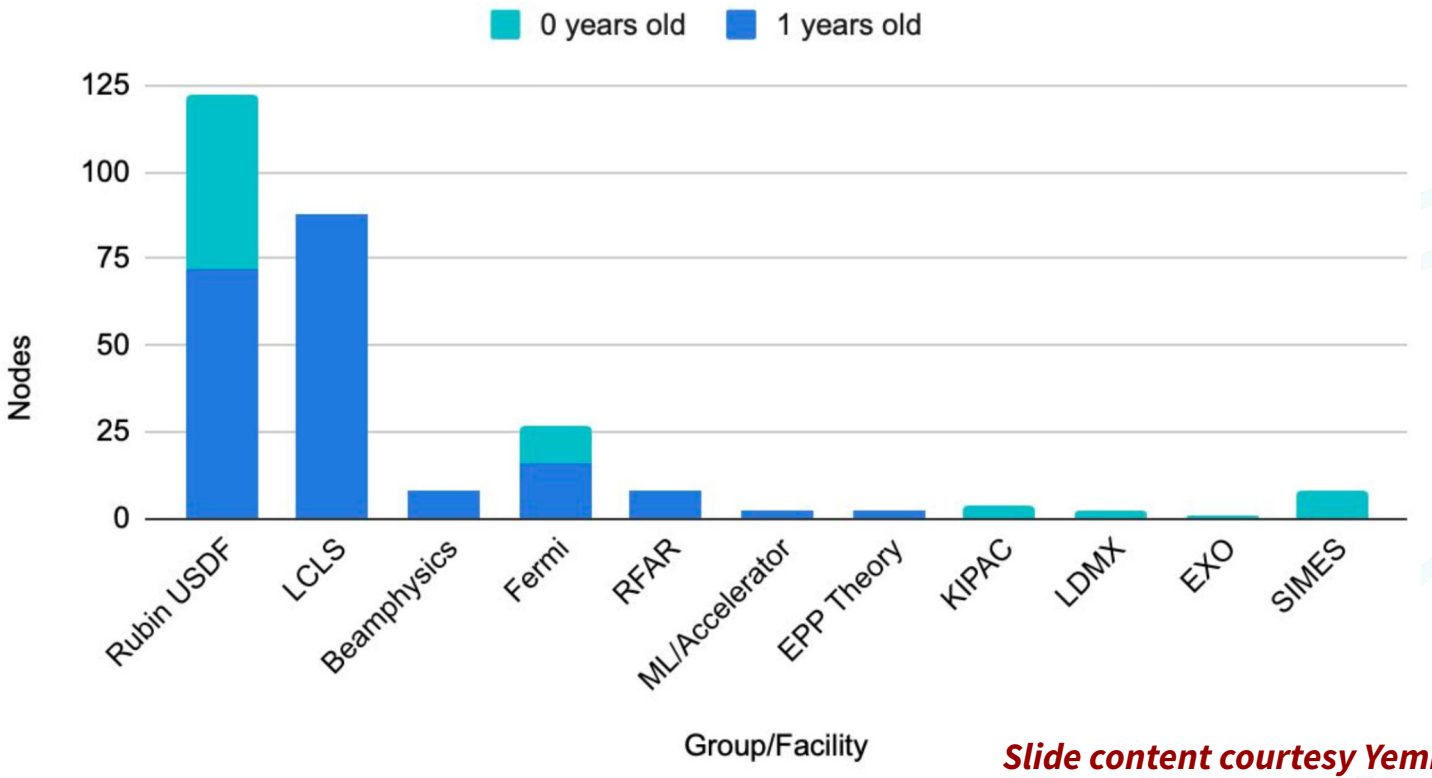


- Compute Clusters (Partitions) are available to Facilities
- Facilities (e.g. Rubin) are entities that contribute to the S3DF
- Users within a Facility are assigned to Repos - which are a group of resources for a common project/experiment

- Access to compute resources is gated by Slurm Batch System
- Users specify Partition + Repo to gain Slurm access to compute resources
- Coact portal to view and modify all user allocations and setup, resource management, etc.

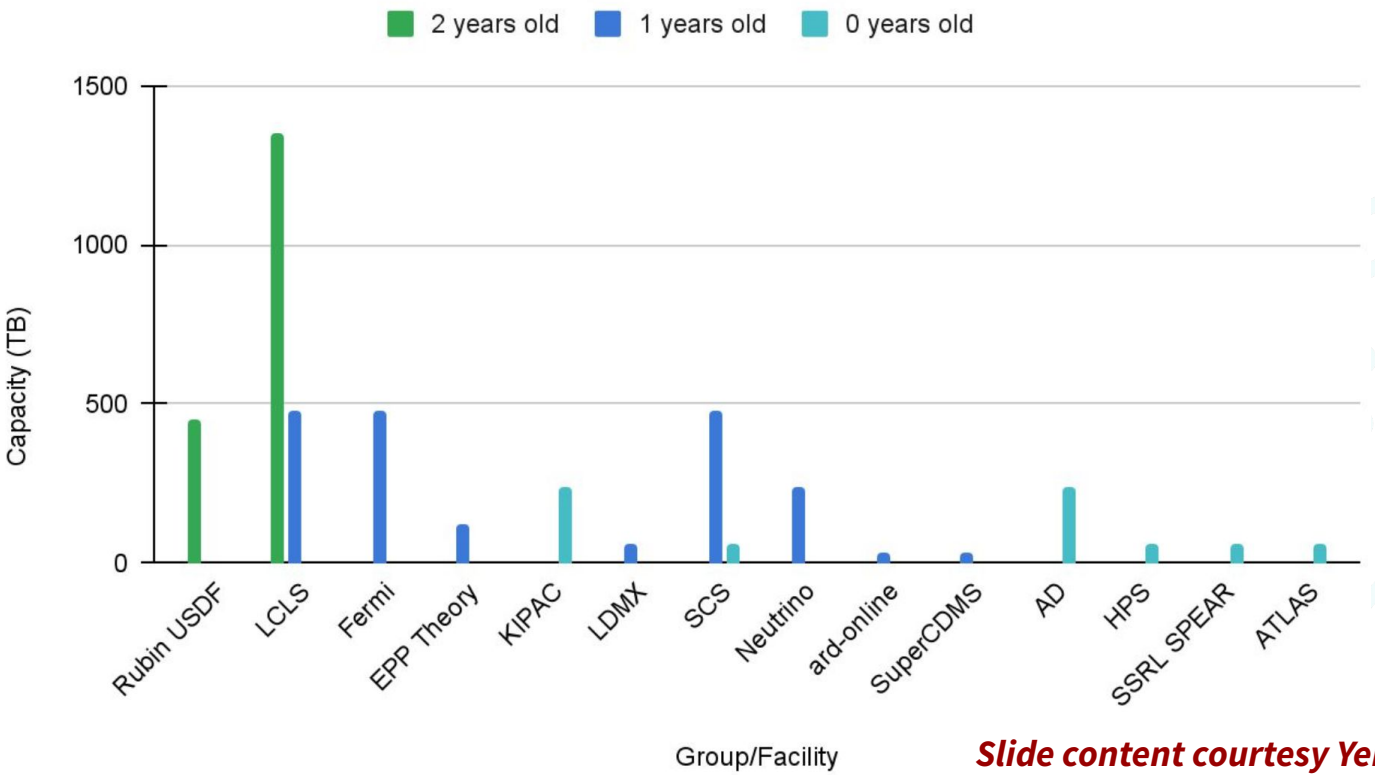
Slide content courtesy Jay Srinivasan

S3DF CPU node distribution (Milan)



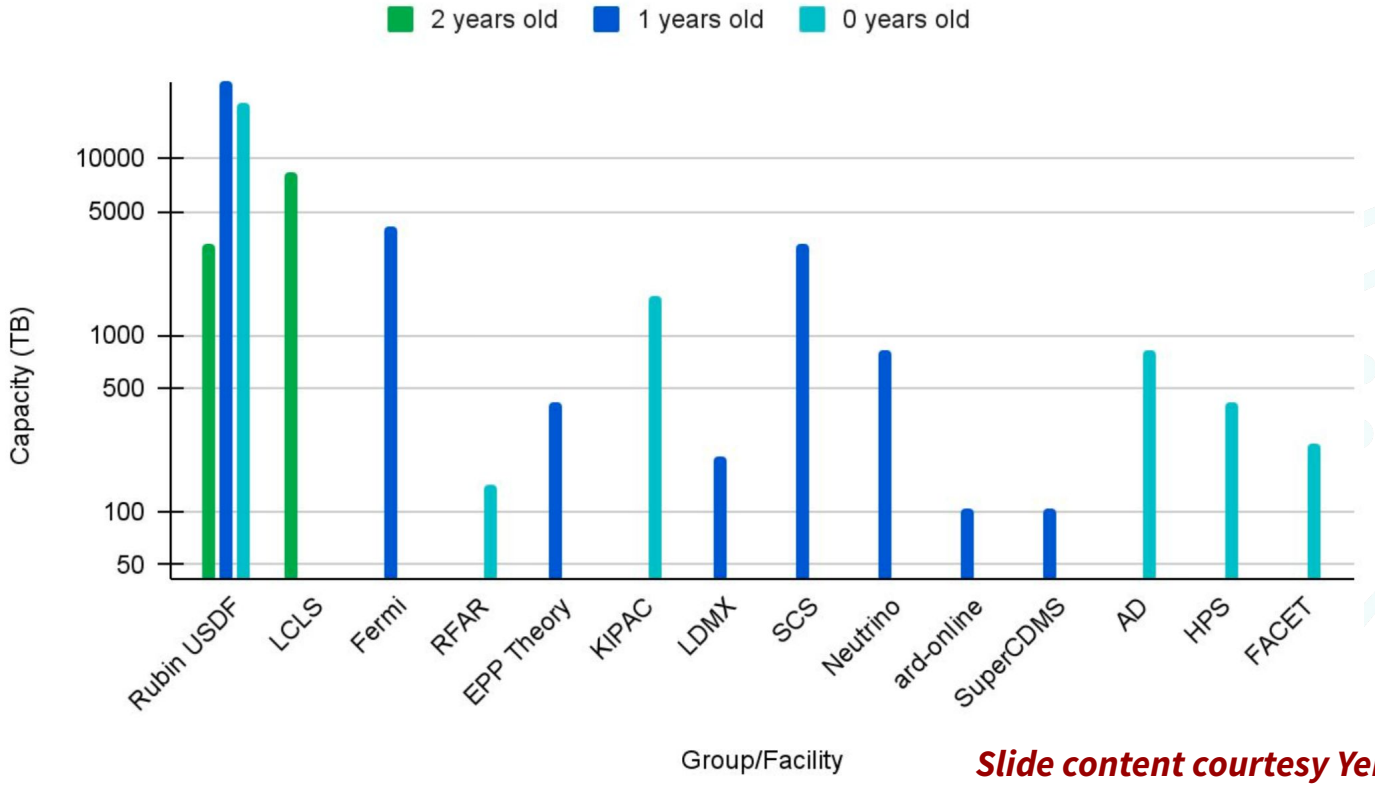
Slide content courtesy Yemi Adesanya

S3DF NVMe capacity distribution



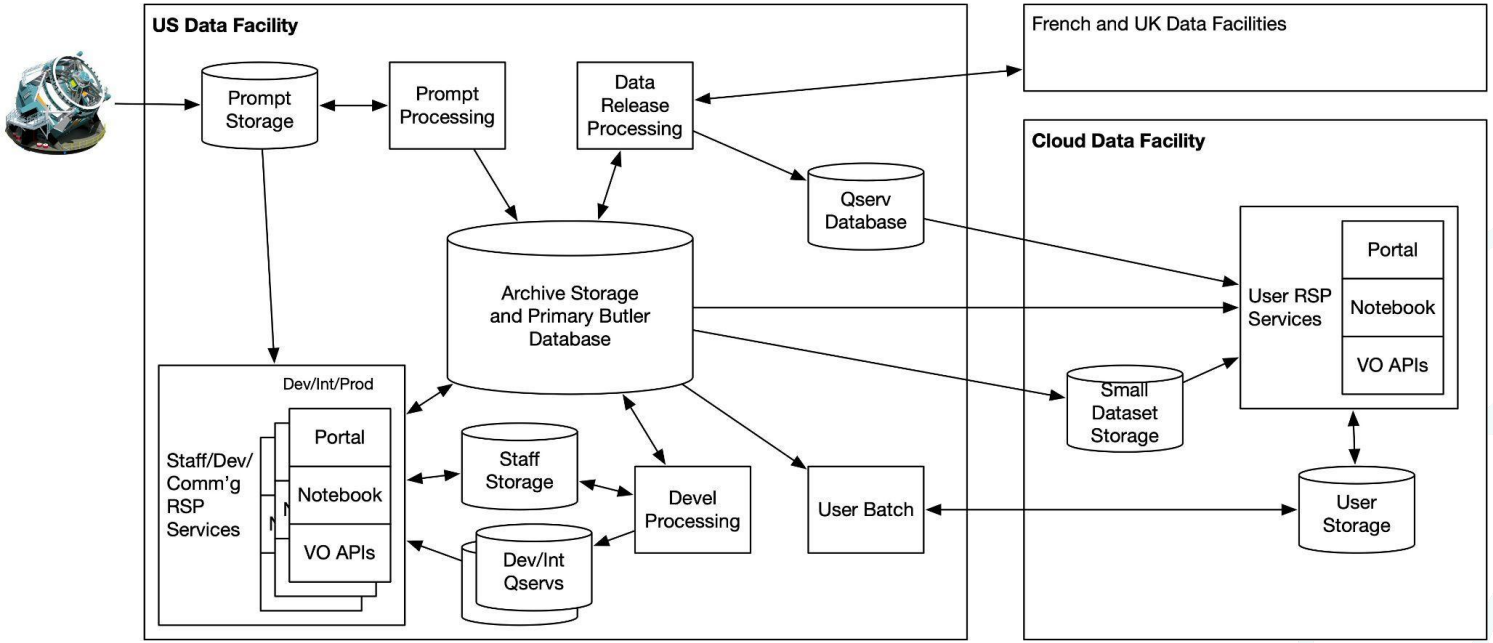
Slide content courtesy Yemi Adesanya

S3DF hard drive capacity distribution



Slide content courtesy Yemi Adesanya

Hybrid-cloud model



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

Allows:

- Separation of security concerns
- Burst response
- Reduced risk

(see DMTN-209)

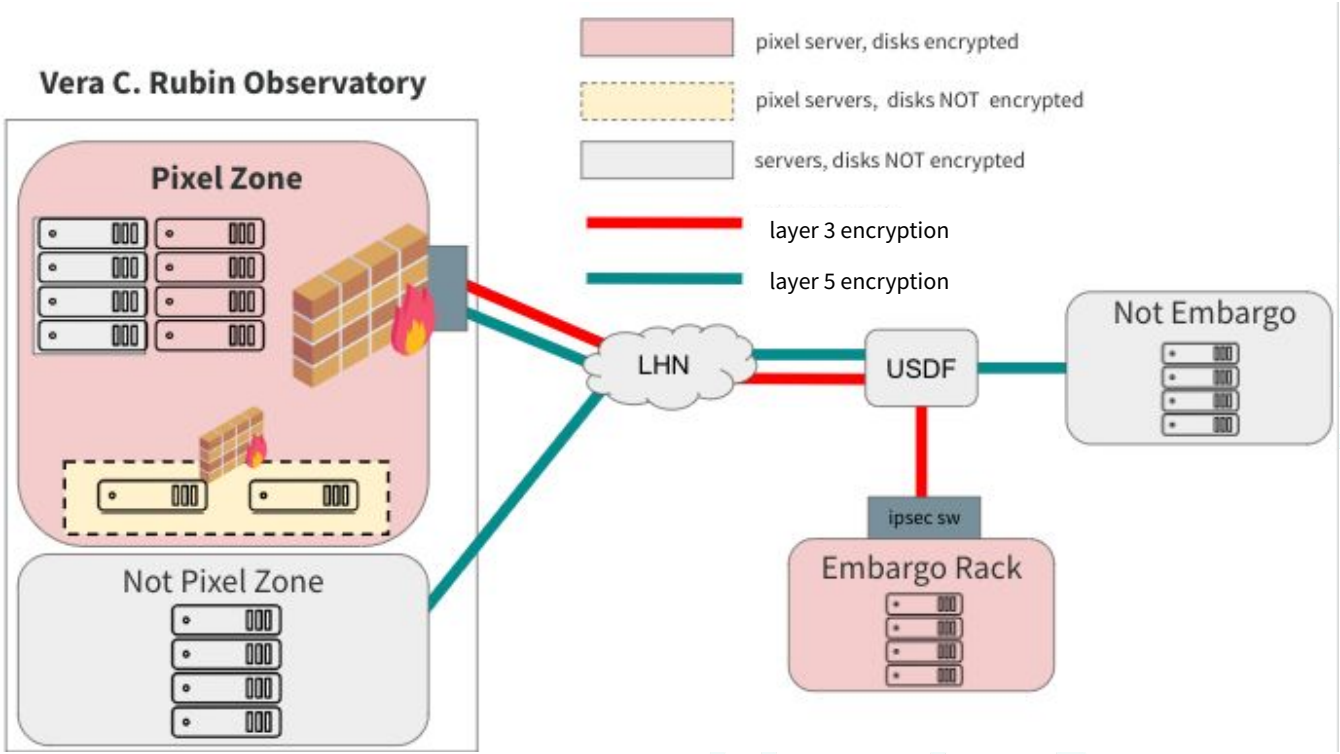
Data Embargo

Prevent identification or alerting on specified satellites

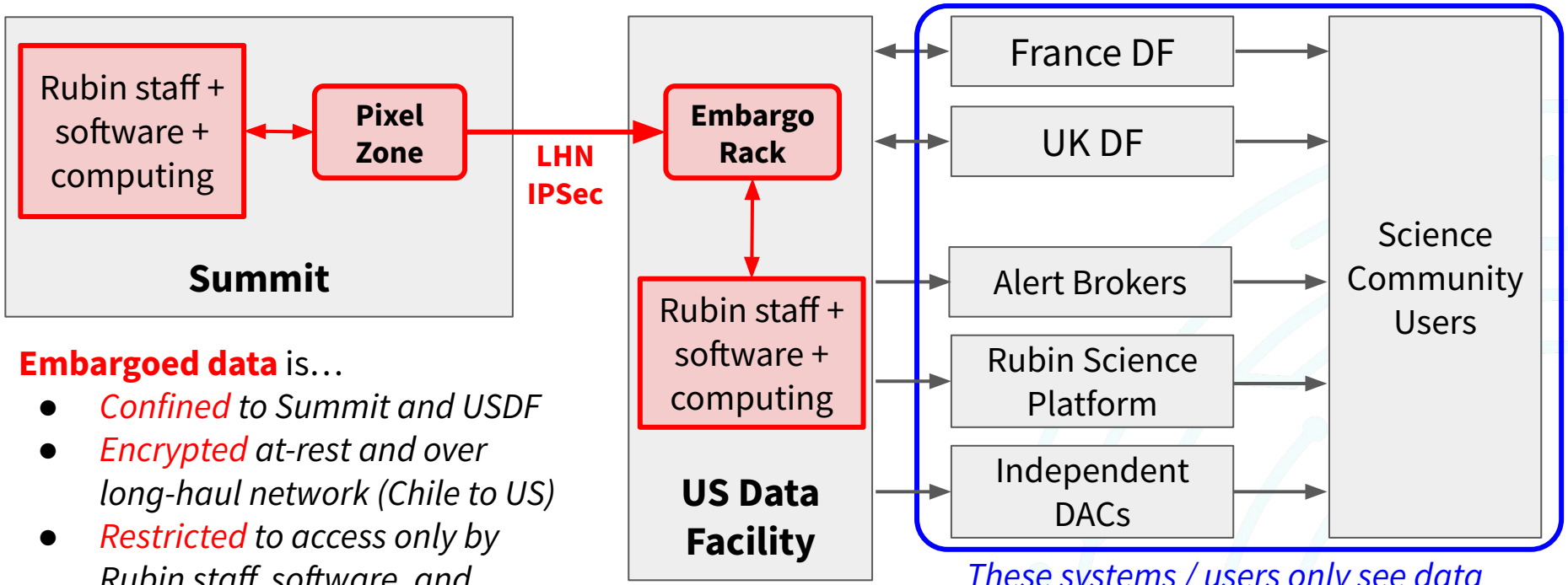
Pixel data encrypted and access-controlled during embargo period

- 30 days during on-sky engineering
- 80 hours after that

Pixels sent over an AES256 encrypted link from Summit to USDF



Data Embargo



Embargoed data is…

- *Confined to Summit and USDF*
- *Encrypted at-rest and over long-haul network (Chile to US)*
- *Restricted to access only by Rubin staff, software, and Rubin-dedicated compute*

These systems / users only see data after expiration of embargo period

S3DF ongoing development

- S3DF infrastructure split across SRCF-I and SRCF-II
  - “East-west” traffic has to go up to a higher network level
  - Impacts Ceph storage cluster that provides disk tier to Weka
  - Impacts communication between compute nodes and data storage
- Procuring more performant internal network h/w across S3DF
  - Currently delayed by vendor timescales (Arista), expected in March
  - Will also unblock deployment of expanded storage and compute
  - Recent maintenance window allowed prepositioning of existing hardware
  - Deployment will be minimally disruptive (hiccups for cutover, no downtime)
- Ceph storage cluster optimization: bringing in additional Ceph expertise
- Next major infrastructure upgrade in ~2 years
  - Planning beginning now
  - Need to maintain Rubin engagement with S3DF roadmap
- Model for operations support (Slack, ServiceNow, Jira) still being developed
  - *This is a hard problem – could be a session of its own!*

[end / transition]

Top 5 Concerns (Fabio/George/Adam synthesis)

- Demonstrating multi-site processing at scale
  - Access control, science validation, plan to scale up systematically
- Roles and responsibilities for execution and support during operations
  - Significant complexity, limited person power, site-specific and inter-site concerns
- Science user demand after releasing real data
  - Managing resource demands, coordinating IDAC contributions, responding to unanticipated priorities
- Potential mismatch between sizing model and actual needs
  - Fixed capital budgets, S3DF development roadmap
- Need for additional focus on core data-wrangling and stewardship
  - Identify responsible parties and key data sets, distinguish core needs from auxiliary