



Vera C. Rubin Observatory

SA3CC 2026

Cristian Silva - DevOps and Cybersecurity Manager



U.S. DEPARTMENT OF
ENERGY

SLAC

CHARLES AND LISA SIMONYI FUND
••• FOR ARTS AND SCIENCES •••

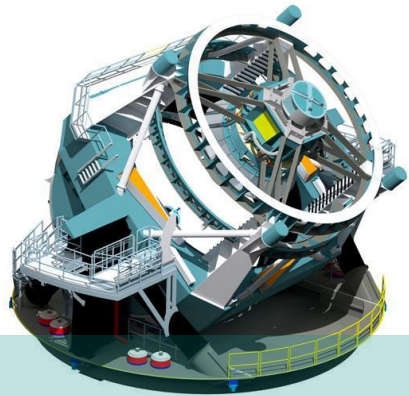


Data Production System Vision

Raw Data: 20TB/night



Sequential 30s images covering the entire visible sky every few days



Prompt Data Products

Alerts: up to 10 million per night

Results of Difference Image Analysis (DIA): transient and variable sources

Solar System Objects: ~ 6 million

Data Release Data Products

Final 10yr Data Release:

- Images: 5.5 million x 3.2 Gpx
- Catalog: 15PB, 37 billion objects



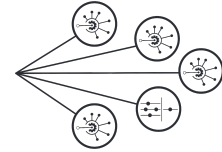
via nightly alert streams



via Prompt Products Database



via Data Releases



Community Brokers
Alert Filtering Service

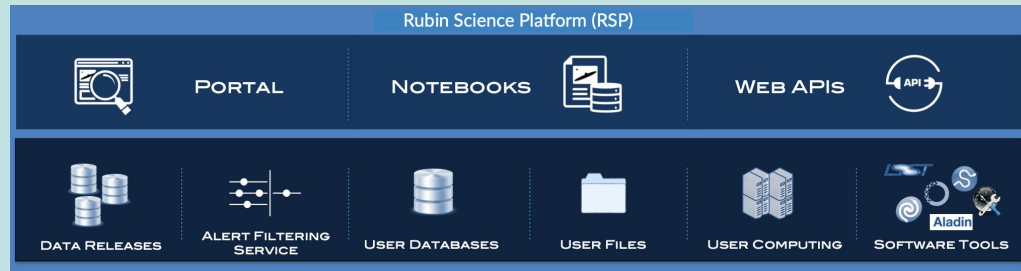


Rubin DACs (DFs & Chile)
Independent DACs (iDACs)

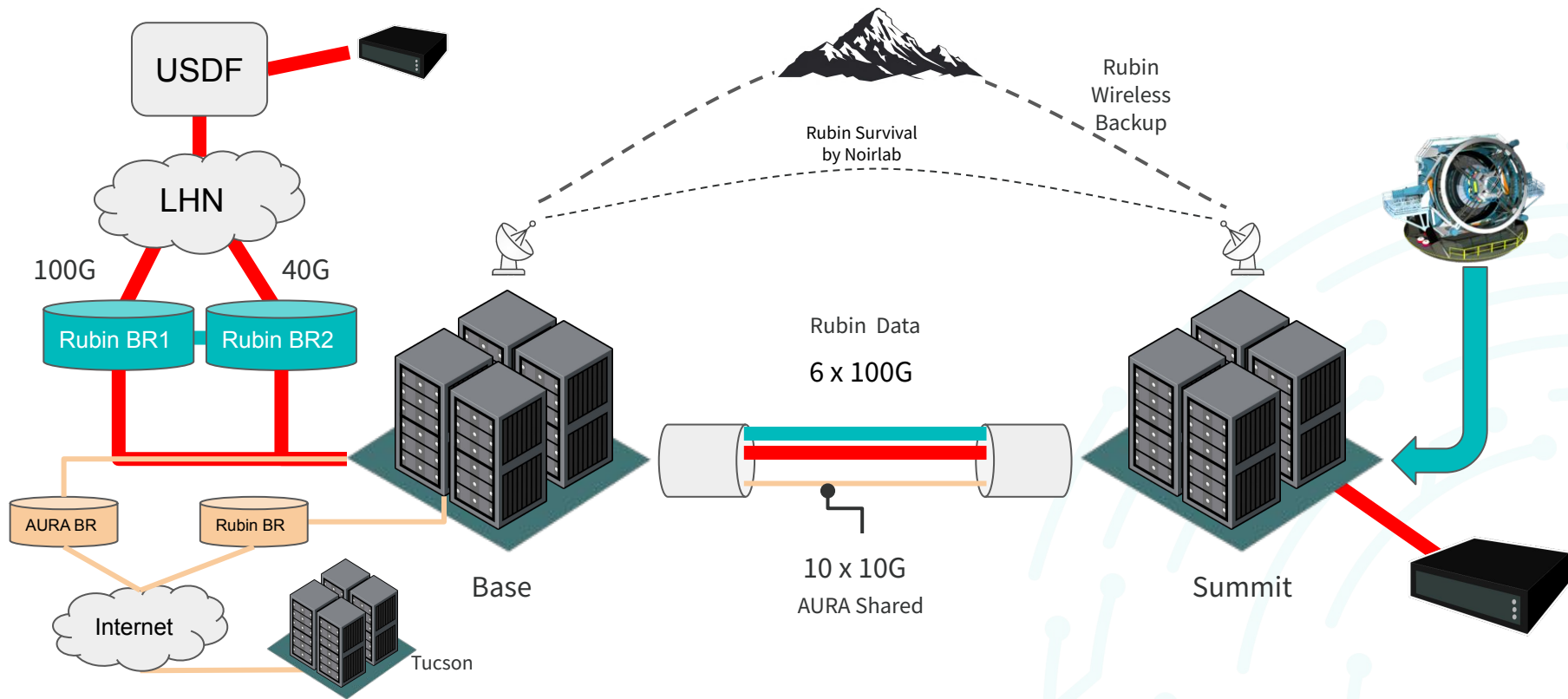
Access to proprietary data and the Science Platform require Rubín data rights

Rubin Science Platform

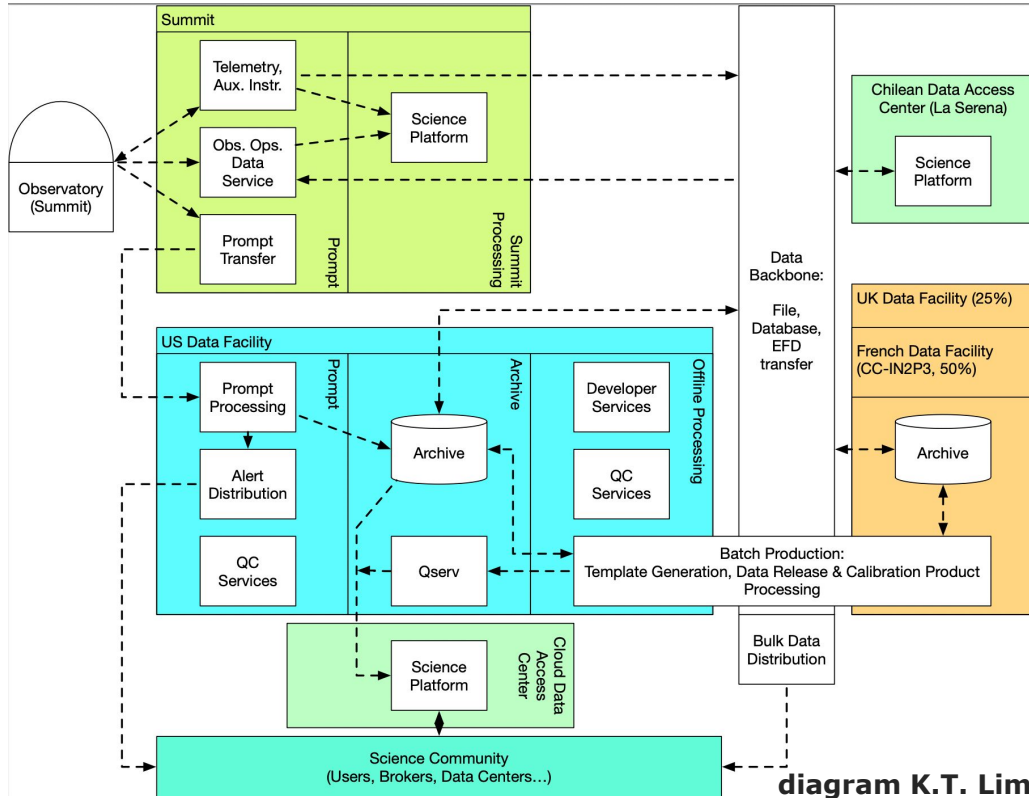
Provides access to Rubín Data Products and services for all science users and project staff



Vera Rubin Observatory Network - Topology



Rubin DM build and deploy



DM must build everything to get [Rubin products](#) to users.

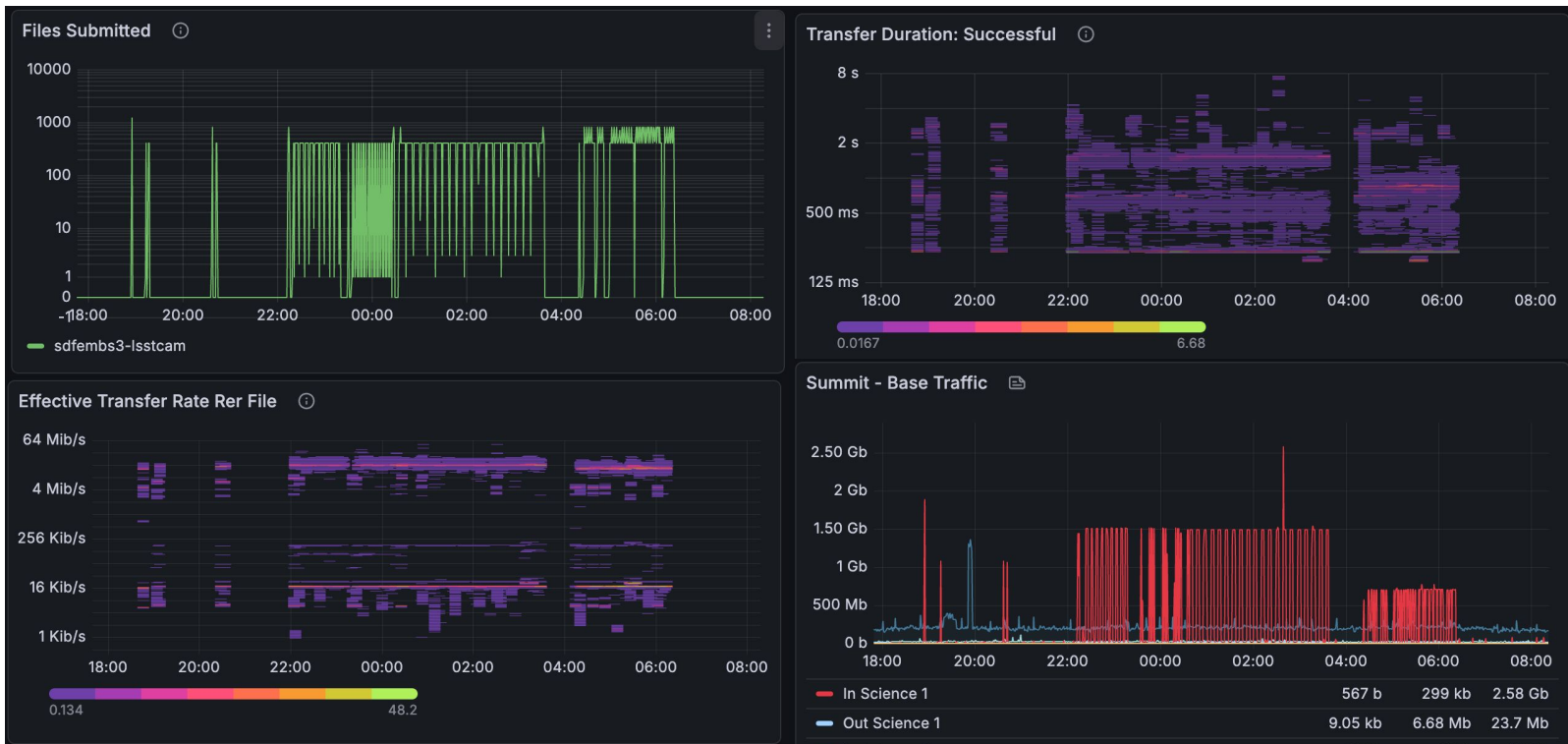
- large data sets (20TB/night)
- complex analysis
- aiming for small systematics
- Science Alerts in under 2 minutes .. (aiming for 1 minute)

All code on github: <https://github.com/lst>

Initial seconds

- **HEADER Service**
 - Picks up information from the kafka streams (DDS in the past)
 - Image target, RA, DEC,
 - Makes available to any subscriber like Camera
- **Camera readout is 2s**
 - Camera assembles a fits file per CCD
 - Uses the header info from HEADER service to provide a useful File e.g. has WCS.
 - More information is put in a JSON sidecar
 - ID is Day + Sequence e.g. 2024121100607 - 2024 December 11 seq 607; compromise between human understandability and machine tracking
 - On mountain previews (just ISR Instrument Signature Removal) ~ 6seconds
- **The Image and sidecar are transferred to SLAC**
 - Image over IPsec channel (Arista AES 256)
 - S3 protocol using persistent HTTPS connections to minimize connection setup overhead
 - Immediately go to Prompt Processing and Butler ([Jenness 2022](#)) within 2 minutes

Network Performance



The Infrastructure

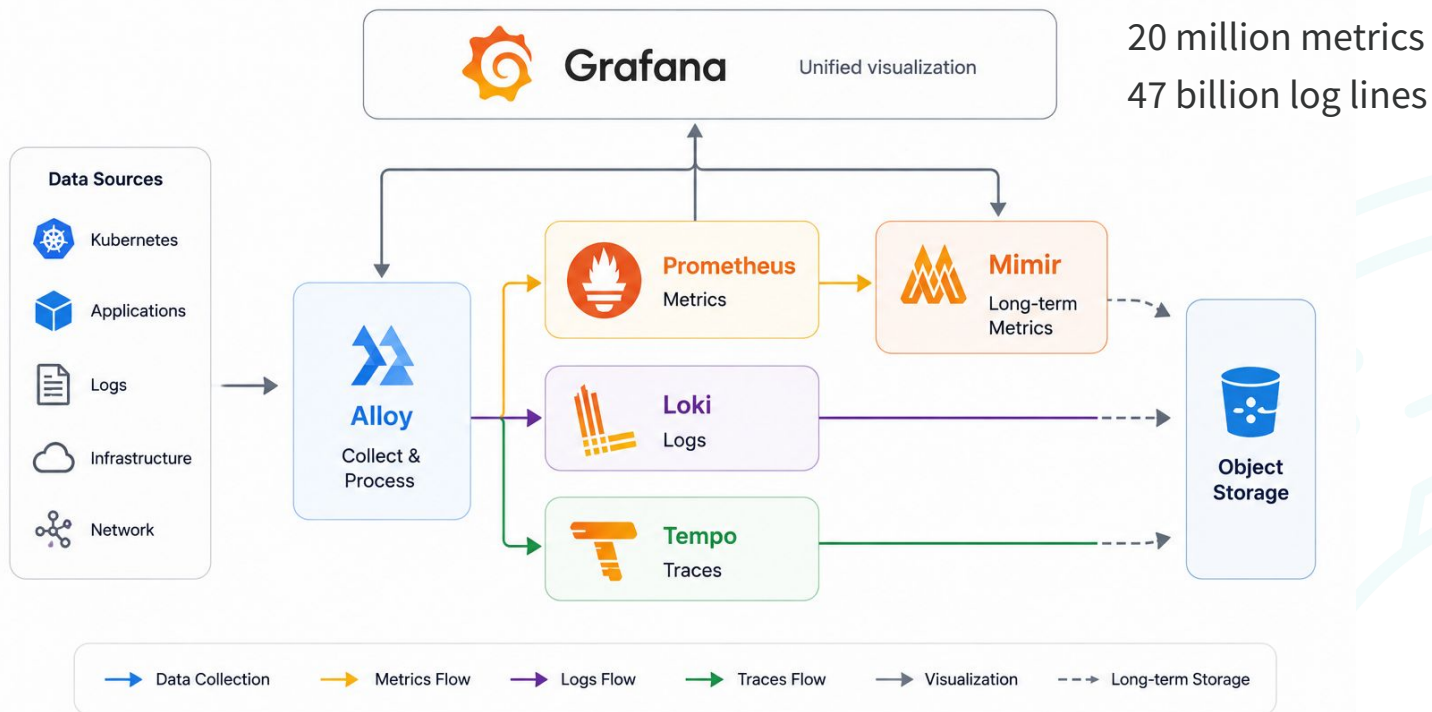
- All servers and network devices are building blocks (Legos)
 - Failed Nodes are easy to replace by any other node (they are all the same model), so we don't rely on vendor support to keep running.
 - Provisioning a new node is straightforward and predictable, with no unexpected issues..
- Infrastructure as Code
 - Rubin deployments are consistent, with the same configuration applied uniformly across all environments (dev, staging and prod)
 - Automation with rapid setup. We don't duplicate work and avoid manual configs.
 - All changes are tracked and easy to audit (who did it?, why?)
 - We can quickly rebuild after a failure. The config is not stored locally.(github)
- Rubin runs on open source
 - Reduces cost. We pay \$0 in licenses in the control system.
 - Large support from the community (our problem it was likely the problem of someone else)
 - Avoids vendor lock-in, so we can choose what works better for us.
 - If there's a feature we need and the maintainers don't have plans to make it, we can build it ourselves and contribute it back to the community. (yes, we have done it)

The observatory generates a bunch of data about itself

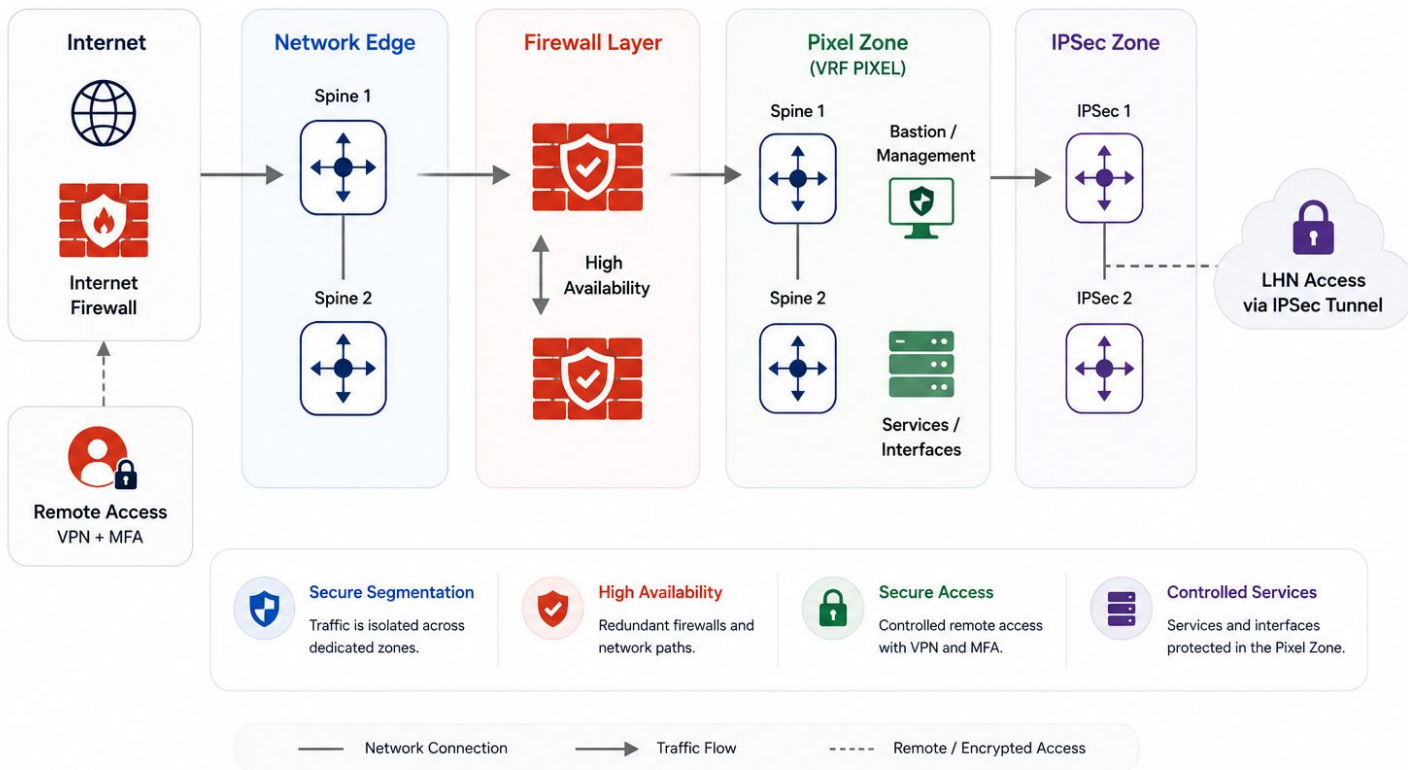
- Total of 1148 DDS topics
 - 350 commands, 531 events, 267 telemetry topics
 - Various frequencies up to 50hz (mirror actuators)
 - All values broadcast on the Data Distribution System (DDS) captured in EFD via Kafka stored in influxdb
- Total throughput ~21GB/h → real-time monitoring
 - ~15TB per month → offline analysis
 - ~1.5PB for the 10yr of operations → trend analysis (relational)
 - So joins can be made with astronomy data if needed



Observability Stack



Security - Pixel Zone



AI accelerates vulnerability discovery

Security flaws and exploitable code are identified much faster than traditional methods.

Faster patching cycles

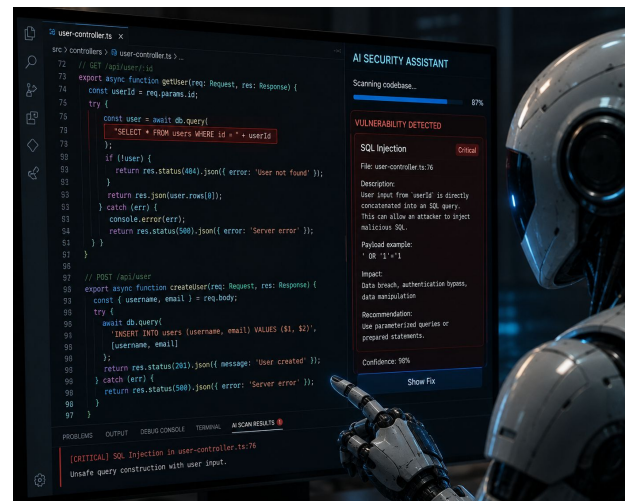
Reduce remediation time as vulnerabilities are discovered and weaponized more quickly.

Continuous security assessments

AI enables ongoing analysis of applications, infrastructure, and dependencies instead of periodic reviews.

Increased pressure on security teams

Faster threat discovery requires stronger patch management and response processes.



Fiber Outages

Several events on the primary link
Mainly in the Ovalle - Vicuña road.

Multiple carriers
Rubin fiber was the only one a few years ago

Working with Reuna





- Computing = AMD EPYC ZEN4 9554
- Storage = Fast - AMD EPYC ZEN4 9554P ~ 30T per node
- Storage = Slow - AMD EPYC 9124 4th Gen ~ 500T per node
 - ~ 13PB of usable storage
 - Running on top of Ceph and Rubin CI/CD
 - Access to Community will be done by VPN.
 - More nodes every FY