VERA C. RUBIN OBSERVATORY

Rubin Observatory US Data Facility

NET Networking Update

Mark Foster with Richard Dubois & Phil Marshall

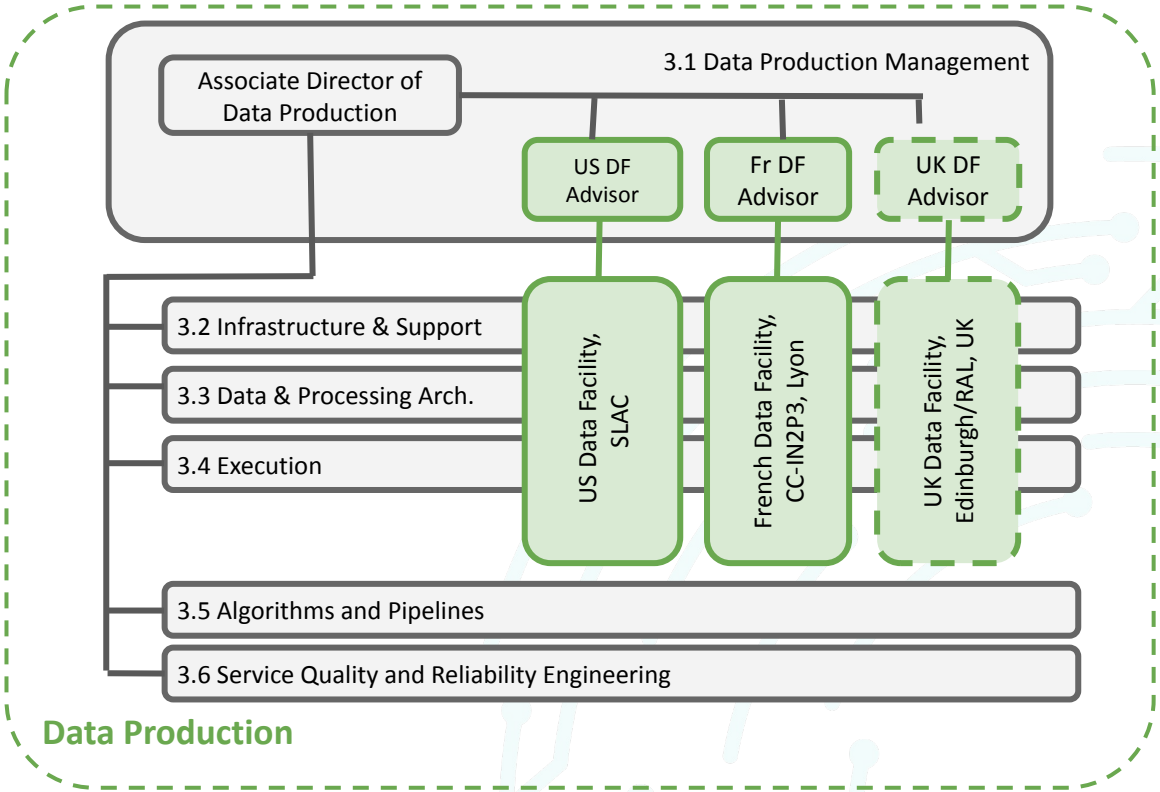
April 2021

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

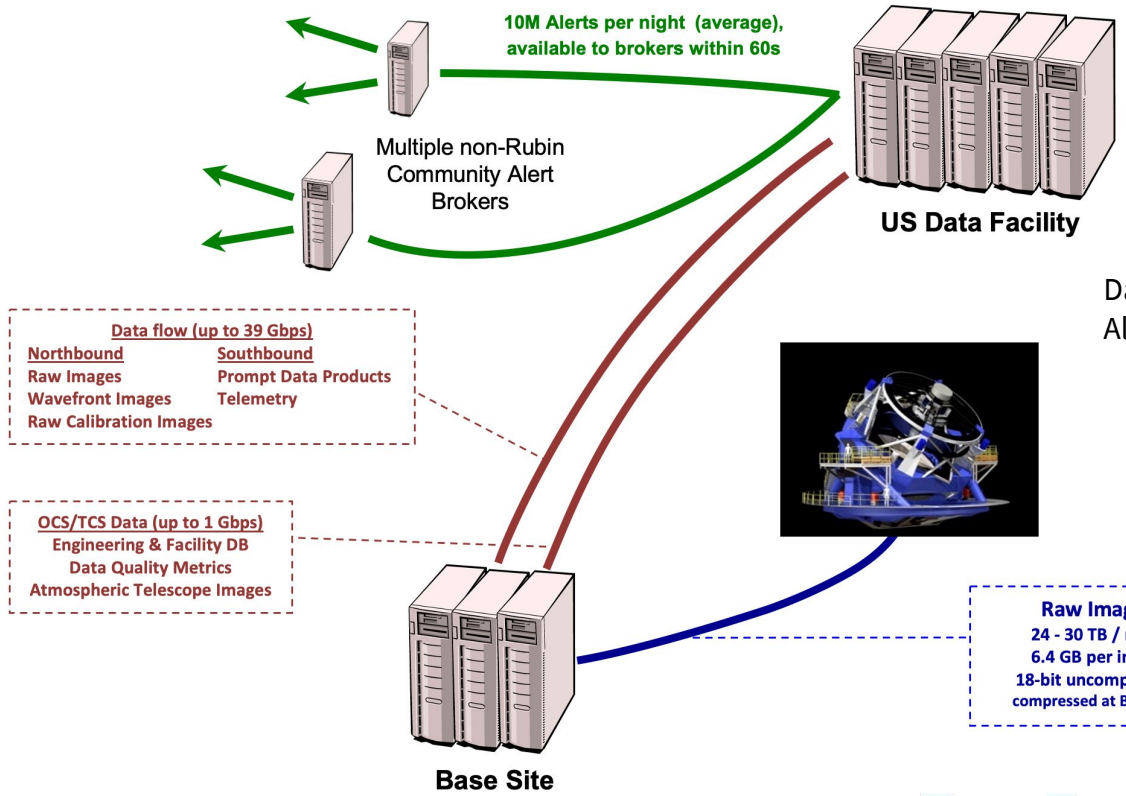
Data Production as a Matrixed, International, Multi-center Department

Multiple data facilities

- **United States Data Facility (USDF)**
- French Data Facility at CC-IN2P3
  - 50% of data release processing
- UK Data Facility
  - 25% of data release processing
- Independent Data Access Centers
  - May serve only a subset of data
- Clouds

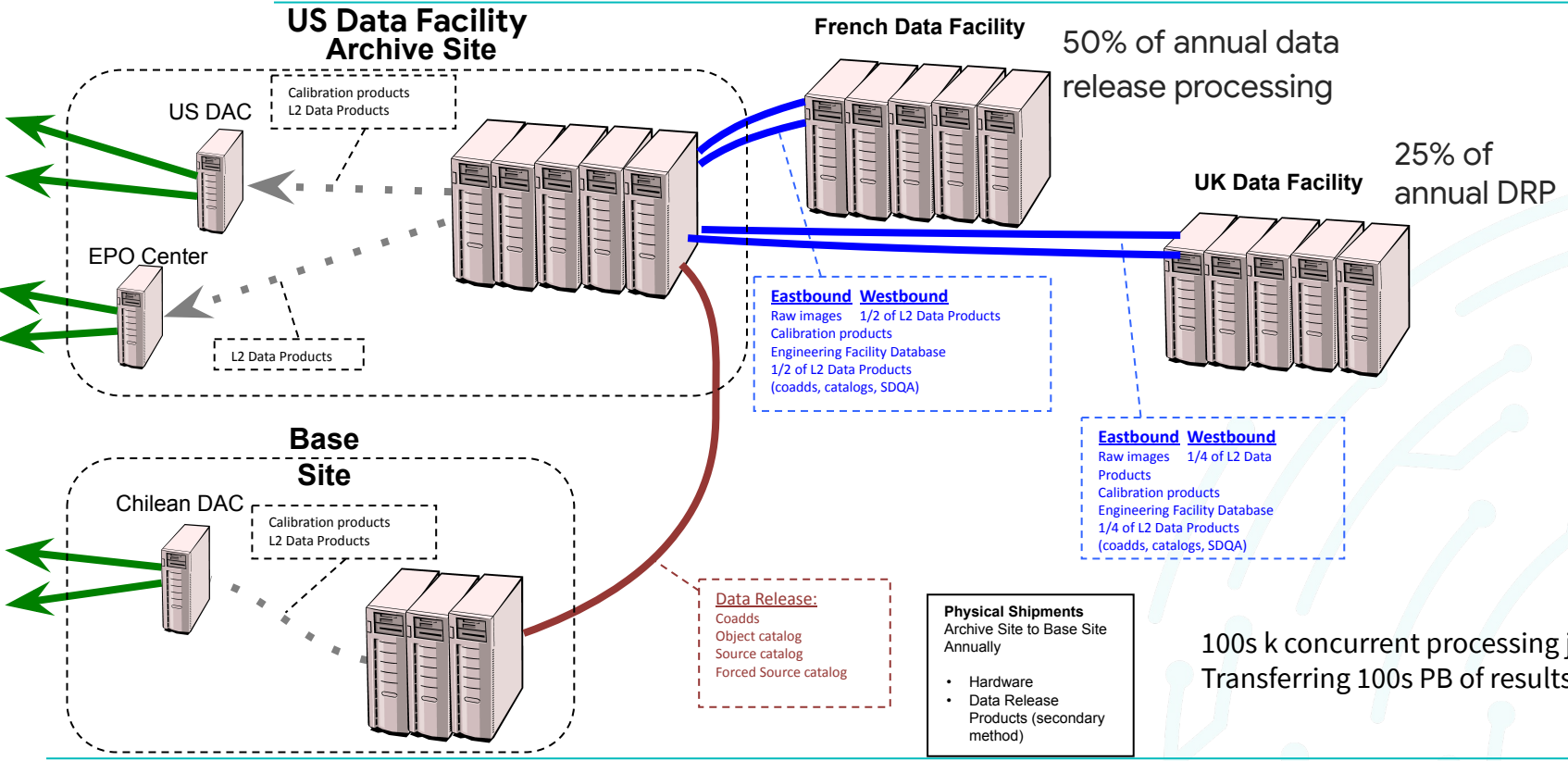


Data Flows: Prompt Processing



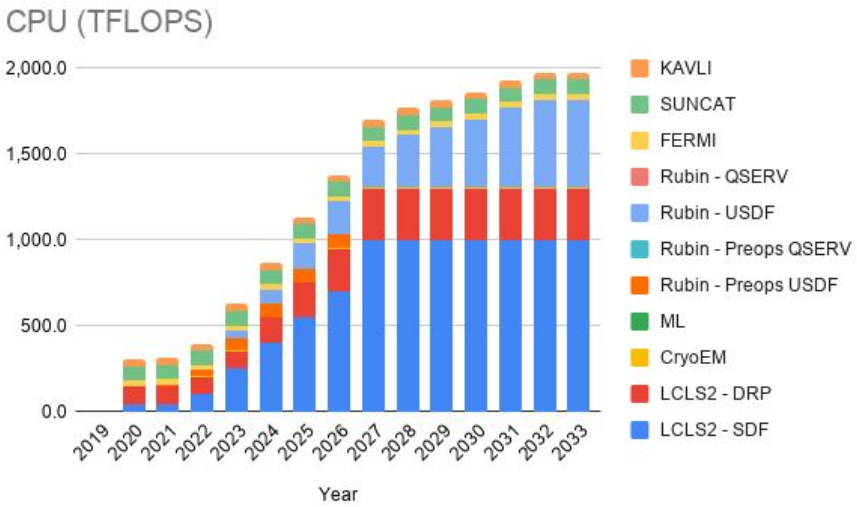
Data to SLAC in 7 secs
Alerts generated in 1 min

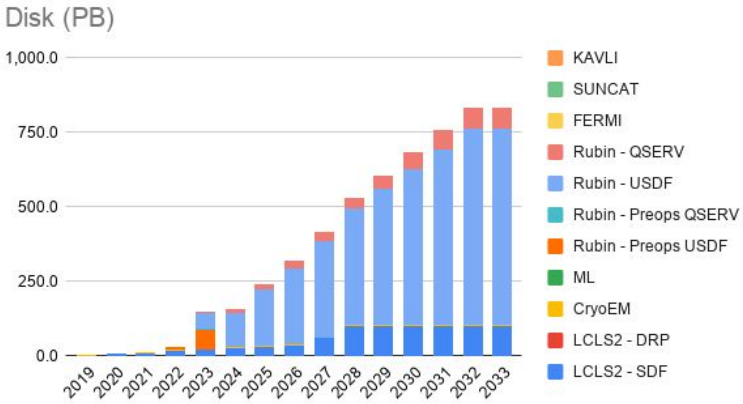
Data Flows: Data Release Processing

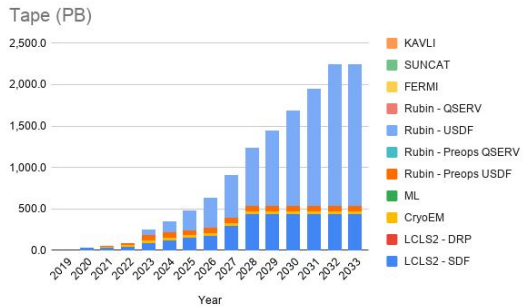


100s k concurrent processing jobs
Transferring 100s PB of results

SLAC Shared Data Facility (SDF)



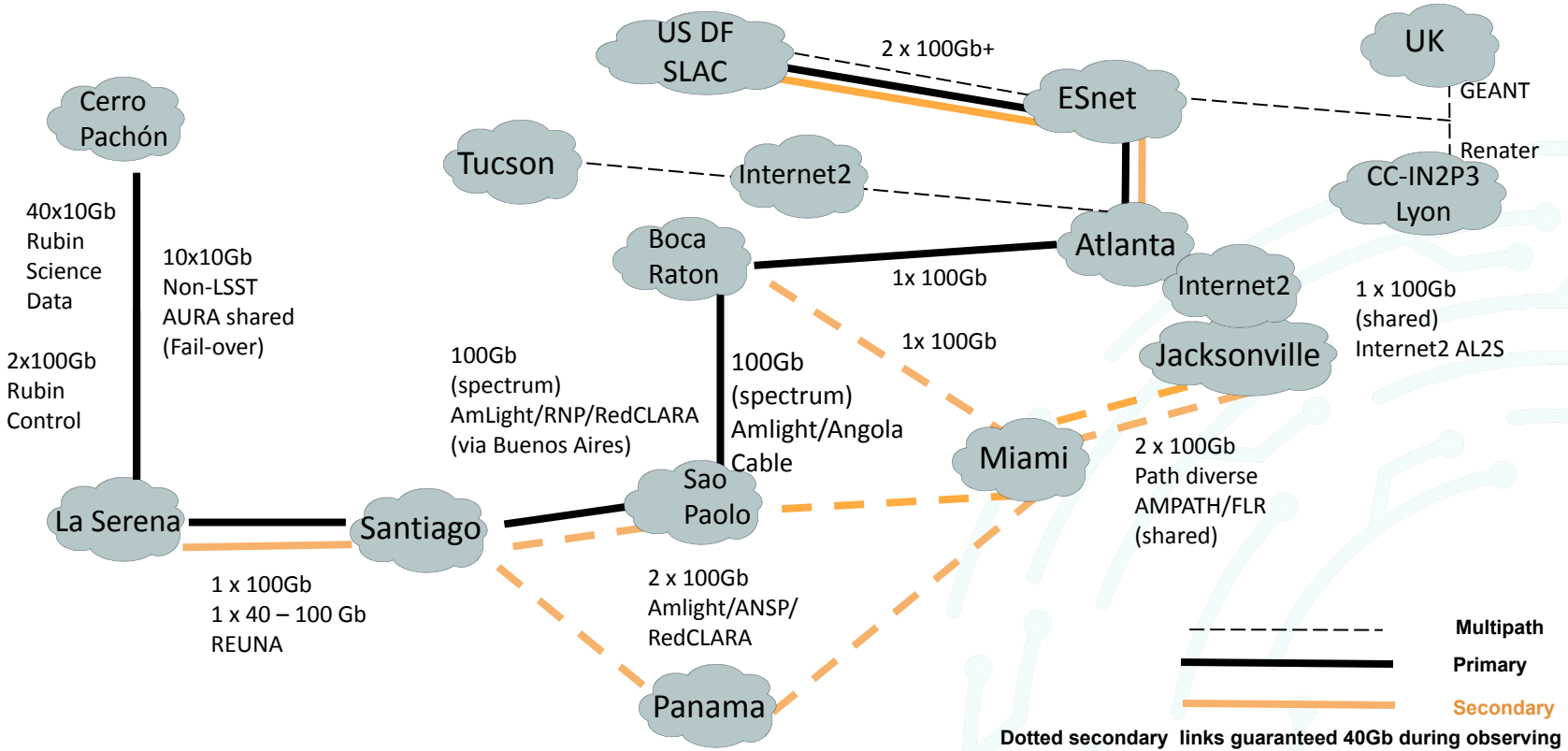




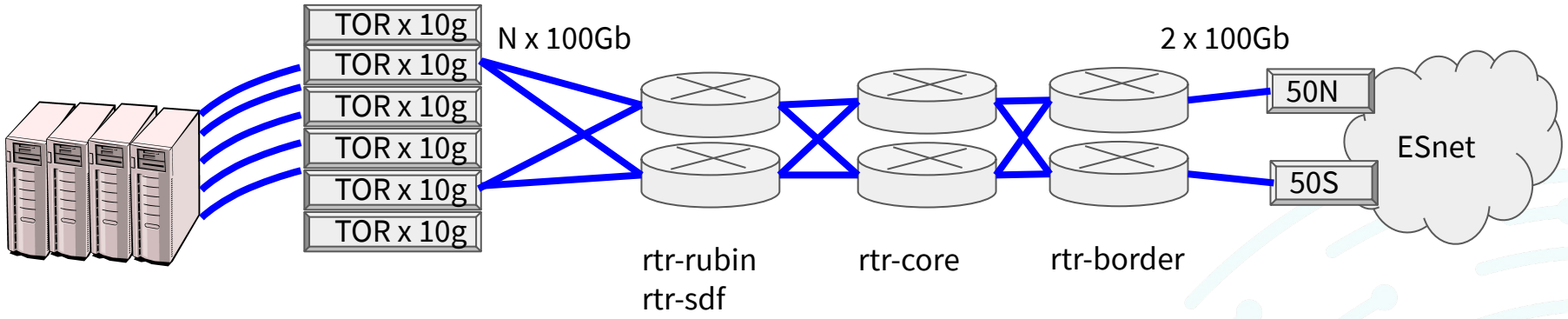
Physical transitions: B50, SRCF, SRCF-II - active planning underway

CPU dominated by LCLS-II
Storage dominated by Rubin

Long Haul Network Links



SLAC / USDF Networking



Rubin Server/Storage enclave: TOR switches: Nx10G with Nx100G (layer2) to routing infrastructure via SLAC core and SLAC border

Existing: support 200Gbps aggregate capability between SLAC and other sites with multiple ESnet 100Gbps links; ability to scale Nx100Gbps now, Nx400Gbps future.

ESnet6: two optical nodes on SLAC premises: part of Bay Area optical ring (multi-Tbps optical capacity)

Questions?

Rubin Observatory Sept 2019

http://www.lsst.org

http://community.lsst.org

Backup

Data Facility Functions

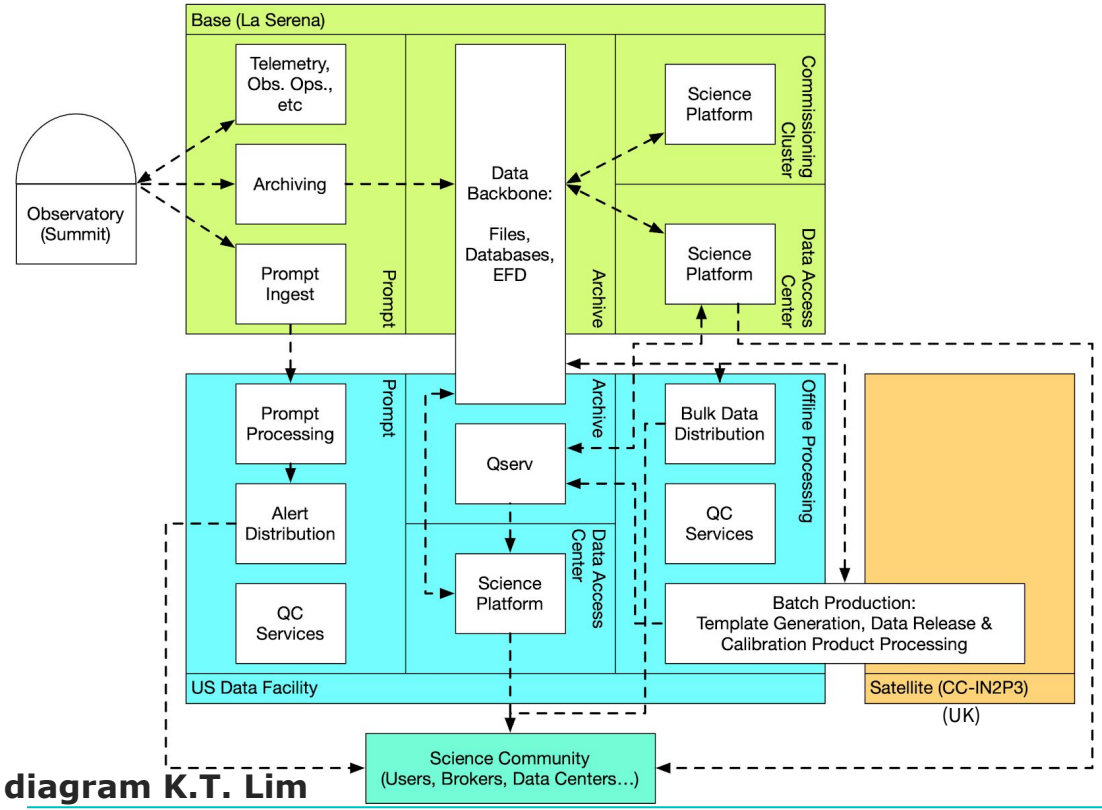


diagram K.T. Lim

- large data sets (20TB/night)
  - complex analysis
  - aiming for small systematics
- Science Alerts in under 2 minutes .. (aiming for 1 minute)
- Annual processing of all data taken to date - ~200 days to execute annually