Centre de Calcul
de l'Institut National de Physique Nucléaire
et de Physique des Particules

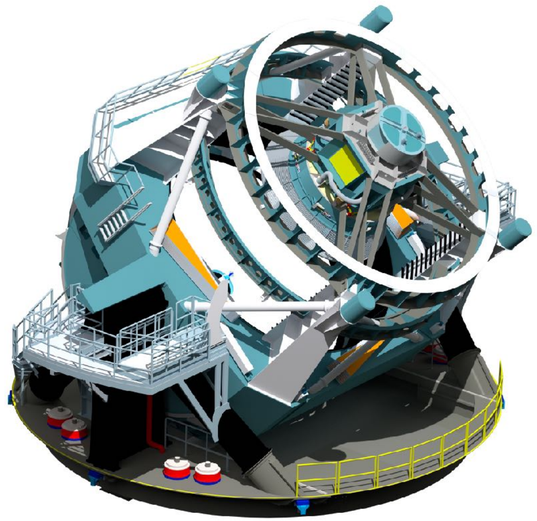
ESCAPE datalake

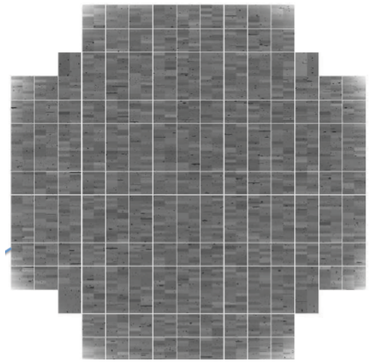
precursor infrastructure for prototyping LSST data distribution

fabio hernandez, lionel schwarz

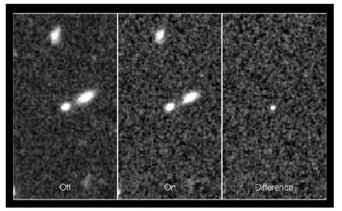
ESCAPE DIOS webinar, March 31st, 2022

RUBIN OBSERVATORY LEGACY SURVEY OF SPACE AND TIME

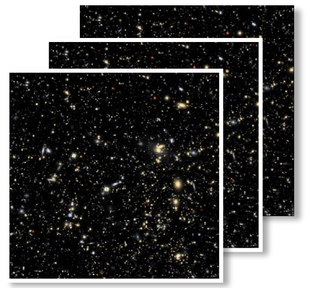




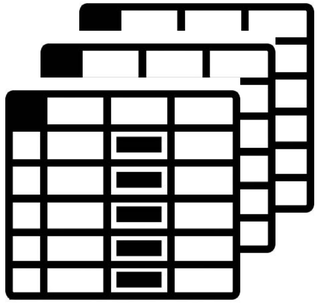
raw images



alerts



co-added images



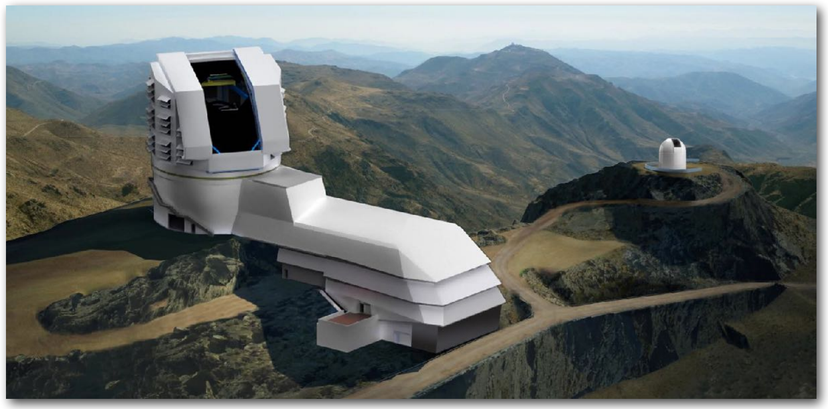
astronomical catalog

science collaborations

LSST aims to deliver a catalog of 20 billion galaxies and 17 billion stars with their associated physical properties

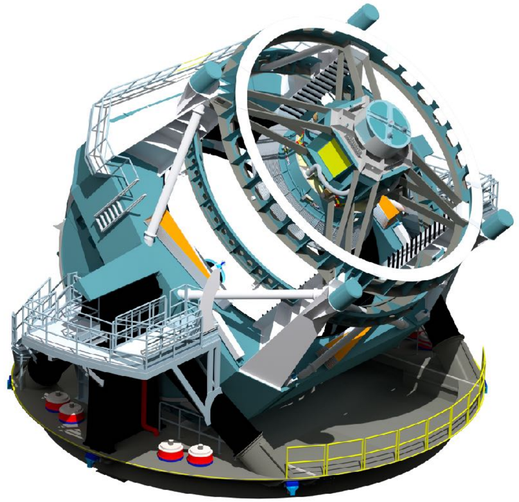
LSST OVERVIEW

OBSERVATORY



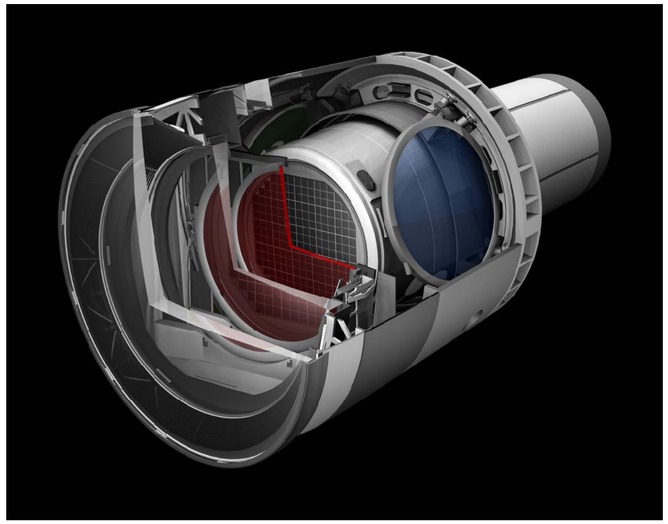
southern hemisphere | 2647m a.s.l. | stable air | clear sky | dark nights | good infrastructure

TELESCOPE



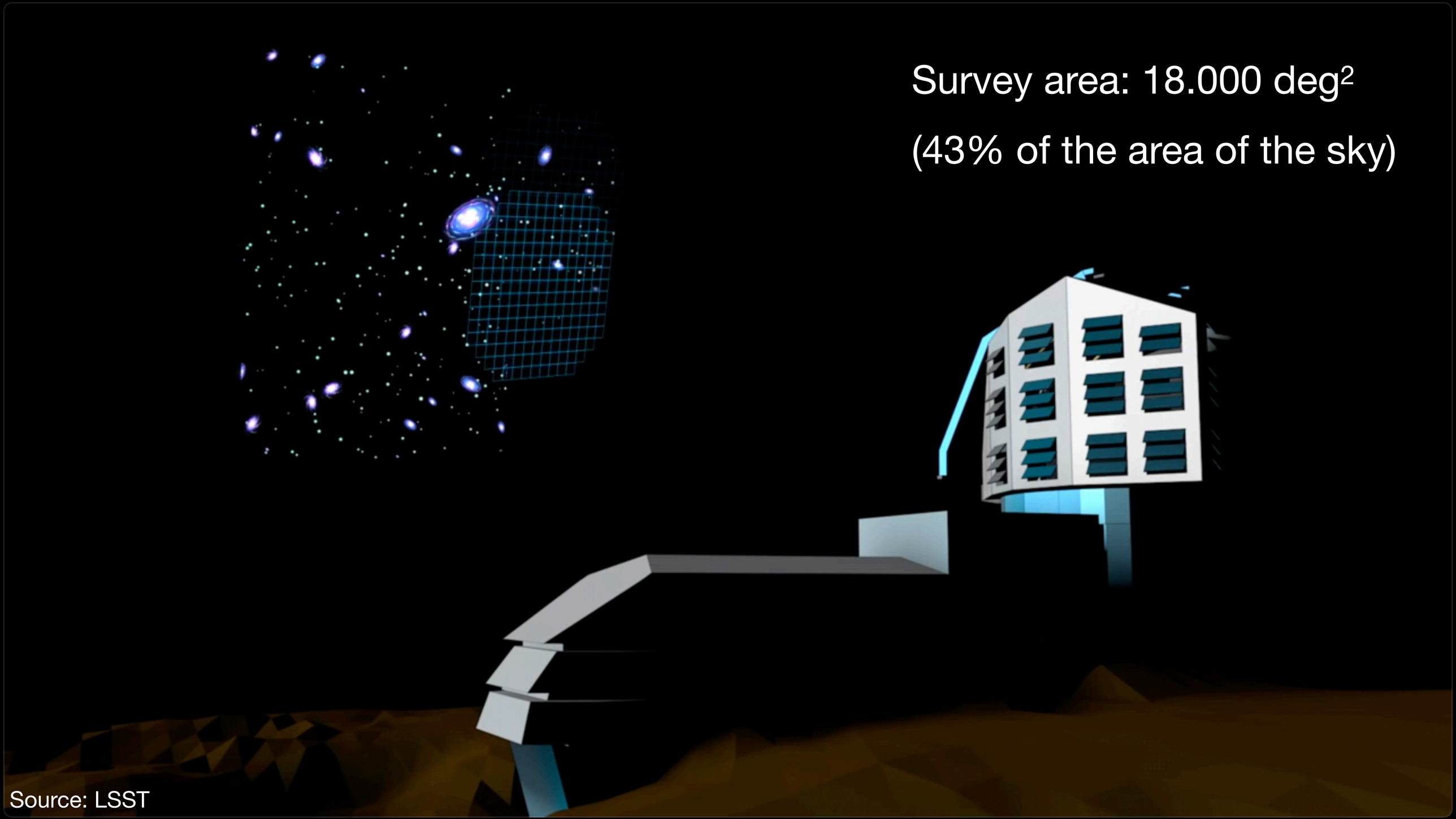
main mirror ⌀ 8.4 m (effective aperture 6.5 m) | large aperture: f/1.234 | wide field of view | 350 ton | compact | to be repositioned about 3M times over 10 years of operations

CAMERA



3.2 G pixels | ⌀ 1.65 m | 3.7 m long | 3 ton | 3 lenses | 3.5° field of view | 9.6 deg^2 | 6 filters ugrizy | 320-1050 nm | focal plane and electronics in cryostat at 173K

Source: LSST


Survey area: 18.000 deg^2
(43% of the area of the sky)
Source: LSST

LSST OVERVIEW (CONT.)

- ## Principle of operations

  *90% of the observing time of the telescope devoted to a **deep-wide-fast survey***

  one complete visit of the southern hemisphere sky every 3-4 nights, from 2024 for 10 years

  each patch of the observable sky to be visited about 1000 times

  43% of the celestial sphere will be covered by this survey

- ## Science themes

  *determining the nature of **dark energy** and **dark matter***

  *taking an inventory of the **solar system***

  *exploring the **transient** optical sky*

  *mapping the structure and evolution of the **Milky Way***

Raw data

6.3 GB per exposure (compressed)

2000 science images + 500 calibration images per night

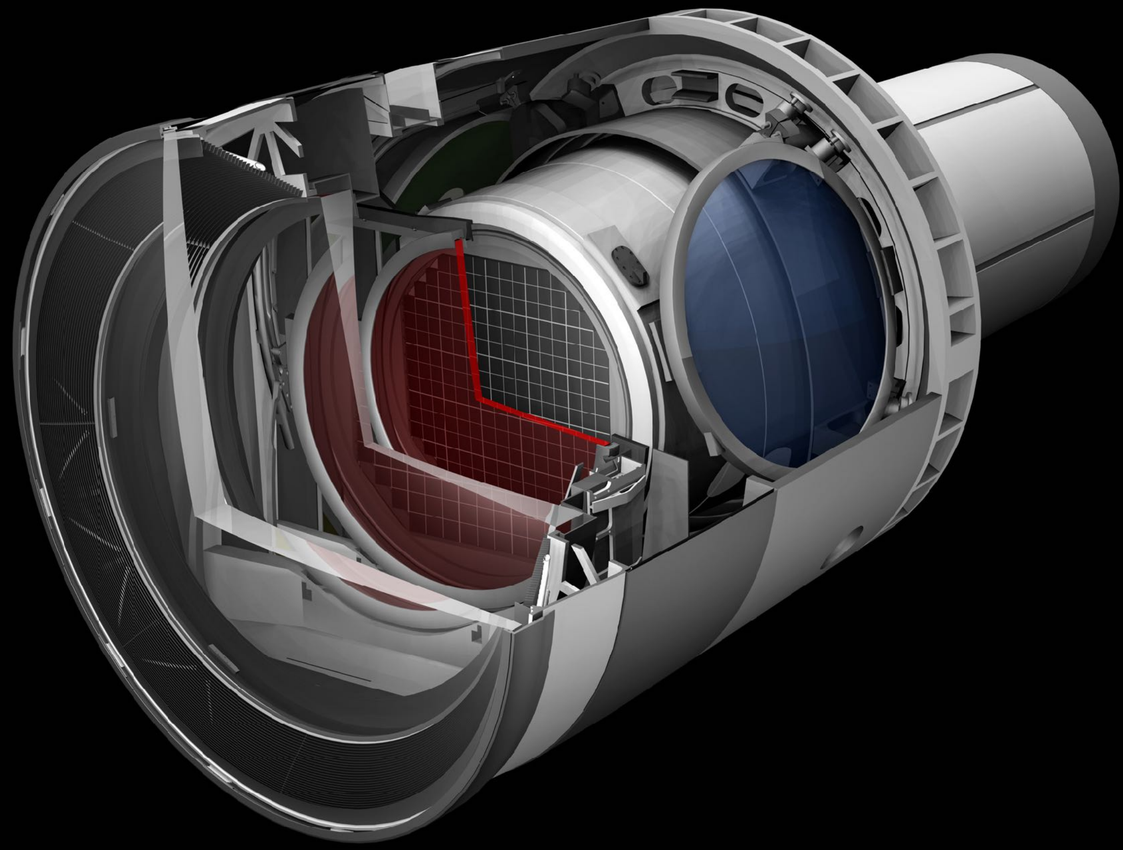
~15 TB per night

300 nights per year, ~5 PB per year

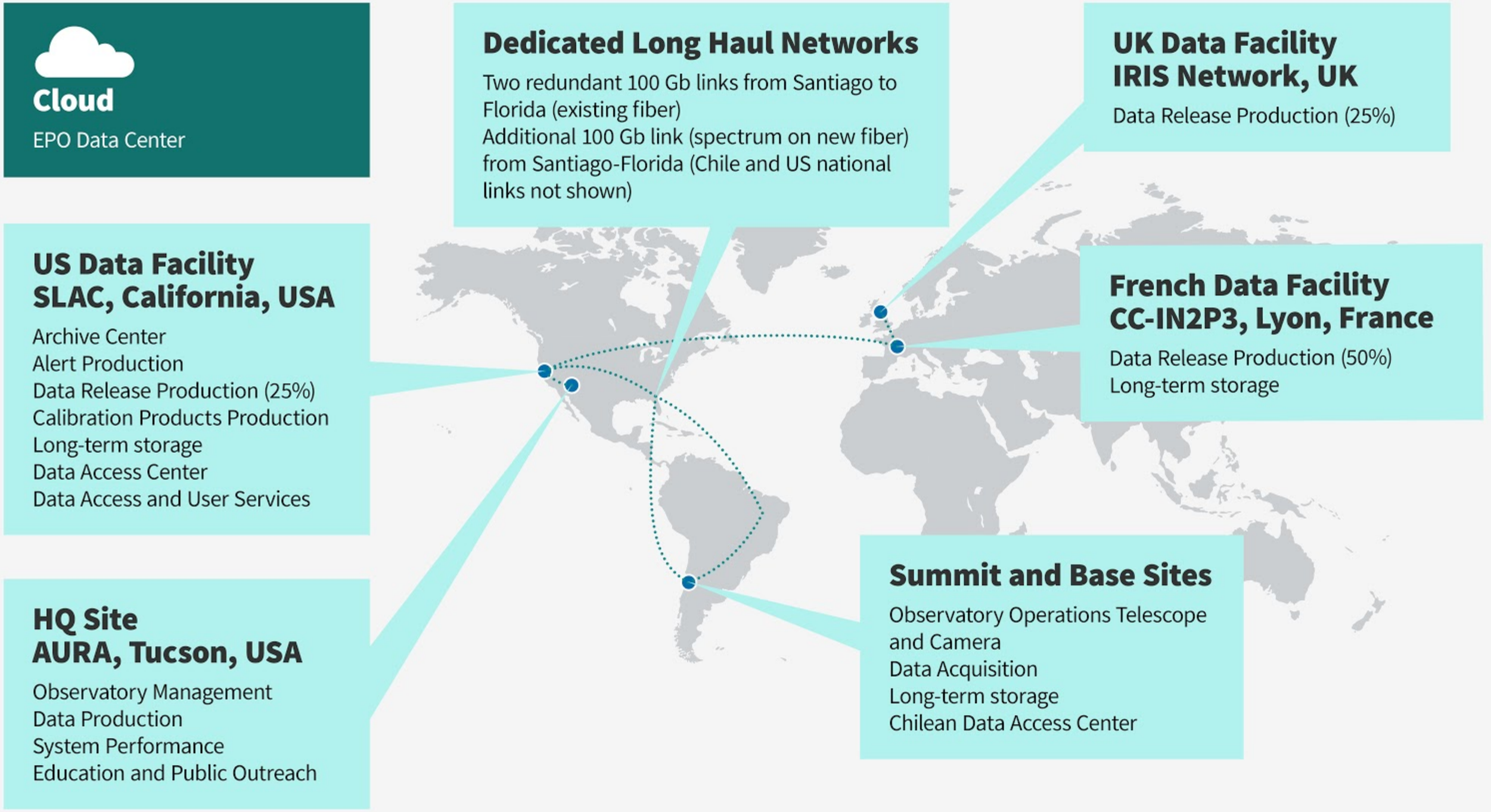
Aggregated data over 10 years of operations*, including derived data

image collection: ~6M exposures, 515 PB

final catalog database: 15 PB



* source: LSST Key Numbers


Cloud
EPO Data Center
Dedicated Long Haul Networks
Two redundant 100 Gb links from Santiago to Florida (existing fiber)
Additional 100 Gb link (spectrum on new fiber) from Santiago-Florida (Chile and US national links not shown)
UK Data Facility
IRIS Network, UK
Data Release Production (25%)
US Data Facility
SLAC, California, USA
Archive Center
Alert Production
Data Release Production (25%)
Calibration Products Production
Long-term storage
Data Access Center
Data Access and User Services
French Data Facility
CC-IN2P3, Lyon, France
Data Release Production (50%)
Long-term storage
HQ Site
AURA, Tucson, USA
Observatory Management
Data Production
System Performance
Education and Public Outreach
Summit and Base Sites
Observatory Operations Telescope and Camera
Data Acquisition
Long-term storage
Chilean Data Access Center

ESCAPE DATALAKE

- Replication of one night worth of LSST raw data, repeatedly over 5 consecutive days

  realistic dataset: 4000 exposures, 800k files, ~15 TB

  equivalent to 1 night of raw data in terms of volume and 2 nights in terms of number of files

  replication time budget: 12 hours

  driven by Rucio & FTS involving storage endpoints connected to ESCAPE datalake

  dataflow: CERN → CC-IN2P3 (RTT: 4 ms)

- Results

  replication of the entire dataset performed in less than 8 hours without errors

monit-grafana.cern.ch

Main Services / Rucio Stats (Replicas)

2021-11-22 00:00:00 to 2021-11-26 23:59:59

Overview

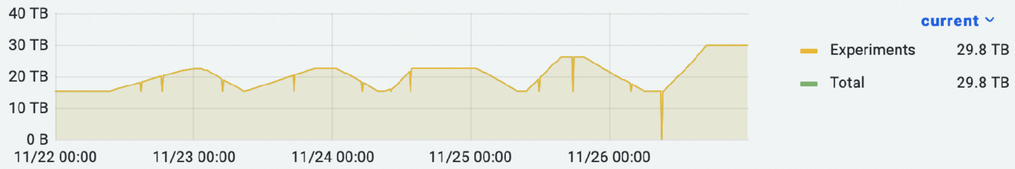
Storage used

29.85 TB

Total replicas

1.56 Mil

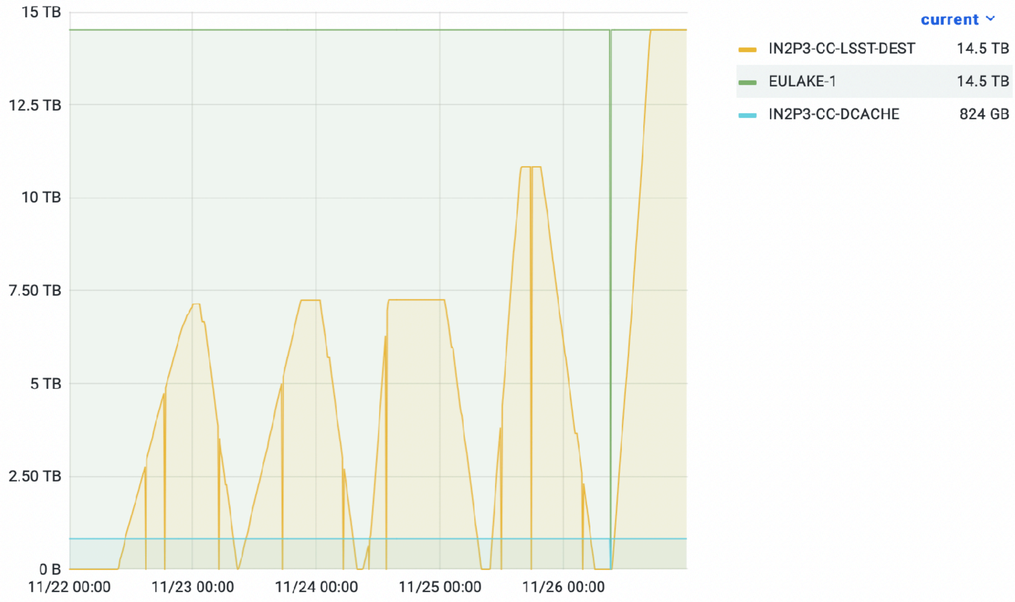
Storage Used



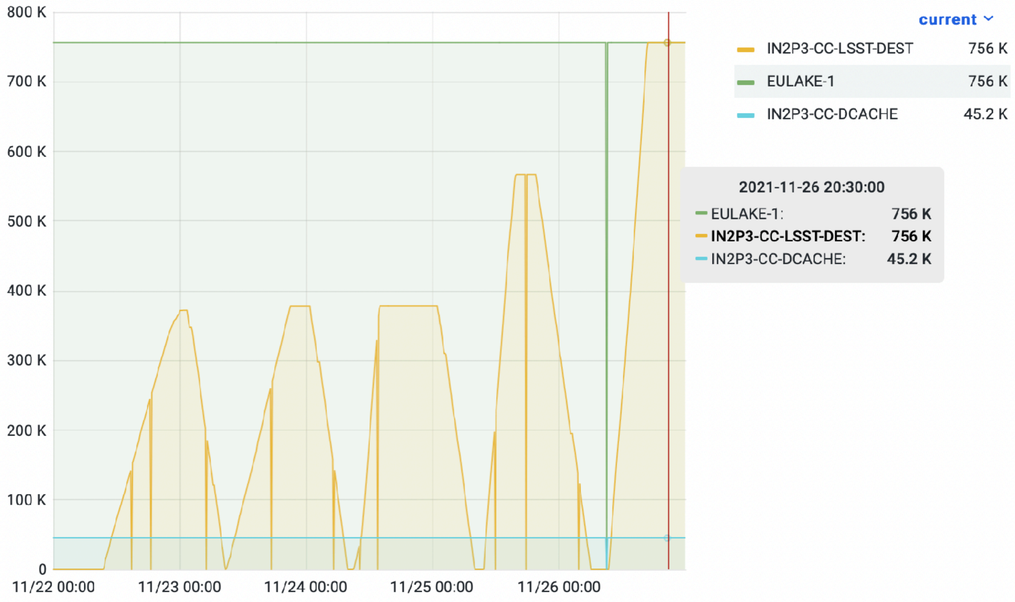
Total replicas

RSES

Used Storage / RSE



Replicas / RSE



SCOPES

REPLICATION USING RUBIN INFRASTRUCTURE

- Similar exercise over transatlantic link using same dataset

  data flow: SLAC National Accelerator Laboratory (US) → CC-IN2P3 (FR)

  replication driven by instances of Rucio & FTS deployed and operated by Rubin US data facility at SLAC

- Ongoing work

  identifying FTS parameters adequate for efficient transfer of high number of small files (~20 MB) over high-latency network links (RTT 150 ms)

  e.g. buffersize, number of streams, min active/max active, optimizer mode

  integration with the butler (the LSST I/O abstraction layer) and preservation of LSST file namespace in Rucio

ESCAPE BENEFITS

- ESCAPE data infrastructure has been instrumental

  for getting familiar with data management tools for evaluation purposes

  access to a ready-to-use, well-maintained, monitored, flexible infrastructure is an accelerator of adoption of those tools by science projects

- ESCAPE provides a forum to share experience

  understanding how other science projects use or intend to use the same tools is enlightening

  direct access to developers and operators of those tools is extremely valuable

  science projects get previews of the upcoming technologies and provide input to developers and operators on atypical use cases

See also

- Vera C. Rubin Observatory
  https://www.lsst.org
- Rubin-LSST France
  https://www.lsst.fr