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

Large Synoptic Survey Telescope

inputs to ESCAPE

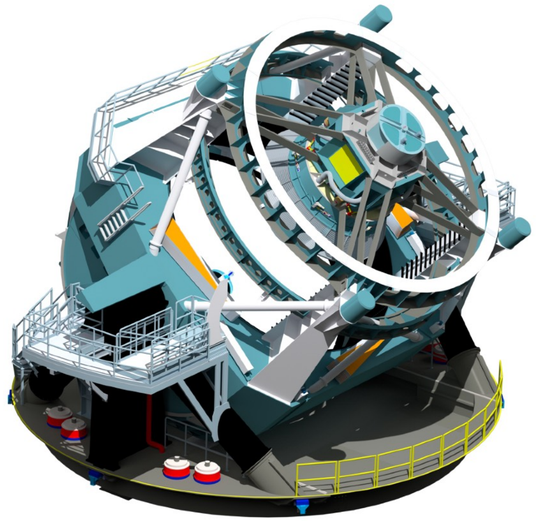
fabio hernandez

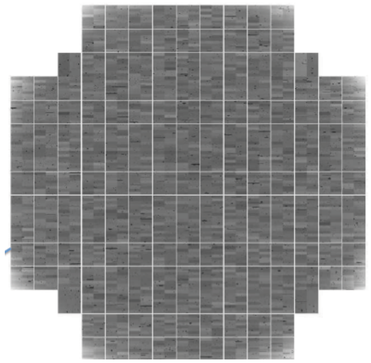
CCIN2P3 cnrs ESCAPE European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures

ESCAPE WP2/WP5 workshop, Amsterdam, July 2nd, 2019

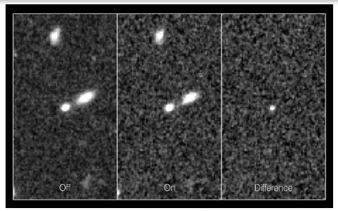
LSST OVERVIEW

LARGE SYNOPTIC SURVEY TELESCOPE

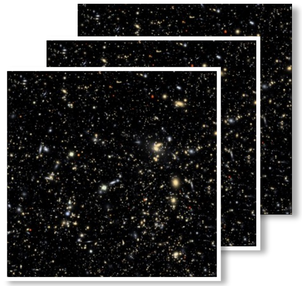




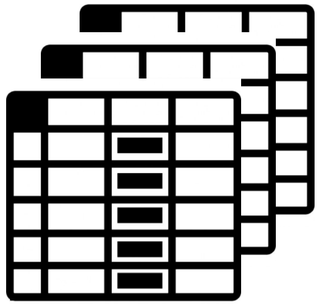
raw images



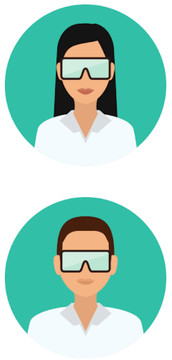
alerts



calibrated exposures
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 (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 2022 for **10 years

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

  each patch of the sky to be visited about 1000 times

- Science themes

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

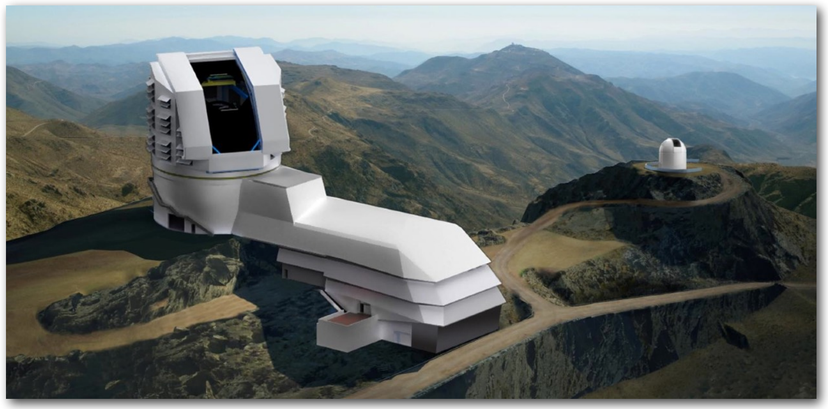
  *taking an inventory of the **solar system***

  *exploring the **transient** optical sky*

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

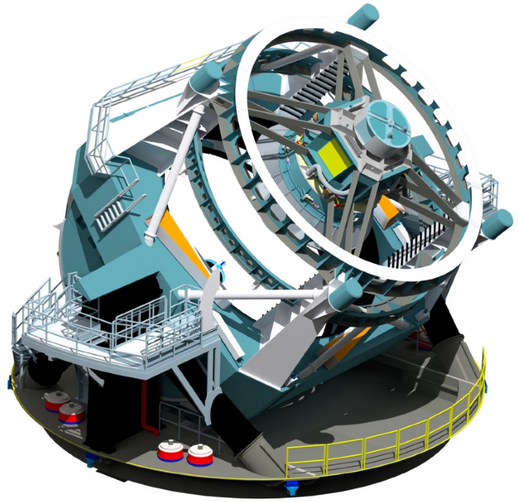
LSST OVERVIEW

OBSERVATORY



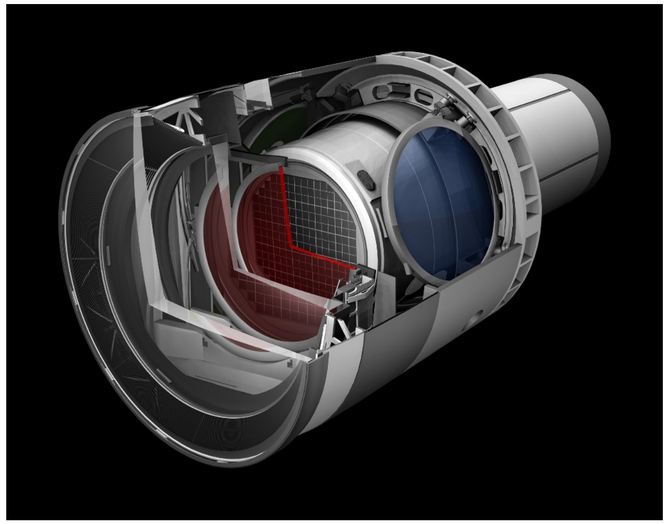
south hemisphere | 2647m a.s.l. | stable air | clear sky | dark nights | good infrastructure

TELESCOPE



main mirror ⌀ 8.4 m (effective aperture 6.5 m) | large aperture: f/1.234 | wide field of view | compact | 350 ton | 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

LSST DATA PRODUCTS

PROMPT: REAL-TIME DIFFERENCE IMAGE ANALYSIS (DIA)

NIGHTLY

Stream of 10M time-domain **events** per night, detected and transmitted to event distribution networks within 60 seconds of shutter close

Catalog of **orbits** for 6M bodies in the Solar System

DATA RELEASE: REDUCED SINGLE-EPOCH & DEEP CO-ADDED IMAGES, REPROCESSED DIA PRODUCTS

ANNUAL

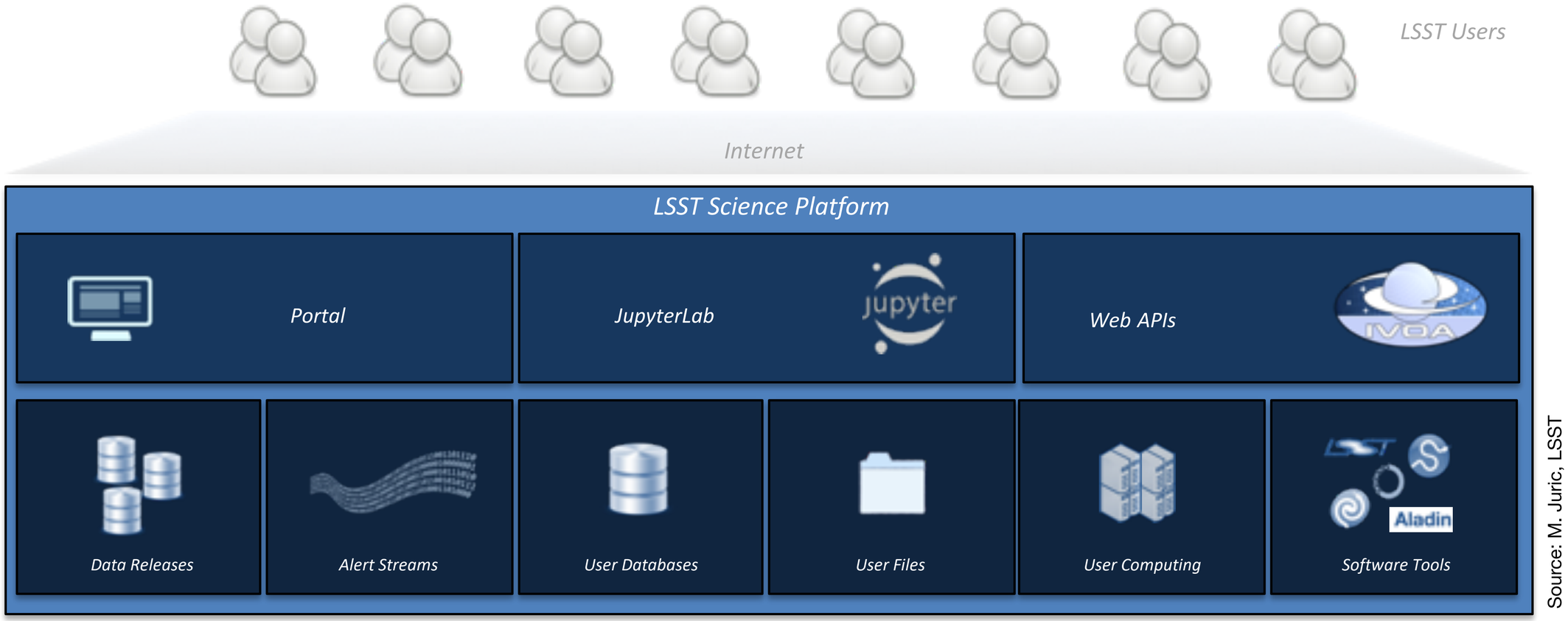
Catalog of 37B objects (20B galaxies, 17B stars), 7T observations, 30T measurements, produced annually, accessible through online databases

Deep co-added **images**

Source: LSST

User-generated data products not shown

LSST Science Platform



Source: M. Juric, LSST

Set of integrated web applications and services, through which the scientific community will access, **visualize**, subset and perform **next-to-the-data analysis** of the data

LSST AT CC-IN2P3

- Main roles

  satellite data release processing *under NCSA leadership*

  CC-IN2P3 to ***process 50% of the raw data***

  both NCSA and CC-IN2P3 will exchange and validate the data produced by the other party

  each site to ***host an entire copy of both raw and reduced data****, i.e. the products of the annual data release processing (images and catalog)*

Large Synoptic Survey Telescope

LSST Operations: Sites & Data Flows

HQ Site
Tucson, AZ

Science Operations
Observatory Management
Education & Public Outreach

Base Site
La Serena, Chile

Base Center
Long-term storage (copy 1)

Data Access Center
Data Access & User Services

French Site
CC-IN2P3, Lyon, France

Satellite Processing Center
Data Release Production
Long-term Storage (copy 3)

LSST Data Facility
National Center for Supercomputing Applications (NCSA), Urbana-Champagne, IL

Processing Center
Alert Production
Data Release Production
Calibration Products Production
EPO Infrastructure
Long-term Storage (copy 2)

Data Access Center
Data Access and User Services

Summit Site
Cerro Pachón, Chile

Telescope & Camera
Data Acquisition
Crosstalk Correction

DATA RELEASE PROCESSING

- Data release processing pipelines
  - *set of stages for extracting information from images: detect astrophysical objects and their physical properties*
  - *produce the data to populate the astronomical catalog*
  - *C++ and Python 3*
- File size in the range 50 - 100 MB
  - *1 file per CCD (there are 189 CCDs in the focal plane)*
  - *currently FITS format*
  - *~10B files aggregated over the 10 years of operations (raw + derived)*
- High-level I/O abstraction layer designed to make life easier for scientists
  - *currently requires POSIX API and needs control the file namespace*
  - *ongoing work to improve this situation to include requirements for bulk processing*
- During annual release processing, only a few production accounts interact with the file catalog
  - *however, access to previous years' data releases covered by embargo only accessible by individuals with data rights*

Answers to Questionnaire

Data Production

- Estimated number of files: ~1billion / year, 10 years of operations

  this includes intermediary data that we don't keep forever, so real number may be less

- Data file sizes

  raw: exposures of 3.8 GB composed of 189 files of 20 MB each

  derived: calibrated exposures composed of 189 files of 110 MB each

  several other kind of products: ~100 MB each

- Primary data already compressed

- Primary data as well as the released data products to be archived

Data Model

- Raw data is collected at the summit where the instrument is located (Chile)

  1st copy stored at the base center in La Serena (Chile)

- Immediately transported to Champaign, IL (USA)

  2nd copy stored there

- Without significant delay, transported to Lyon (France)

  3rd copy stored there

- Raw data produced nightly: 20 TB, 300 nights per year

Data Model (cont.)

- Every year, all the data collected since the beginning of the survey is reprocessed and a new release of the data products is delivered
  both images and astronomical catalog
- Derived data, ready for analysis produced yearly
  images: ~10 PB

  astronomical catalog: ~10 PB

  user-derived data: ?? PB
- Data lifecycles
  a copy of all data products (image and catalogs) archived on tape

  2 more recent data releases are kept on disk
- Searchable metadata
  image metadata already stored in each image

  image registry contains excerpts of that metadata

Data Access and Data Processing

- Need for quasi-online data processing of primary data before analysis?

  yes, for alerts, but this will be taken care of separately

- Protocols to transfer data

  my preference HTTP (or anything else provided it is standard) including for making data downloadable for users

- Protocols for accessing data

  currently POSIX only

  future, object store + POSIX (on local file system to the worker node)

- Data caching for file re-usability ?

  yes, for exposing final products (images, not catalog) to partner sites or even end-users

- Read-ahead cache for latency hiding purposes ?

  maybe, a system that understands the file format (e.g. image vs. tables) could be useful

  not for accessing the catalog database

- QoS for storage ?

  I don't know, I would say throttling data transfer to different sites and end-users would be useful

Data Access and Data Processing (cont.)

- File popularity management service ?
  - *collecting information on the popularity of files would be useful, but I'm not sure if we could automatically take decisions on the data placement based on that information*
  - *how to articulate this mechanism with the I/O abstraction layer on top of which all the LSST software is based ?*
- Data access patterns
  - *copy the data to the local worker node storage ? possible, but also using a network file system (within the same site)*
  - *consume full contents of files ? I have no figures yet, but I guess yes, but not necessarily sequentially (FITS files)*
  - *evaluation of the impact of remote depending on RTT and bandwidth ? not considered so far*
  - *processing of data stored on tape ? scheduled recall campaigns, no direct access to tape*
- Workload management system
  - *at the level of a single site, yes. At the level of the experiment, not clear. Static partitioning of the workload may be possible*
  - *data placement probably managed somehow separately from data processing*

Data Access and Data Processing (cont.)

- CLIs, APIs and/or Web Interfaces?

  I would say that programming language-neutral APIs is the minimum, so that CLIs and web interfaces can be built, if not included in the products

- MPI jobs ?

  currently using MPI as a mechanism for exploiting several CPU cores in a single node

  not tested MPI in a multi-node setting, but this would probably work

- Event-driven data processing ?

  yes: triggering actions based on events emitted by the storage system, e.g. for shipping data products to a partner site, pre-processing images just imported, etc.

- Impact of data loss

  don't want to loose raw data nor released data products, hence at least 3 copies around the world

  temporarily unavailable data acceptable, depending on the length of the period

Data Access Control

- Annual data releases only available to project members during the first 2 years after publication

  then publicly available to anyone

- Anonymous access to data products after the embargo period

  the mechanisms to make the data publicly available still to be precisely defined

- ACL needs from a global perspective

  raw data: read-only except for a few production accounts

  data products: read-only for project members only

  during embargo period, data accessible only to ***authorized members of the project***

- Groups with privileges ?

  production users vs. project members vs. anonymous

AAI QUESTIONS

- Is there an identified AAI expert ? yes
- Number of active users: > 1000
- End-users will access data through notebooks (including terminal sessions) and web-based applications

  I don't think there will be LSST-specific native applications, but I suspect users will want to use their preferred image visualisation tool

  users may want to transparently access published data (e.g. images) from their comfort of their laptop, for instance for visualisation purposes (likely not for bulk processing) ⇒ caching (?)

- Current authentication tools

  CILogon, ssh keys, X.509 certificates (web application), Kerberos tickets

- User registration and lifecycle management for experiment

  not clear for me: I suspect that currently this is handled by each site independently, but eventually a central registration will be needed, or at least, a central directory service where users with data rights will be registered

- Authorization structure

  access to data is currently controlled by what the file system provides (uid/gid, ACLs)

Questions & Comments

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

BACKUP SLIDES

IN2P3

A DISTRIBUTED LABORATORY

in2p3.fr

2500 researchers, engineers and technicians

700 post-docs and PhD students

25 laboratories and research platforms in France, 16 international laboratories

COMPUTING CENTER

IN2P3 COMPUTING CENTER

- CC-IN2P3

  84 people, 80 FTE, 80% permanent positions

  ~15 M€ overall annual budget

  scientific data center, high throughput computing

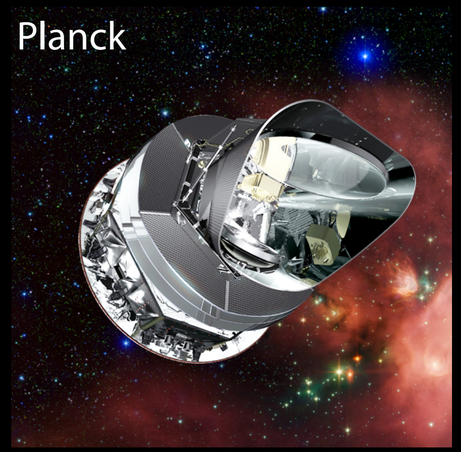
  well connected to national and international networks

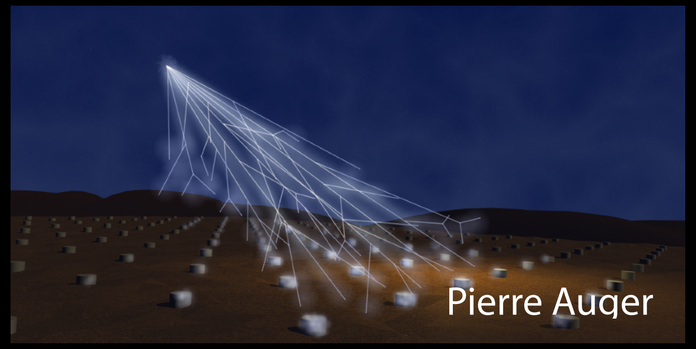
- Shared computing facility supporting the institute's research program

  ~70 projects in high energy physics, nuclear physics and astroparticle physics

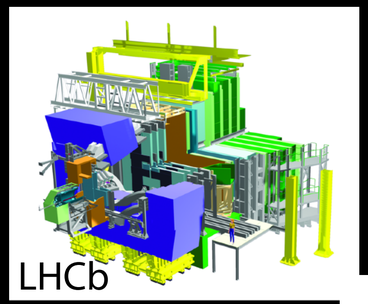
- Operations: 24x7

  unattended during nights and weekends

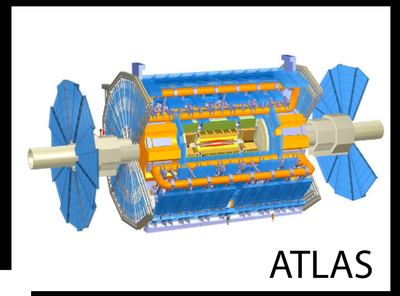


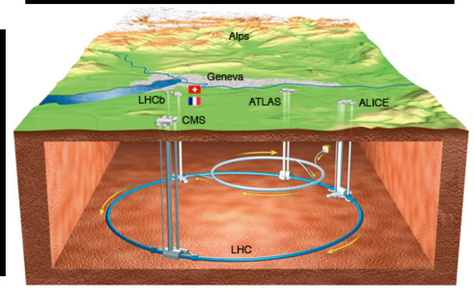


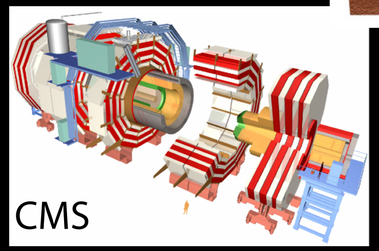

SuperNova Legacy Survey

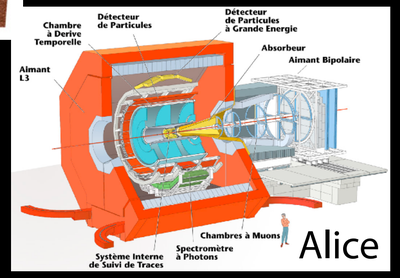


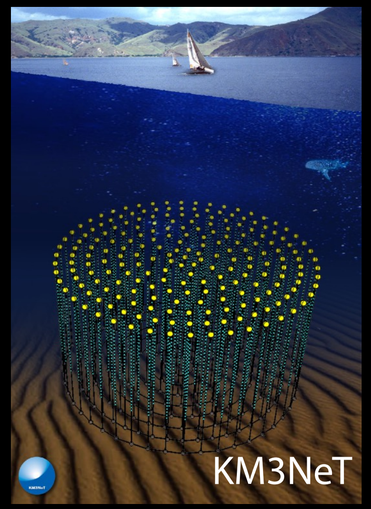
LHC @ CERN



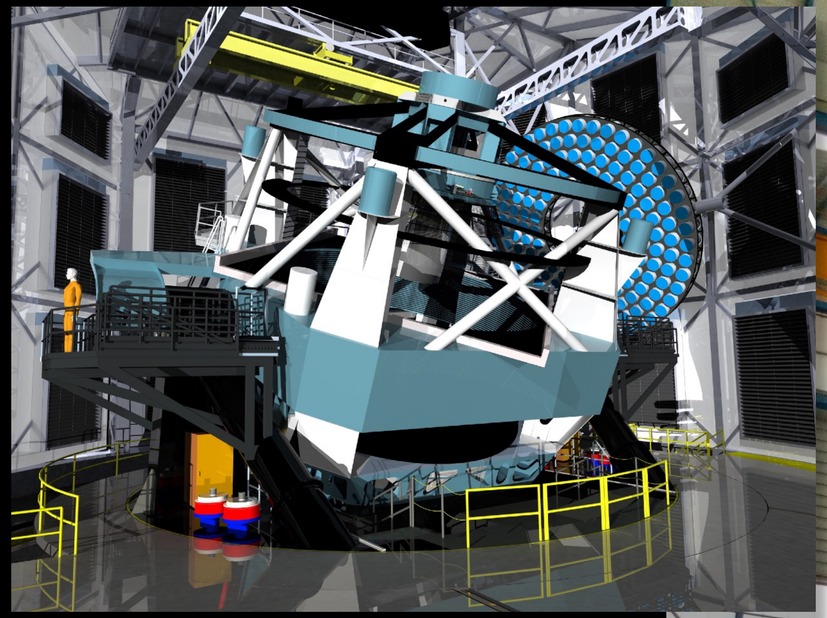






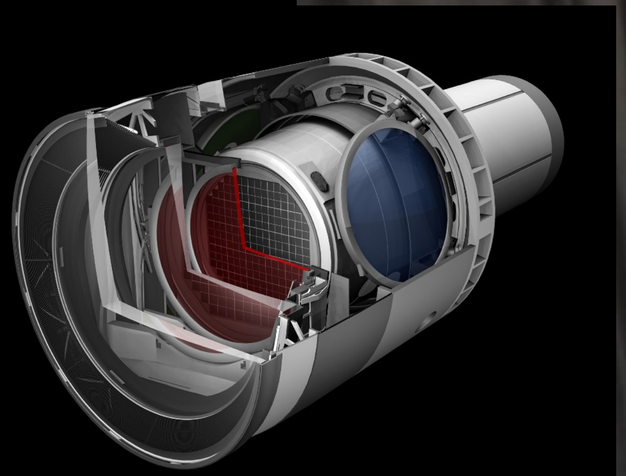


Sources: LSST, Gianluca Lombardi (2018-08)



ROTATION: 10 deg/s
3.5° SLEW-SETTLE: 5 s

Sources: LSST, Asturfeito (2018-03)



LSST DATA MANAGEMENT CONTRIBUTORS

Princeton University

UNIVERSITY of WASHINGTON

National Optical Astronomy Observatory

SLAC National Accelerator Laboratory
Stanford University

Infrared Processing and Analysis Center
California Institute of Technology

DATA RELEASE PROCESSING CENTRES

National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign

CNRS / IN2P3 computing center

LSST DATA MANAGEMENT SUBSYSTEM

- Archival

  *to **record**, **transport** and permanently **store raw data** issued by camera*

- Processing

  *to **detect transients** and **emit alerts** within 60 seconds after observation*

  *once per year, to **produce a data release**: a **self-consistent, immutable dataset**, composed of **processed data since the beginning of the survey***

  *to **develop the software** necessary for processing the data: image processing algorithms (calibration, point spread function, co-addition of images, characterization of objects, processing pipelines, …), catalogue database, middleware (workload management, orchestration, …), data transfer, etc.*

- Publication

  *to **deliver the reduced data** (images + catalogs)*

  to facilitate custom data reduction and individual data analysis

Data Acquisition

- Raw data

  7.2 GB per image

  2000 science images + 450 calibration images per night

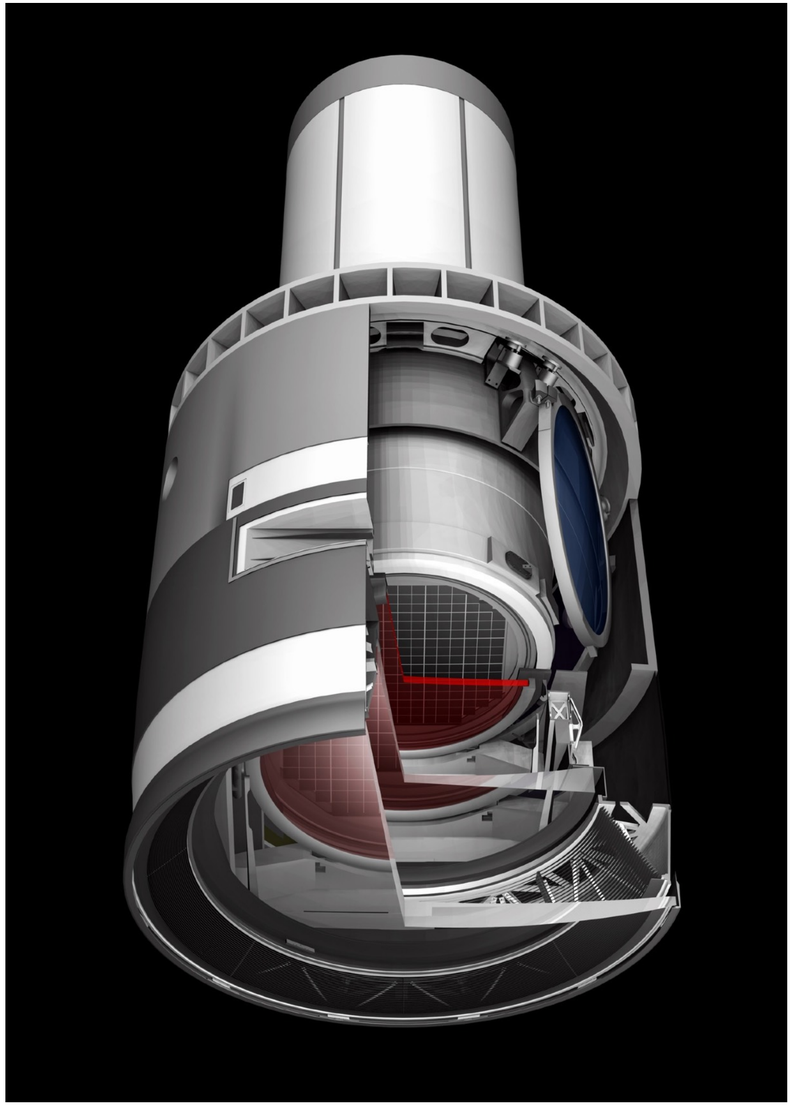
  300 nights per year, ~20 TB per night ⇒ ~6 PB per year

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

  images: ~6M exposures, 515 PB

  final catalog database: 15 PB

  ** source: LSST key numbers*



Source: LSST