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

LSST data processing at CC-IN2P3

status and perspectives

fabio hernandez

CCIN2P3

CC-IN2P3 seminar, Lyon, March 16th, 2018

cnrs

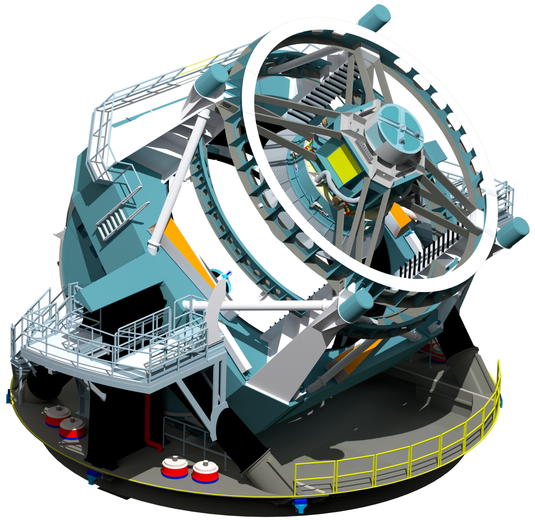
lpsc

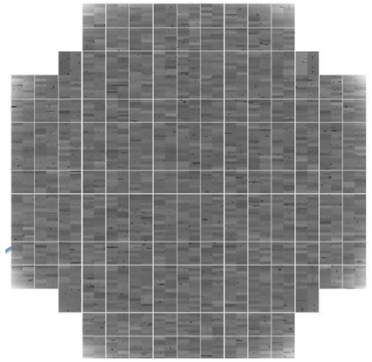
Contents

- LSST project overview
- LSST at IN2P3
- Planning
- Ongoing and foreseen technical work
- Summary

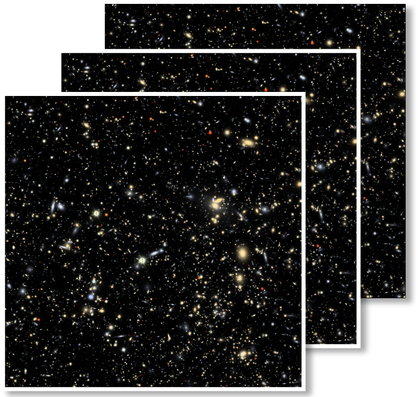
Project overview

LARGE SYNOPTIC SURVEY TELESCOPE

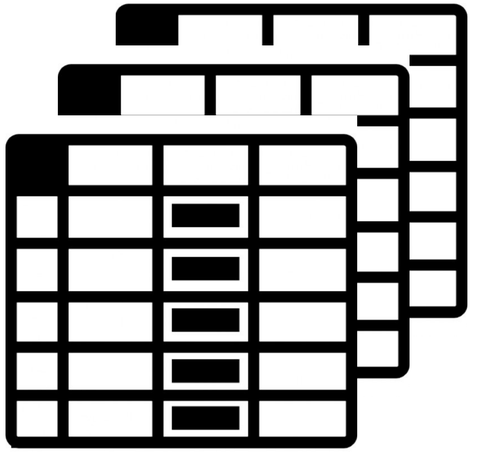




raw data



processed images



astronomical catalog
(stars, galaxies, objects, sources, transients, exposures, etc.)

LSST will build a catalog of 20 billion galaxies and 17 billion stars and their associated physical properties

LSST OVERVIEW (CONT.)

- 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***

- Principle of operations

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

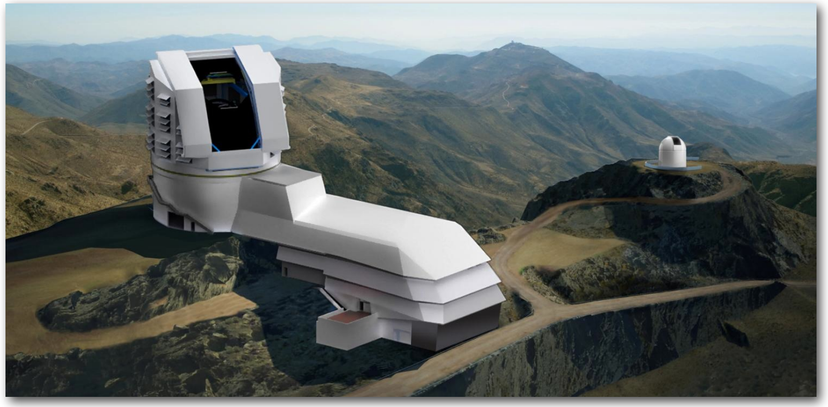
  one complete visit of the southern hemisphere sky every 3-4 nights**, from 2022 for **10 years

  each patch of the sky to be visited about 1000 times

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

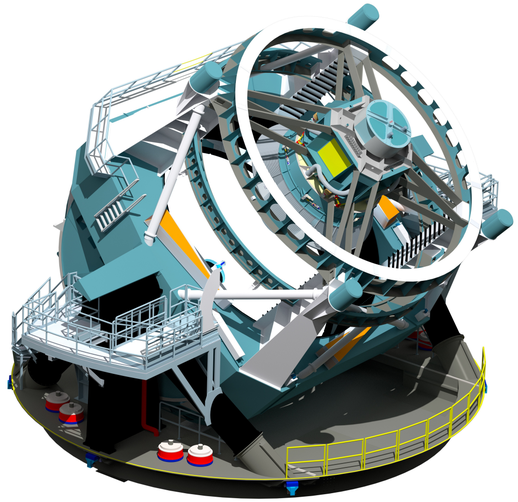
LSST OVERVIEW

OBSERVATORY



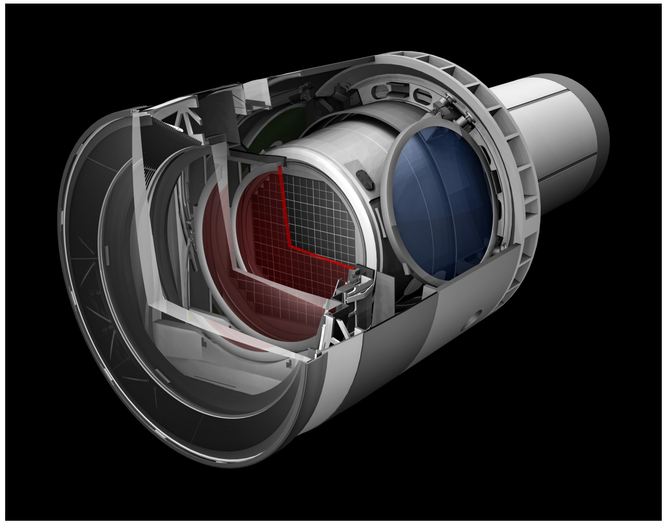
south hemisphere | 2647 m a.s.l. | stable air | clear sky | dark nights

TELESCOPE

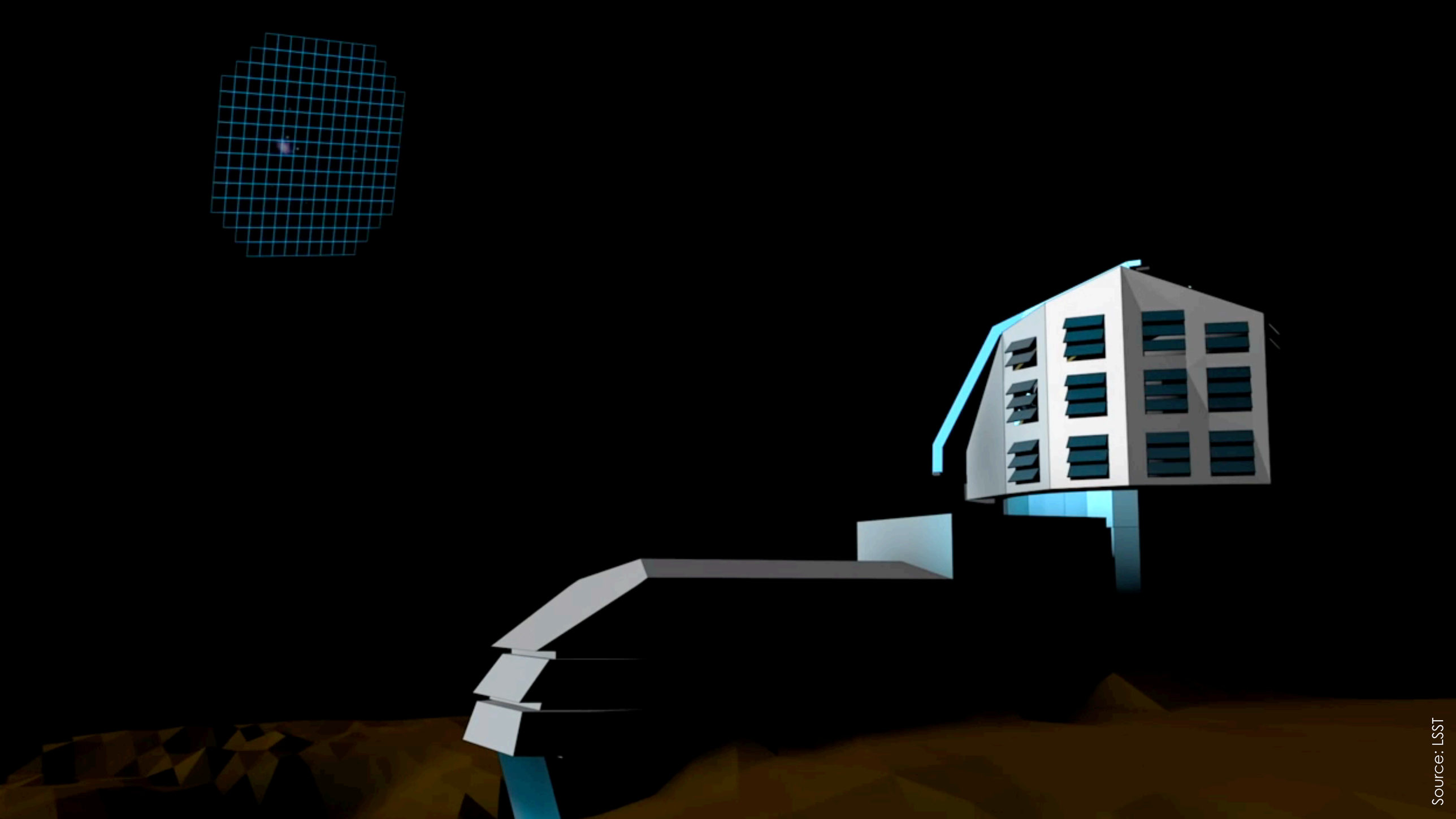


large aperture | wide field of view | compact | 350 ton | to be repositioned about 3M times over 10 years of operations | main mirror ⌀ 8.4 m (effective aperture 6.5 m) | f/1.234

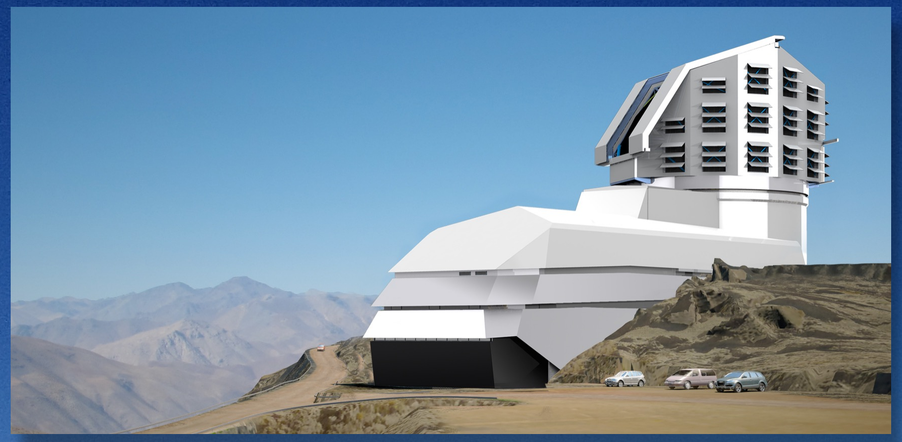
CAMERA



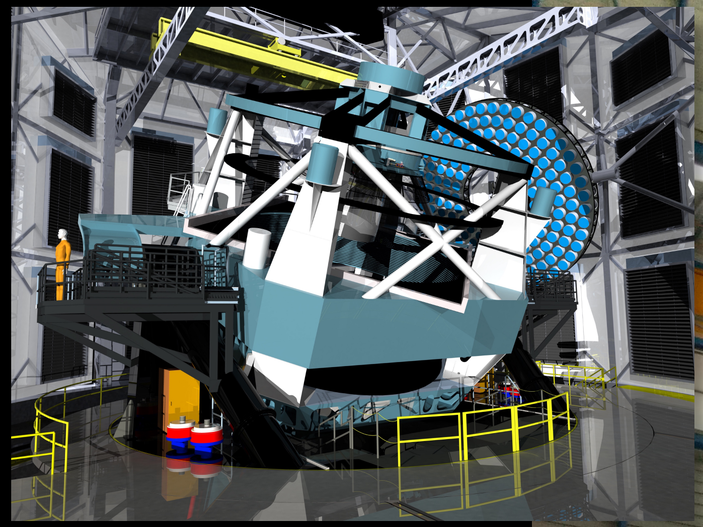
⌀ 1.65 m | 3.7 m long | 3 ton
3.2 G pixels | 3 lenses | 5 embedded filters
3.5° field of view | 9.6 degree^2
focal plane and electronics in cryostat at 173K



Source: LSST



Source: LSST (2018-02)



Source: LSST (2017-12)

LSST OVERVIEW: DELIVERABLES

- Deliverable

  *the science-enabling, **ultimate deliverable** of the project will be the **fully reduced data***

  the scientific exploitation of the processed data will be performed by the scientific community

- Open data

  *complete **cumulative data set** (images and catalogs), open to the scientific community of the participating countries, once per year, with no proprietary period*

  ***alerts** of detected variable sources (transients) made available for world-wide distribution **within 60 seconds of observation**, published via standard protocols*

- Open source software: `github.com/lsst`

LSST OVERVIEW: FUNDING AND BUDGET

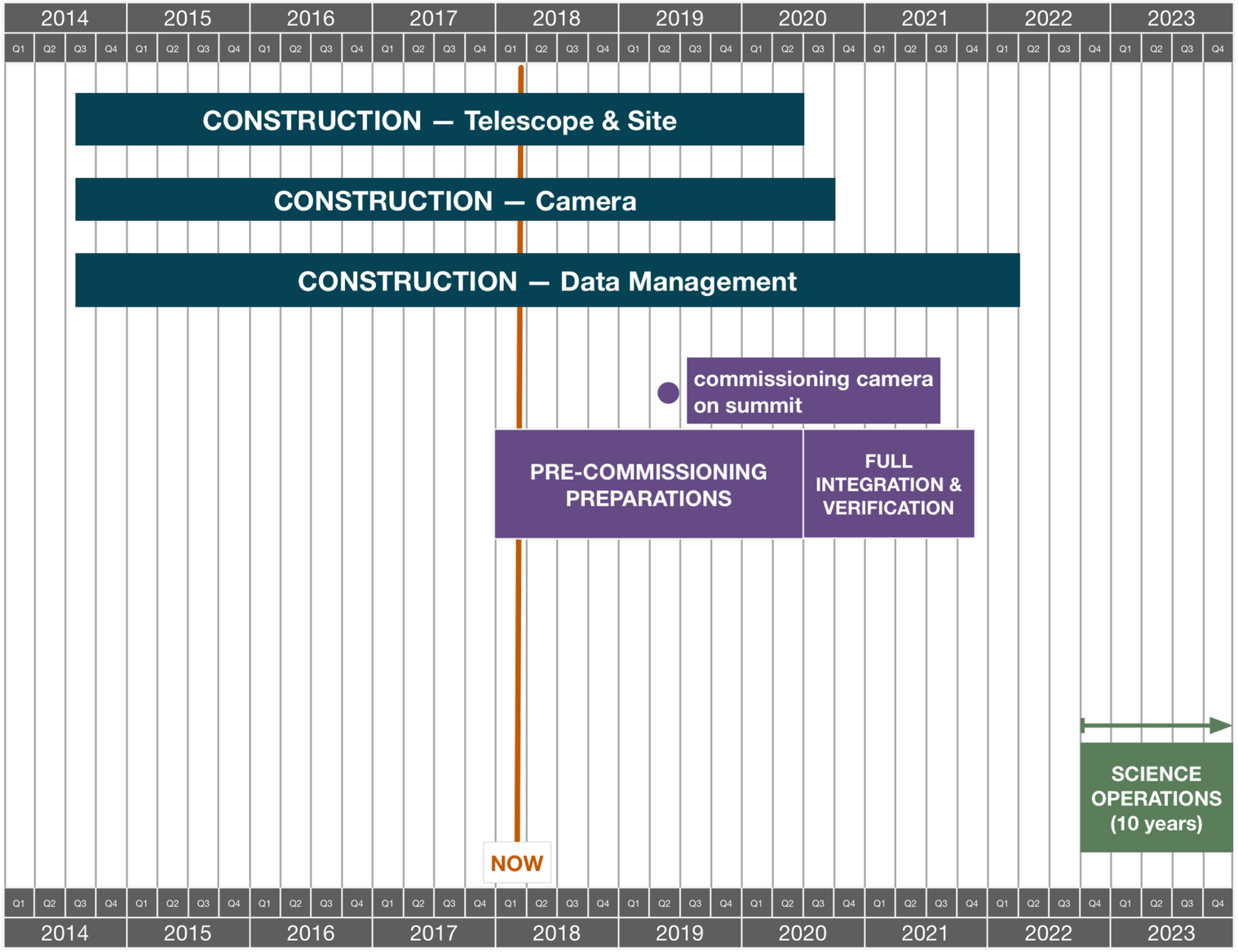
2014-2022 — Construction phase budget: US$ 671M

About 20% of the construction budget devoted to the DATA MANAGEMENT subsystem

2019-2034 — Operations phase budget: US$ 41M/year

International collaboration: 25 countries, 39 research institutions

LSST OVERVIEW: SCHEDULE AND STATUS



Description	% complete*
Data management construction	36 %
Telescope and site construction	69 %
Camera	77 %
Education and public outreach construction	18 %
Systems engineering and commissioning	22 %

** as of 2018-01-31*

Adapted from V. Krabbendam

LSST AT IN2P3

LSST AT IN2P3

- IN2P3 contributes to the construction of the LSST camera

  CCD electronics, filter carousel and changer (design, construction, command and control software)

  sites: APC, CCPM, LAL, LAPP, LMA, LPC, LPNHE, LPSC, LUPM

- IN2P3 formally committed to contribute to offline data processing

  site: CC-IN2P3

INSTITUTIONAL COMMITMENT

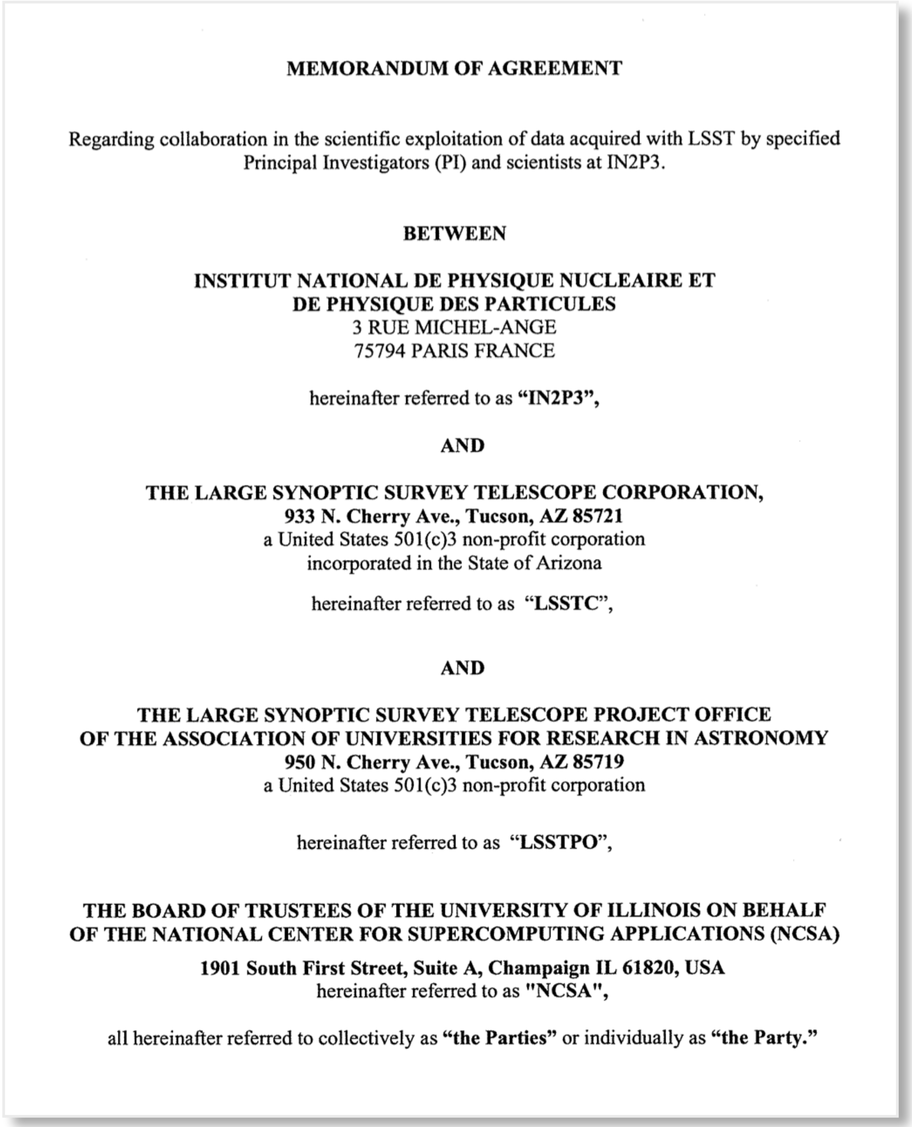
- Formal agreement between IN2P3, LSST Corporation and NCSA

  signed in 2015

  *contribution by IN2P3 to the **annual LSST data release processing** during the operations phase*

  computing equipment and labor

  in exchange of data rights



MEMORANDUM OF AGREEMENT

Regarding collaboration in the scientific exploitation of data acquired with LSST by specified Principal Investigators (PI) and scientists at IN2P3.

BETWEEN

INSTITUT NATIONAL DE PHYSIQUE NUCLEAIRE ET DE PHYSIQUE DES PARTICULES
3 RUE MICHEL-ANGE
75794 PARIS FRANCE

hereinafter referred to as **"IN2P3",**

AND

THE LARGE SYNOPTIC SURVEY TELESCOPE CORPORATION,
933 N. Cherry Ave., Tucson, AZ 85721
a United States 501(c)3 non-profit corporation
incorporated in the State of Arizona

hereinafter referred to as **"LSSTC",**

AND

THE LARGE SYNOPTIC SURVEY TELESCOPE PROJECT OFFICE OF THE ASSOCIATION OF UNIVERSITIES FOR RESEARCH IN ASTRONOMY
950 N. Cherry Ave., Tucson, AZ 85719
a United States 501(c)3 non-profit corporation

hereinafter referred to as **"LSSTPO",**

THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS ON BEHALF OF THE NATIONAL CENTER FOR SUPERCOMPUTING APPLICATIONS (NCSA)

1901 South First Street, Suite A, Champaign IL 61820, USA
hereinafter referred to as **"NCSA",**

all hereinafter referred to collectively as **"the Parties"** or individually as **"the Party."**

ATRIUM-216220

LSST DATA PROCESSING

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

LSST DATA PRODUCTS

Nightly

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

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

Annual

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

Deep co-added **images**

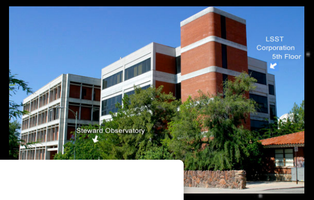
On demand

Services and computing **resources** to enable user-specified custom processing and analysis

Software and APIs enabling development of analysis code

Source: LSST

LSST DATA CENTERS



HEADQUARTERS SITE

HQ facility
- *observatory management*
- *science operations*
- *education & public outreach*

ARCHIVE SITE

Archive center
- *alert production*
- *data release production*
- *calibration products production*
- *long-term storage (copy 2)*
- *education & public outreach*
- *infrastructure*

Data access center
- *data access and user services*

SATELLITE RELEASE PRODUCTION SITE

Archive center
- *data release production*
- *long-term storage (copy 3)*

BASE SITE

Base facility
- *long-term storage (copy 1)*

Data access center
- *data access and user services*

SUMMIT SITE

Summit facility
- *telescope & camera*
- *data acquisition*
- *crosstalk correction*

Sources: LSST, Google

LSST DATA MANAGEMENT CONTRIBUTORS

Stanford Linear Accelerator Center

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

Princeton University

Infrared Processing and Analysis Center
California Institute of Technology

UNIVERSITY of WASHINGTON

National Optical Astronomy Observatory

IN2P3 / CNRS computing center

LSST AT CC-IN2P3

- Main roles

  satellite data release production *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

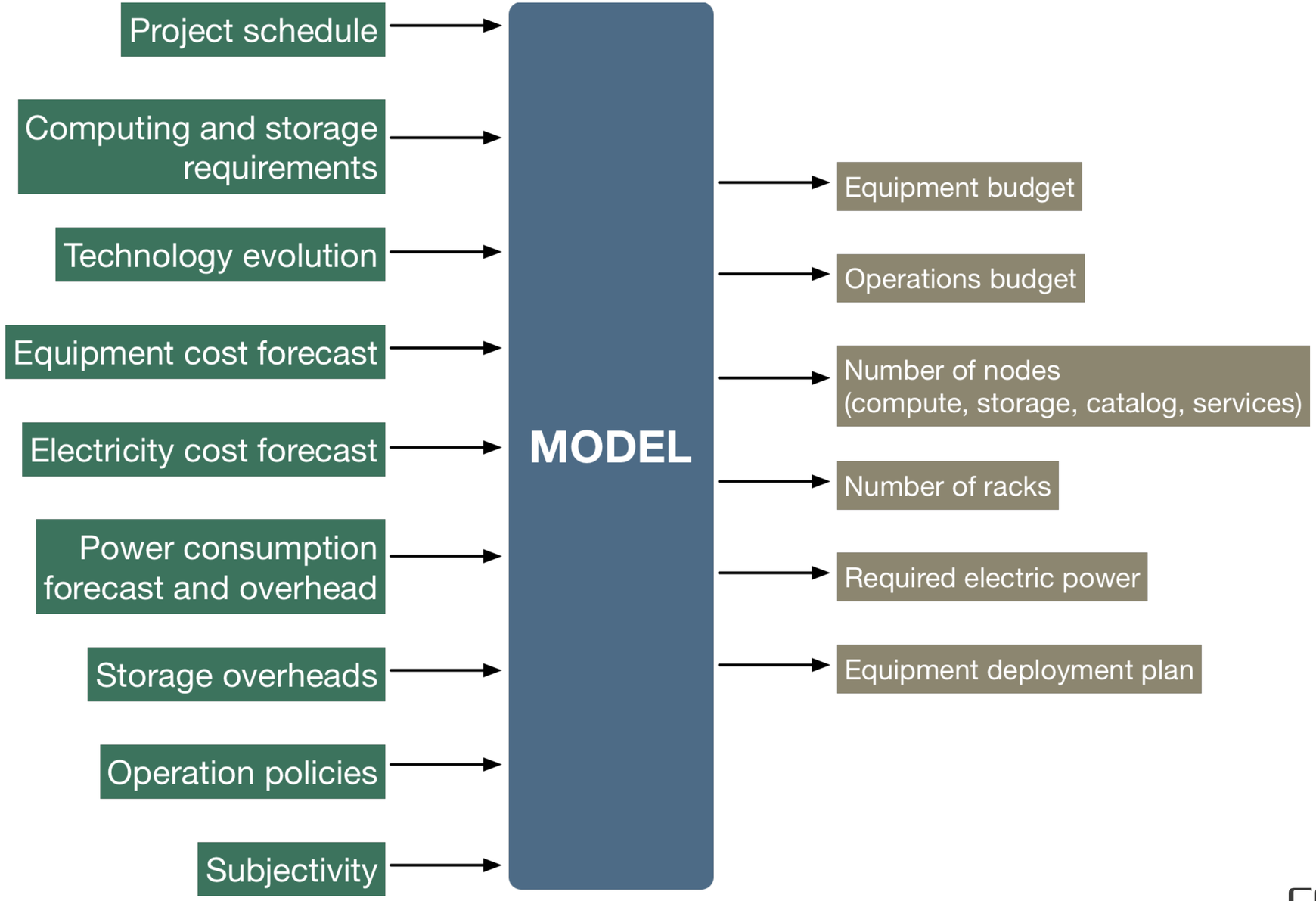
  each site to ***host an entire copy of both raw and reduced data****, i.e. the products of the annual data release processing (images and catalog)*

PLANNING

PLANNING

- Required equipment, estimate budget and initial deployment plan established
  - *covers period **2018-2031***
  - *mostly for **data release processing** but it does include equipment for catalogue database*
- Delivered and reviewed in 2017Q4
  - *many thanks to M. Betoule (LPNHE), F. Chollet (LAPP) and G. Rahal (CC-IN2P3) for their detailed work and very useful feedback*
- Documents
  - *model and companion explanation: ATRIUM-215611*
  - *review report: ATRIUM-280394*
  - *the model in ATRIUM does not yet include the reviewers' recommendations*

PLANNING: MODEL

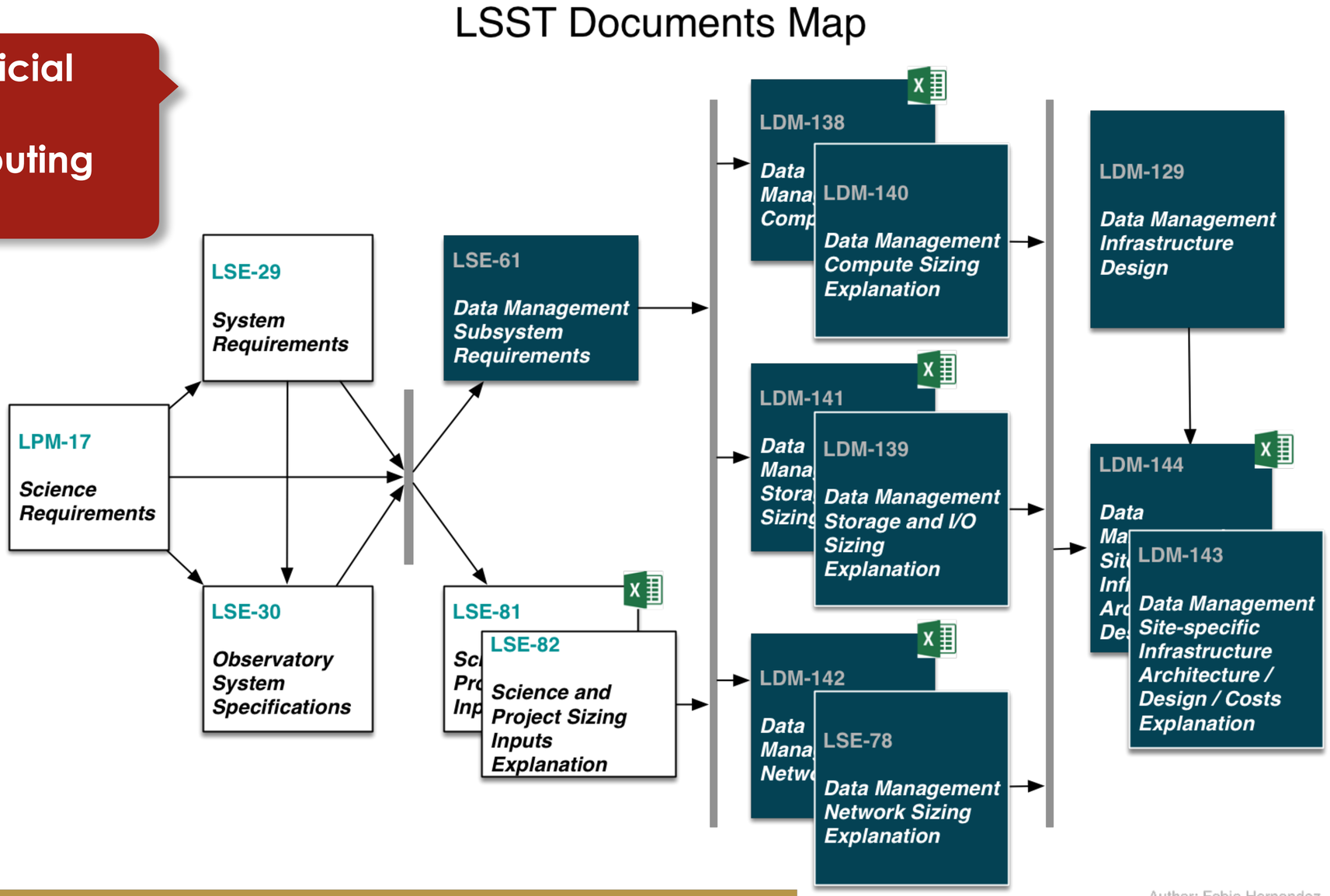


“All models are wrong; some models are useful”

George Box

PLANNING: REQUIREMENTS

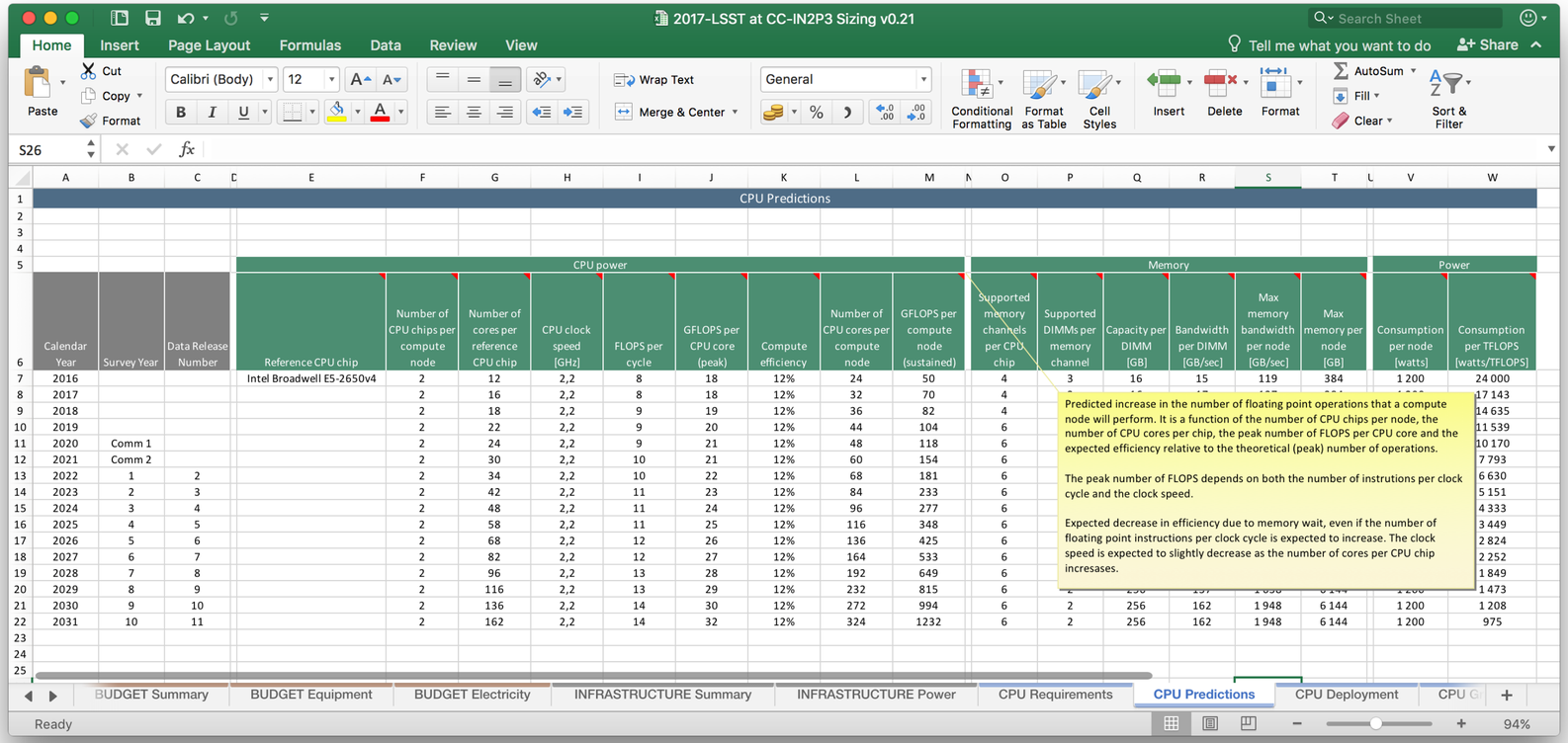
Subset of the LSST official documents used for estimating the computing requirements



To retrieve e.g. LDM-144 use the URL : https://docushare.lsstcorp.org/docushare/dsweb/Get/LDM-144

Author: Fabio Hernandez
Last update: 2018-03

PLANNING: MODEL



ATRIUM-215611

LSST SATELLITE DATA RELEASE PROCESSING AT CC-IN2P3

Required Equipment, Estimate Budget and Deployment Plan

WARNING: this document is under review and has not yet been approved

Fabio Hernandez, CC-IN2P3
fabio@in2p3.fr

November, 2017

This document presents the main assumptions made for estimating the computing resources, their cost and deployment schedule, necessary at CC-IN2P3 for performing its share of the annual data release processing of the Large Synoptic Survey Telescope, over the period 2018-2032. It is intended to serve as a companion document to the model itself which is implemented as a Excel workbook file associated to this document and located at the same address.

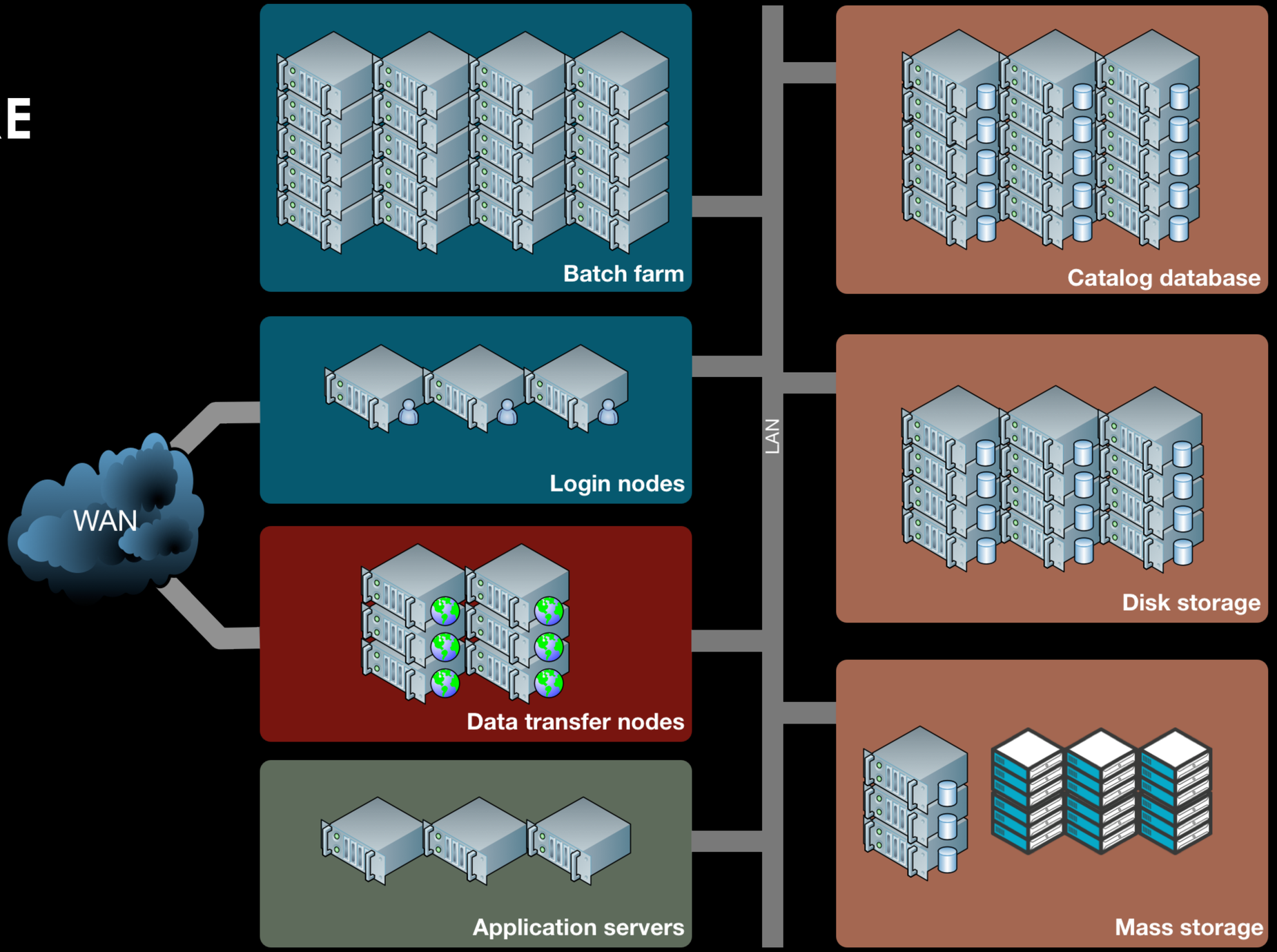
Document ID[1]: ATRIUM-215611

NOTE: This document accompanies v0.18 of the Excel file.

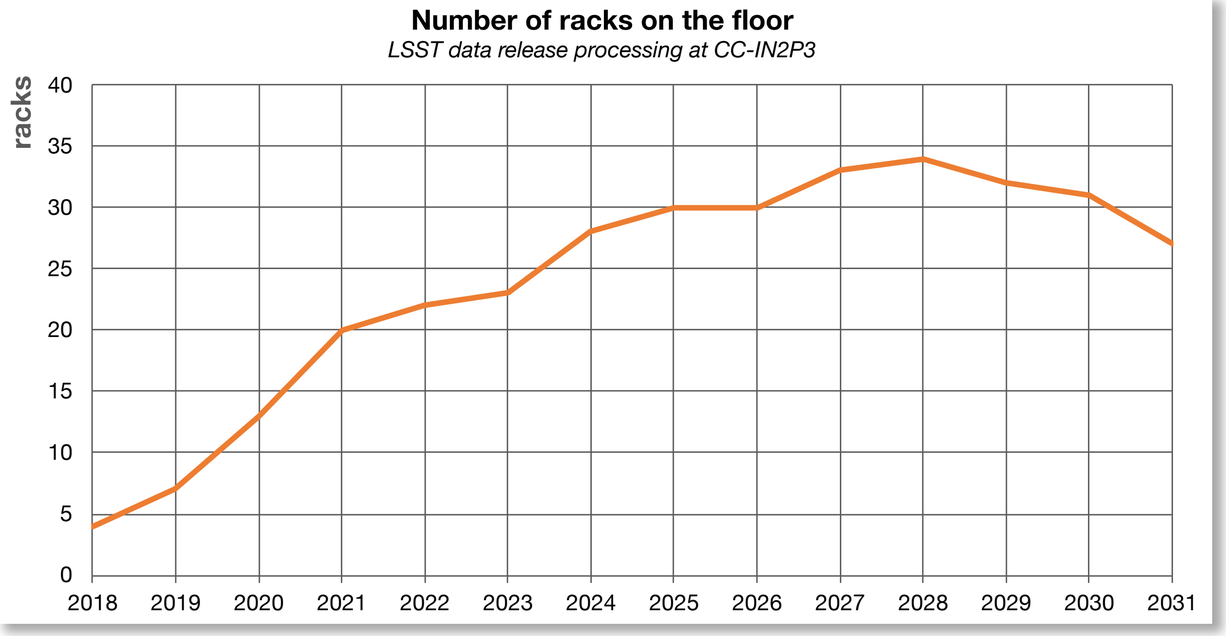
Contents

[1] This document source location: http://gitlab.in2p3.fr/fabio/lsst-drp-sizing

Envisioned architecture

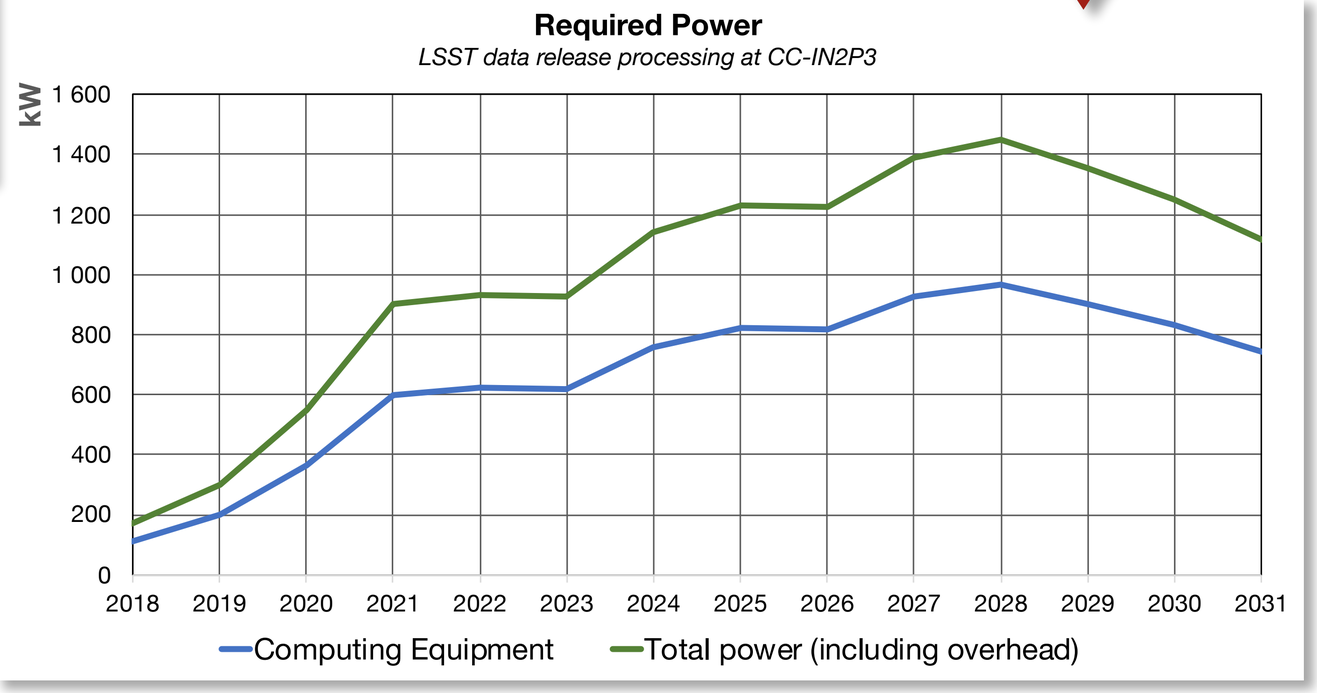


PLANNING: MACHINE ROOM INFRASTRUCTURE

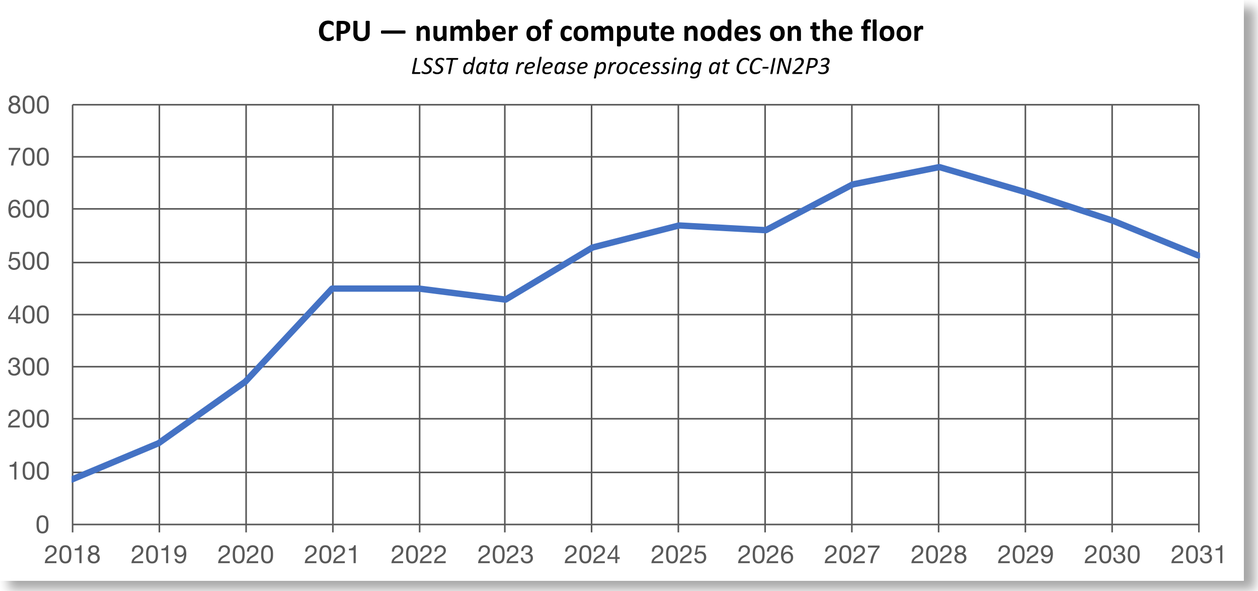


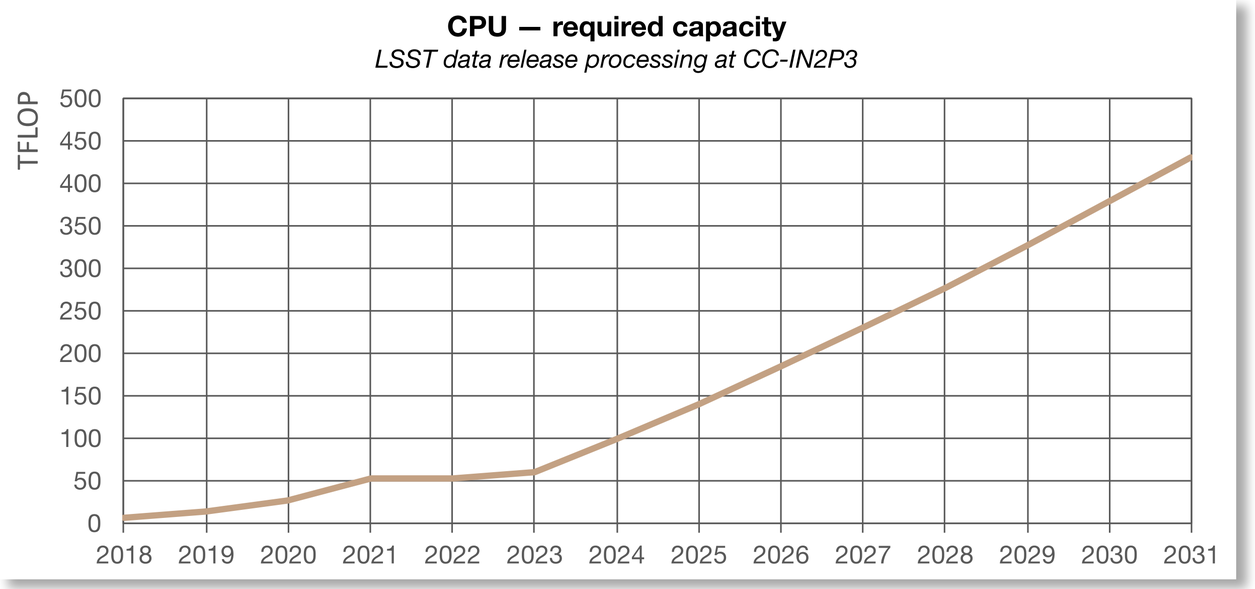
Racks
peak 35 racks

Power
peak 1.4 MW



PLANNING: COMPUTE





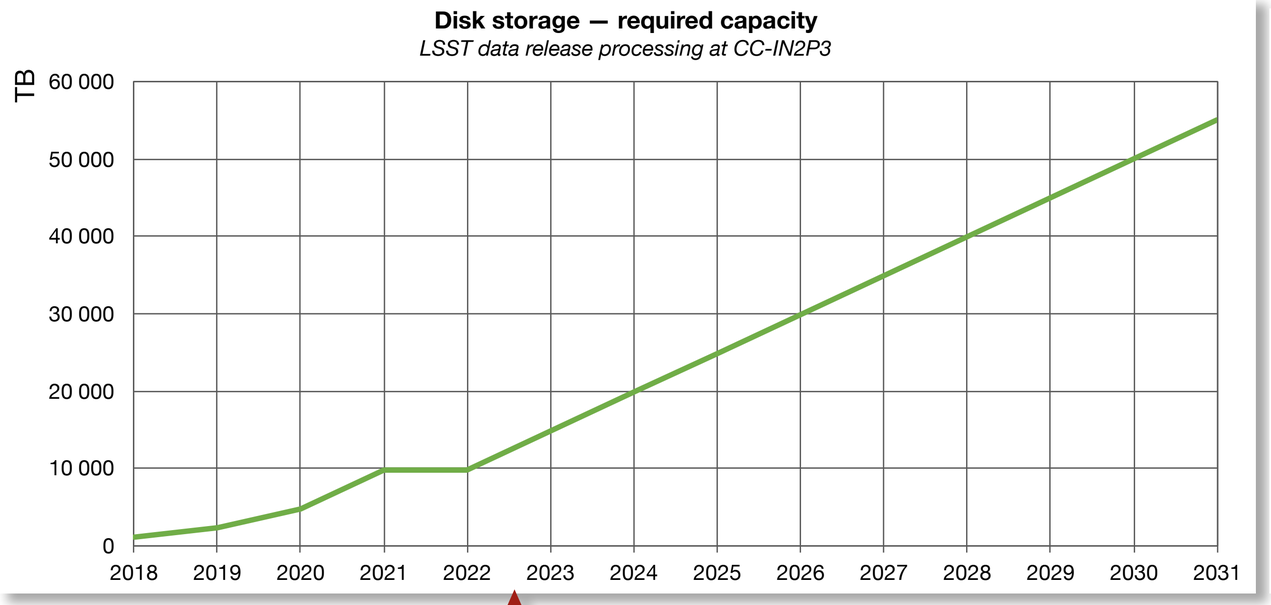
CPU nodes: peak 681
CPU cores: peak 122 k
Racks for CPU: peak 15

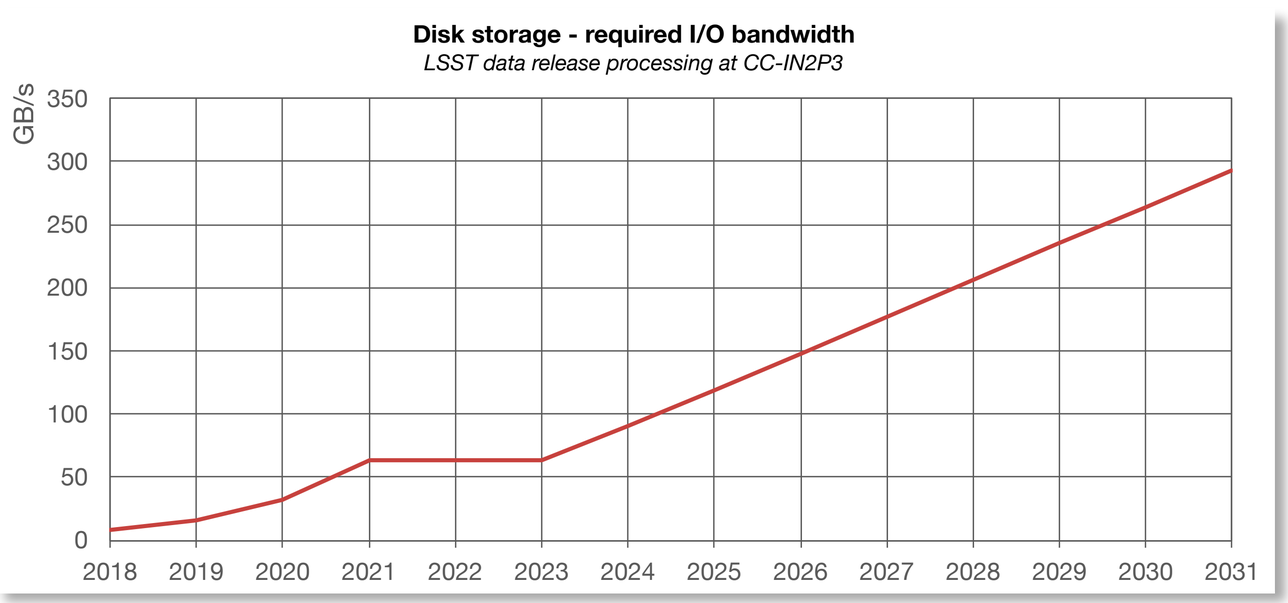
Size of CC-IN2P3 batch farm*:

CPU nodes: 866
CPU cores: 38 k
CPU power: 293 TFLOP (381 k HS06)

** as of Mar. 11th, 2018*

PLANNING: DISK STORAGE





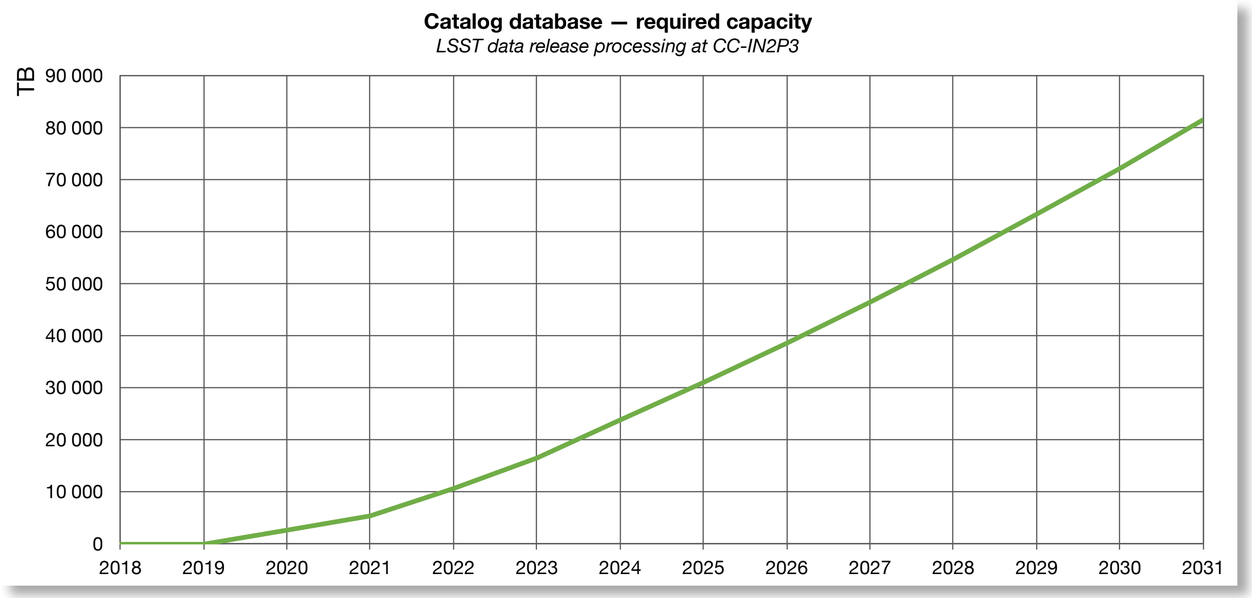
Disk servers: peak 65 (7 racks)

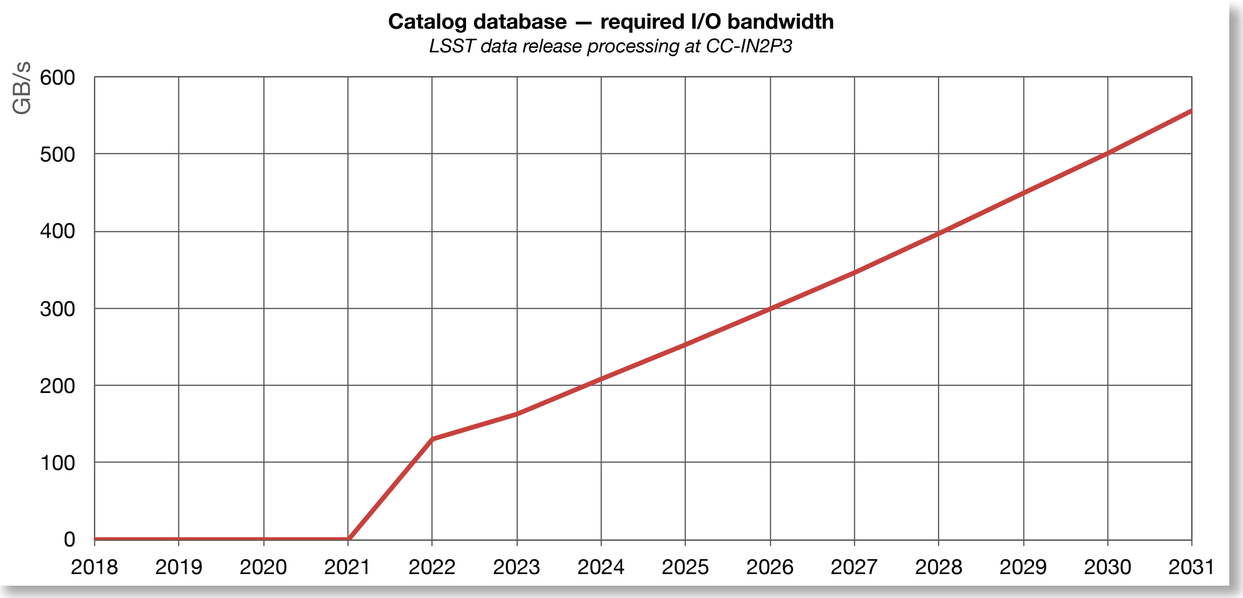
CC-IN2P3 disk storage installation*:

Allocation:	26 PB
Number of disk servers:	257 (est.)
Delivered bandwidth:	???

** as of Mar. 11th, 2018*

PLANNING: CATALOG DATABASE

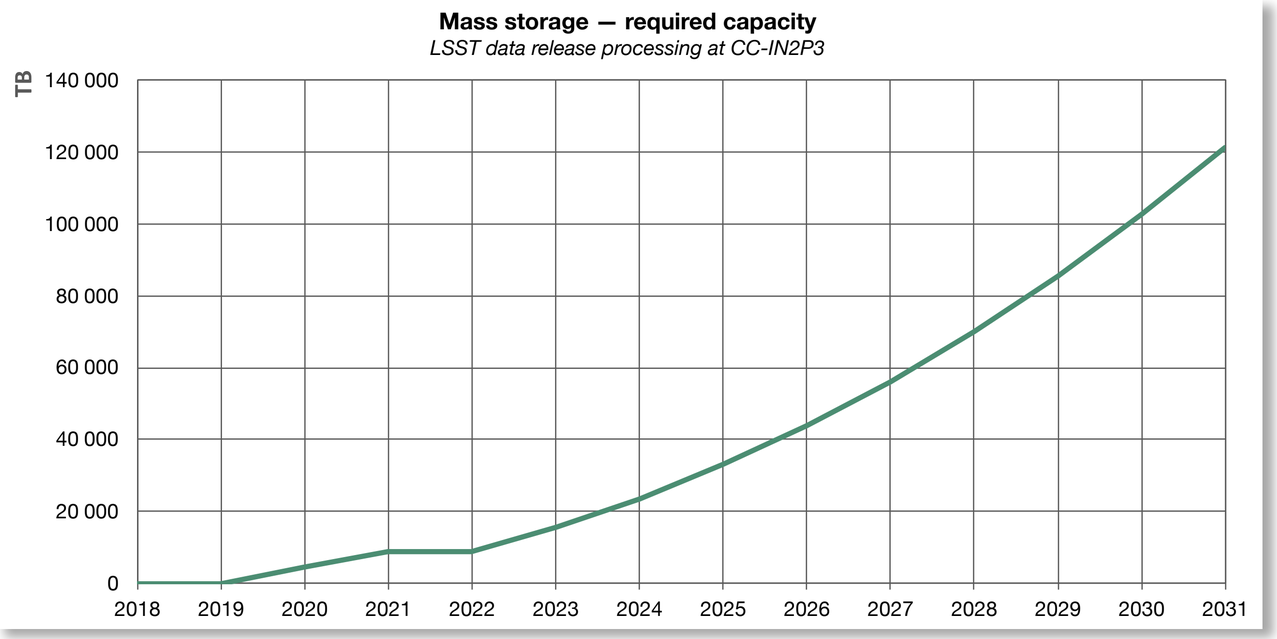


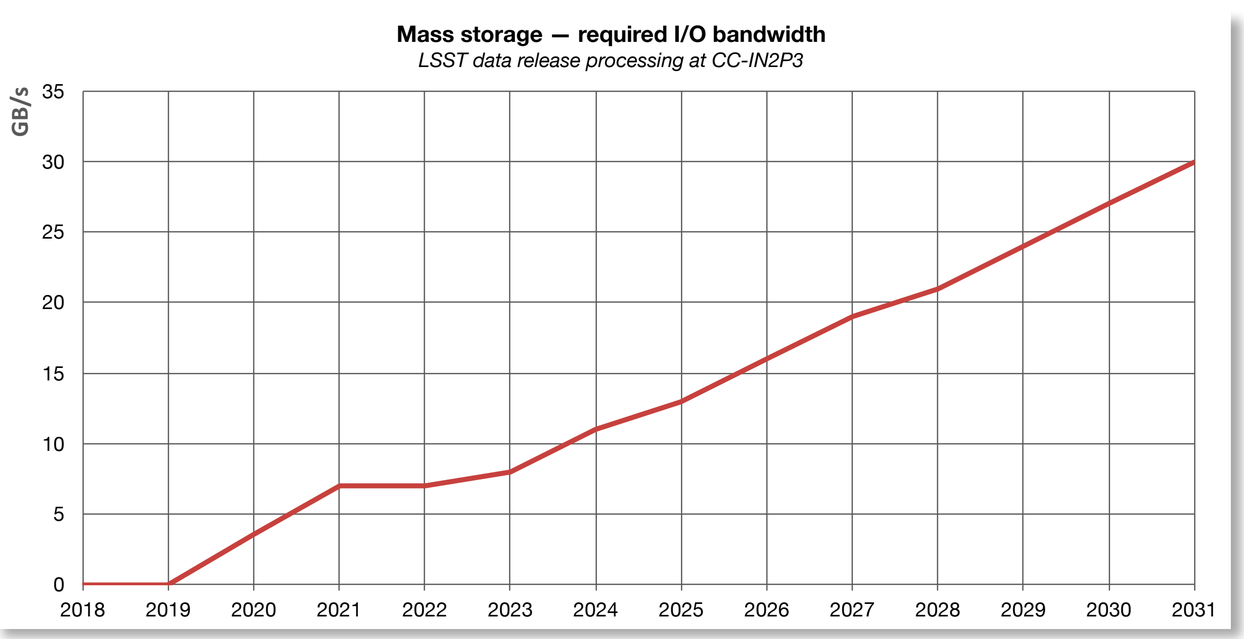


Catalog DB servers: peak 197 (10 racks)

The LSST catalog database will be a new component of the CC-IN2P3 infrastructure we don't have experience operating

PLANNING: MASS STORAGE





Tape drives: peak 66
Tape movers: peak 22
Tape cartridges: peak 3900

CC-IN2P3 MSS installation*:

Used storage:	53 PB
Number of disk movers:	13 (1.7 PB)
Number of tape drives:	50
Number of tape movers:	???
Delivered bandwidth:	???

** as of Feb, 2018*

Model assumptions

- Model built on several assumptions which may or may not hold for the long time span of this project

  inputs and assumptions to be revisited annually

- Operation policies

  different lifetimes for different types of equipment: typically 4 or 5 years

- Purchase schedule

  funding and purchase based on annual cycles

  equipment required for year N to be purchased on first half of same year, to be ready for production on Oct. 1st year N

- Machine room infrastructure

  budget for machine room infrastructure (i.e. UPS, water chiller, water and power distribution, etc.) not included in this model

  model does include cost of racks

Model assumptions (cont.)

- **CPU efficiency**

  we assume LSST software to be able to exploit 12% of the raw CPU capacity

  LSST initial estimation is 25%

  LHC observation is < 10%

- **Electric consumption**

  envelope of power consumption per compute and storage node assumed constant over the period

- **Tape storage**

  used LTO technology as the baseline

  model assumes we need to purchase tape drives, cartridges and tape movers

  purchase of tape library is not included

- **Software licensing costs**

  software licensing costs (e.g. for HPSS, GPFS, GridEngine) not included

Model assumptions (cont.)

- **Contingency**

  model includes contingency according to the IN2P3 project management practice

  level of contingency varies over time and related to the perceived severity of risk

  typically from 15% to 30%

- **Commissioning**

  we expect CC-IN2P3 to contribute to the commissioning effort, even if not part of the formal agreement

  considered required to be prepared for operations

FORESEEN IMPACT

FORESEEN IMPACT

- Disk storage

  connectivity per server to be upgraded to 40 Gbps

  it may become possible to use a shared storage area not exposing POSIX API, provided there is enough local disk in the compute nodes to stage the data in and out

  assuming GPFS as the baseline, but uncertainty regarding the future of GPFS at CC-IN2P3 need to be cleared as soon as possible

  number of objects (files, directories, links) to be counted in billions

- Compute nodes

  required per node network connectivity to storage servers of 10 Gbps with minimum number of network hops

  increase local disk capacity (scratch) to reach ~15 TB

FORESEEN IMPACT (CONT.)

- Catalog database

  new category of service

  good connectivity required to both login and compute nodes (for queries) and to storage servers (for populating the catalog)

  nodes for catalog database planned to have less storage capacity and more RAM, relative to typical file servers

- Data analysis and visualisation

  Python notebook-based environment for analysis, visualisation, quality control

- Mass storage

  the decision on the future of mass storage at CC-IN2P3 will directly impact LSST

  small files will be pervasive: 20 MB to 100 MB likely to be typical

- Authentication and authorisation

  we need to comply with the Authentication & Authorisation mechanism being deployed by the project

FORESEEN IMPACT (CONT.)

- We intend to allocate a fraction of the budget for purchasing modern equipment for R&D purposes

  repurposing obsolete equipment is not always suitable and may lead to questionable conclusions

- Application servers

  logging, monitoring, web applications, etc.

- Inter-site data transfer

  more servers will likely be needed

TECHNICAL WORK

CONTEXT

- The activities reported herein all aim…
  - *to **understand the mechanisms** of LSST data processing*
  - *to **get experience** with the software being developed by LSST for data reduction*
  - *to **understand the needs** of the computing infrastructure to perform bulk data processing efficiently and effectively*
  - *to **provide feedback** to developers*
  - *to **iterate***
- We need to acquire and demonstrate our…
  - *capacity to continuously **import and export** large amounts of data (mainly NCSA)*
  - *capacity to **perform bulk processing** locally at CC-IN2P3*
  - *capacity to **store large amounts of data** and deliver them to the applications*
  - *capacity to **archive and restore** large amounts of data*
  - *capacity to **serve large amounts of data** to **authorised members** of the projet*

ONGOING ACTIVITIES

- Special effort devoted to end user documentation
- Systematic deployment of LSST software

  both weekly and stable releases, local to CC-IN2P3 and in the cloud via CernVM FS

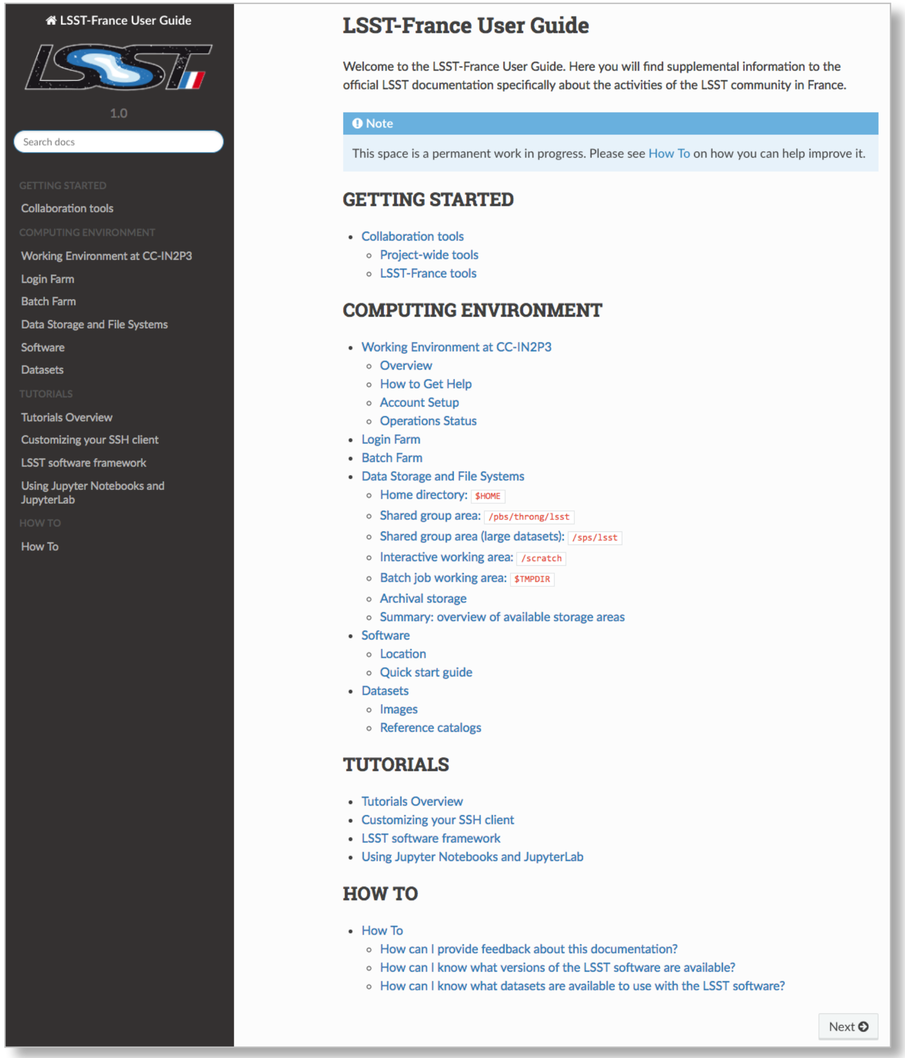
  allows for repeated processing of precursor or simulated data (e.g. CFHT, HSC, DESC) by IN2P3 scientists

  makes easier the detection of regressions

- Started work aiming at understanding I/O activity induced by LSST image processing software

  produced technical note DMTN-053

  another note in preparation



doc.lsst.eu

Ongoing activities (cont.)

- Hosting of one of the Qserv development platforms

  joint work by F. Jammes (LPC Clermont) and F. Wernli

  two logical clusters used by the developers of Qserv for code development, packaging, continuous benchmarking (see Fall2016 report)

  currently being reconfigured for use by IN2P3 scientists

- Continuous interaction with scientists and with the project as a whole is extremely important

  agile instant communication via Slack proven extremely effective

  dedicated channel for IN2P3-specific conversations

  forum for larger conversations via Discourse: `community.lsst.org`

ONGOING ACTIVITIES (CONT.)

- Exploratory work for prototyping the LSST Science Platform

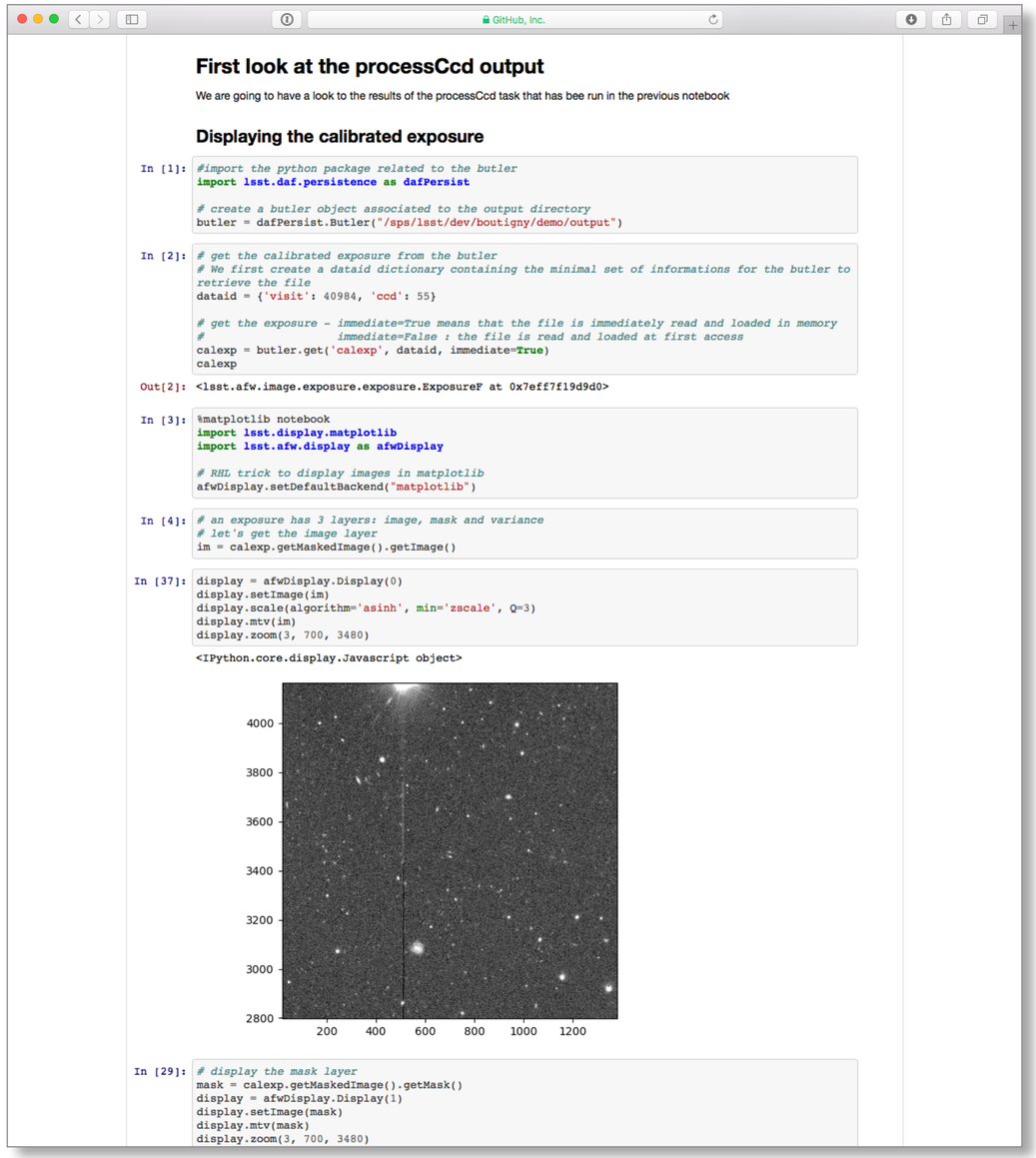
  based on Python Notebook as the main interface for data exploration and analysis

  prototype already used for several training sessions

  requires further integration with Firefly, the LSST interactive data exploration web application

- Experimentation with big memory machines (Dell partnership)

  for processing of images taken with Subaru telescope (see presentation) and for object index of LSST catalog (see report)



First look at the processCcd output

We are going to have a look to the results of the processCcd task that has bee run in the previous notebook

Displaying the calibrated exposure

In [1]:
```
#import the python package related to the butler
import lsst.daf.persistence as dafPersist

# create a butler object associated to the output directory
butler = dafPersist.Butler("/sps/lsst/dev/boutigny/demo/output")
```

In [2]:
```
# get the calibrated exposure from the butler
# We first create a dataid dictionary containing the minimal set of informations for the butler to retrieve the file
dataid = {'visit': 40984, 'ccd': 55}

# get the exposure - immediate=True means that the file is immediately read and loaded in memory
#                    immediate=False : the file is read and loaded at first access
calexp = butler.get('calexp', dataid, immediate=True)
calexp
```

Out[2]:
```
<lsst.afw.image.exposure.exposure.ExposureF at 0x7eff7f19d9d0>
```

In [3]:
```
%matplotlib notebook
import lsst.display.matplotlib
import lsst.afw.display as afwDisplay

# RHL trick to display images in matplotlib
afwDisplay.setDefaultBackend("matplotlib")
```

In [4]:
```
# an exposure has 3 layers: image, mask and variance
# let's get the image layer
im = calexp.getMaskedImage().getImage()
```

In [37]:
```
display = afwDisplay.Display(0)
display.setImage(im)
display.scale(algorithm='asinh', min='zscale', Q=3)
display.mtv(im)
display.zoom(3, 700, 3480)
```

```
<IPython.core.display.Javascript object>
```


In [29]:
```
# display the mask layer
mask = calexp.getMaskedImage().getMask()
display = afwDisplay.Display(1)
display.setImage(mask)
display.mtv(mask)
display.zoom(3, 700, 3480)
```

ONGOING ACTIVITIES (CONT.)

- Past preliminary work on exploring containers as a packaging mechanism for LSST processing pipelines

  reports on usage of Docker (see report) and Singularity (see report)

  work to be resumed in the framework of project HPCEuropa3

- Setting the foundations of the future of the LSST data transfer platform

  used for transferring data in real-life conditions from NCSA and NERSC (see presentation)

- Actively contributing to DESC data challenge 2

  involves: data transfer, bulk data processing and storage

  excellent opportunity for us to exercise all the required components, uncover issues, implement turnarounds / solutions and provide feedback

HIGHLIGHTS FOR 2017

- International scientific event: lyon2017.lsst.eu
  - *school, workshop and hackathon*
  - *24 speakers, 60 participants, 13 countries*
  - *all presentations video-recorded and available online*
- Training programme on the usage of Python and the LSST software framework
  - *targeted at IN2P3 scientists*
  - *June and October 2017, to be reiterated as need arise*
- Participation to the joint NSF / DOE review of LSST Data Management subsystem
  - *invited to present status of IN2P3 contribution to the data management subsystem*
  - *very helpful for IN2P3 visibility within the project and for keeping up-to-date with the project's plans*
  - *a lot of information and intentions are in people's heads and are not written in project documents*

SUMMARY

SUMMARY

- LSST is a world-class, high-profile project in optical astronomy

  high expectations from the scientific community and from the funding agencies about what LSST will bring over the next decade

- Focus of our work so far have been understanding what is needed in our infrastructure for LSST

  we have made significant progress and now have a quantified roadmap

- The challenge is serious

  ***concrete organisational commitment** is urgently needed and of prime importance to succeed*

  contribution by CC-IN2P3 likely to be very visible, for better or for worse

Questions & Comments