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

Source: LSST

LSST activities at CC-IN2P3

overview

fabio hernandez

fabio@in2p3.fr

CCIN2P3

LSST-France, Paris, December 8th, 2015

cnrs

lsst

Contents

- LSST data management overview
- LSST at CC-IN2P3
- Ongoing work
- Planned activities
- Conclusions

LSST data management

LSST data management

- 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 release a self-consistent, immutable dataset, composed of processed data since the beginning of the survey (“Data Release Processing”)

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

- Publication

  to deliver the reduced data (images + catalogs)

  to facilitate custom data reduction and individual data analysis

Data products

Level 1

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

Alert Processing

Level 2

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

Deep co-added **images**

Data Release Processing

Level 3

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

Software and APIs enabling development of analysis code

Source: M. Juric

Partners

Stanford Linear Accelerator Center

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

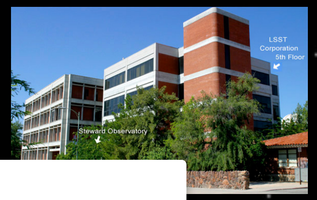
Princeton University

Infrared Processing and Analysis Center
California Institute of Technology

National Optical Astronomy Observatory

IN2P3 / CNRS computing center

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

Data volume

- Raw data

  2000 images science + 450 calibration images per observing night

  300 nights per year

  6.4 GB per image, 15 TB per night, 4.5 PB per year

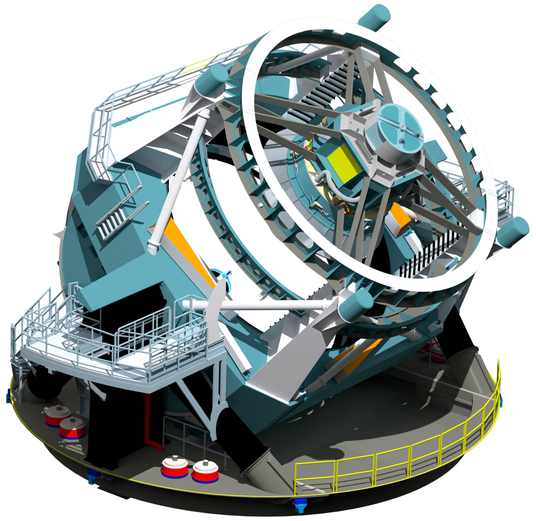
- Aggregated data over 10 years (2022 - 2032)

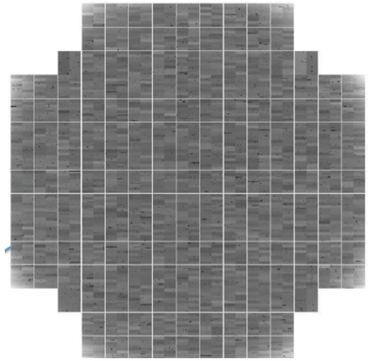
  images: 515 PB, ~6M exposures

  catalog: 83 PB

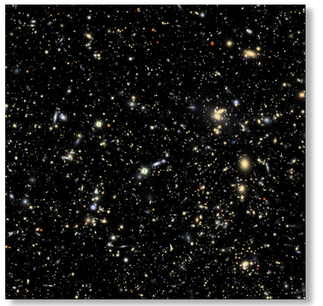
  only a fraction of the generated images are permanently stored

Catalog





raw data



processed image

catalog
(objects, sources, transients, exposures, etc.)

Qserv catalog: custom, distributed relational database

spatial partitioning by sky coordinates, with overlaps

very high number of rows: 32 trillion

LSST at CC-IN2P3

LSST at CC-IN2P3

- Formal agreement signed between LSST Corp., NCSA and IN2P3 in March 2015
- Data release processing (i.e. level 2) to be jointly performed by both NCSA and CC-IN2P3

  NCSA has lead responsibility for LSST data release processing

  CC-IN2P3 to process 50% of the data and store the full dataset, both raw and derived data

  both NCSA and CC-IN2P3 will validate the data produced by the other party and will each host an entire copy of every annual data release

- Note that end-user analysis is not covered by this agreement

LSST at CC-IN2P3 (cont.)

- Task force launched in May 2015
- Mandate

  coordinate the activities to perform the agreed LSST data processing at CC-IN2P3

  those activities require the intervention of experts from all teams

  task force leader liaises with LSST project and LSST-France and reports to CC-IN2P3 steering committee

Ongoing activities

Ongoing activities

- Ongoing work on several fronts

  in close interaction with LSST data management, LSST-France computing and NCSA

- *Joint Coordination Committee* established to coordinate activities among data centers

  composed of LSST data management architect, NCSA LSST project leader, LSST-France computing coordinator, CCIN2P3 LSST project leader

  remotely meets monthly

  in addition to face-to-face workshops between NCSA and CCIN2P3 experts, twice a year

Ongoing activities (cont.)

- Operations of a development and integration cluster for celestial objects catalog (Qserv)

  50 nodes, 400 CPU cores, 800 GB of memory, 500 TB of disk + virtual machines for building, packaging, deploying, HTTP proxying, etc.

  hardware donated by DELL in the framework of an institutional partnership

  used for large scale tests performed in August 2015

  Fabrice Jammes (LPC Clermont) is member of the development team

- Review of infrastructure sizing and equipment costs

  following an update of the experiment sizing assumptions, requirements and technology evolution made in 2013

Ongoing activities (cont.)

- Studies for understanding I/O activity induced by LSST software

  developed synthesised file system for tracing I/O activity and associated tools to summarise the collected data

  used data collected to develop a trace replay simulator on top of SimGrid

  used replay tool to simulate execution of LSST demo application

  currently working on simulating I/O of the simultaneous astrometry application, integrated to the LSST software stack by P. Astier and D. Boutigny

- Studies for understanding impact of using FITS formatted data served by a networked file system

  investigated suitable configurations of GPFS for serving FITS files to avoid cache trashing

  goal is to provide input to LSST data management on the impact of file format

  thanks to P. Astier for his numerous inputs on this issue

Ongoing activities (cont.)

- Exploration of MapReduce model for LSST data processing

  developed prototype application on top of Apache Spark to scan 9M FITS files containing data from the Canada-France-Hawaii Telescope (CFHT)

  extracted FITS header metadata and populated a relational database (400GB)

  interesting for exploring the feasibility of using databases for serving metadata currently stored in FITS files, instead of stressing the file system

  what we found:

header size*	36 KB
number of key-value pairs in FITS header*	438

* 98th percentile

Ongoing activities (cont.)

- Operations of the experimental CernVM FS-based infrastructure for distribution of binary LSST software releases

software specifically built from sources for distribution via this channel

targeted to both individual users and compute nodes

currently stable releases only, weekly releases may be added in the future

individual users don't need to build from source but use the software online

details: https://github.com/airnandez/lsst-cvmfs

other binary distributions: http://sqr-002.lsst.io

Planned activities

Planned activities

- [science] To exercise LSST workflows for processing CFHT data

  data already present at CC-IN2P3

  foreseen scale realistically enough to make inferences about what the LSST workflows may look like

  CC-IN2P3 is very interested in this exercise, which is an opportunity to detect issues early on and iterate

- [science] To participate in the precursor data processing being planned by Dark Energy Science Collaboration (DESC)

  baseline: 3 data challenges of increasing complexity and size from 2016 to 2019, comparable to the first year of LSST data processing

  operational data exchange with US sites is required for these challenges

Planned activities (cont.)

- To prototype inter-site data exchange platform: NCSA ⇄ CC-IN2P3

  use object stores as emission/reception buffers and HTTP2 (or gridFTP) as transport protocol

  start with OpenStack Swift and gain experience, evaluate commercial solutions afterwards

  investigate object store built-in vs. externally-driven georeplication

  this requires good network connectivity with NCSA: we need first to determine what bandwidth we can actually use

- To explore suitable mechanisms for orchestrating LSST workflows

  first within a single data center, then over several geographic locations

  exploit workflow-specific applications packaged as Docker containers, evaluate candidate orchestration mechanisms

Planned activities (cont.)

- To explore what it takes to use cloud-based platforms (on premises and public) for executing LSST workflows
- Emerging interest group on LSST data processing including NCSA, ANL, NERSC and CC-IN2P3

Conclusions

Conclusions

- It is CC-IN2P3's intention to **contribute building** the computing infrastructure for the experiment and we are working towards that goal

  as opposed to only position ourselves as a data processing service provider

  close collaboration with LSST-France community is required for us to succeed

- Work performed so far has given us some visibility

  CC-IN2P3 is perceived as a contributor by both LSST and by LSST-France

- Technical work has been instrumental to get familiar with the context, the constraints, the internals, the technical challenges and the opportunities

  still a lot to be done and new opportunities emerge very often

- New opportunities of collaboration with big players of the US high performance computing world may emerge

  this also supposes some challenges

Questions & Comments