

Overview of LSST-related activities at CC-IN2P3

Presented by Christian Arnault

Fabio Hernandez
on behalf of the CC-IN2P3 team

LSST-France, Montpellier, April 8th, 2015

Contents

- MoA data processing
- Qserv integration cluster
- Binary distribution of LSST stack
- Analysis of file I/O activity
- Issues

Memorandum of Agreement

- MoA data processing signed on March 5th, 2015 by Steven Kahn (LSST corp.) and Jacques Martino (IN2P3)
- Data processing to be jointly performed by both CC-IN2P3 and NCSA
 - NCSA has lead responsibility of LSST data processing*
 - CC-IN2P3 to process 50% of the data and store the full data set*
 - NCSA and CC-IN2P3 will jointly validate the data produced by the other party and will each end up with a full data release*
- NCSA team to visit CC-IN2P3 this Spring
 - details being finalised*
- CC-IN2P3 intends to contribute also to the outreach effort

Qserv integration cluster

- Goals

to provide a realistic testbed platform for developing, integrating, operating and validating Qserv

to exercise the system with realistic use cases (e.g. CFHT data processing using LSST software stack) and provide feedback to developers

- Aggregated capacity: 800 GB memory, 500 TB disk, 400 CPU cores

eventually composed of 50 boxes + virtual machines for building and packaging the software

25 boxes already operational, expected delivery of remaining hardware late April or May — slightly updated hardware configuration

- Qserv development team routinely working with it

Binary distribution of LSST stack

- Goal

to provide a ready-to-use distribution of stable versions of the LSST stack

- Benefits

no need to build the software from scratch in order to use it

self-contained, unique binary distribution to be used for personal computers, worker nodes, virtual machines, containers, etc., for consistency

usable also for developing or extending modules on top of the built-in stack components

- CernVM FS-based LSST repository hosted and maintained at CC-IN2P3

LSST v9.2 available, release candidate of v10 in preparation

works on Ubuntu 14.04, Scientific Linux 6 & 7 and CentOS 7 — MacOS X to be tested

requires installation of CernVM FS client software on the execution machine (one time process)

- More information: <https://github.com/airnandez/lsst-cvmfs>

Analysis of file I/O activity

- Goal: to understand file I/O patterns induced by LSST stack
to identify the characteristics of the data storage platform which best suits the needs of LSST

or alternatively, to modify LSST software to take into account the intrinsic limitations of the available platforms
- Status
developed tool for collecting machine-parseable data on file I/O access

Linux and MacOS X
- Next step
finalise the tools for analysing the data collected when using the LSST stack for processing CFHT data
- More information: <https://github.com/airnandez/cluefs>

Issues

- We still don't have an appropriate means for remote collaboration via video-conference

RENATER services do not seem suitable for remote participation from personal computers at the scale needed by LSST-FR

RENAvisio: no convenient Linux client, no telephone bridge for calling from abroad without fees

Rendez-vous: browser-based, currently technically limited to 12 simultaneous participants

- More information: https://services.renater.fr/voix_et_image/quel_equipement_utiliser