LARGE SYNOPTIC SURVEY TELESCOPE

Large Synoptic Survey Telescope (LSST)

Chiliaen Data Access Center

William O'Mullane, Donald Petravick

LDM-564

Latest Revision: 2017-08-09

Abstract

This document describes the Chilean DAC

Change Record

Version	Date	Description	Owner name
1.0	2017-08-20	Initial version.	W. O'Mullane, D. Petravick

Document source location: `https://github.com/lsst/LDM-564`

1 Introduction

The Memo of understanding with Chile roughly defines a Data Access Center at the base facility in Chile. In this document we would like to begin to specify more precisely with this data access center would look like and the full range of services it will provide. This should be seen as an initial discussion document to allow the Chilean colleagues to collaborate with us to define the Chilean DAC.

2 The Chilean DAC

The Chilean DAC should be very similar to the US DAC at NCSA. We will deploy the science platform in Chile as the interface to the system. The science platform is now well documented with the vision given in **?** more formal requirements in **?** and the design in **?** .

The DAC will require other software, some hardware and of course data to function as a DAC. We go through these in the following sections.

2.1 Software and Services

The components of the science platform: Portal, Access Services(DAX) and Notebooks (JupyterLab) will be deployed in Chile in the same manner as NCSA. Currently using Kubernetes to deploy Docker Containers. The platform gives access, through the notebook, to several versions of the LSST Software Stack allowing processing of image data etc. This interface will also allow users to spawn batch jobs,to process large amounts of data. such batch jobs will have to be written using the Data Management batch system used by LSST large processing. This is accessed through super task (see **?**).

The science platform allows users to upload files and images as well as store temporary results. It most be noted that users of the Chilean DAC will have access to the US DAC however the *user data* is bound to the site where it was created.

All the regular security services etc will also be deployed in Chile as in NCSA. The entire system will be administered from NCSA in Illinois thus if new versions of software are deployed in Illinois they will also be deployed in Chile.

In addition we will need to work together to get the authentication system of the Clean Grid integrated withe the DAC - thus we should be able to Authenticate users using Grid credentials furthermore we should be able to access grid resources from the DAC using the same credentials.

The MOU also mentions last mile networking to Chilean astronomers. We believe this is covered with the current networks to Santiago but we should discuss and agree.

2.2 Hardware and infrastructure

The Chilean MOU provides for a DAC which is 10% of the size of the US DAC. The current DR2 US DAC is intended to comprise:

- Computing:2,400 cores (≈ 18 TFLOPs)
- File storage: ≈ 4PB (VOSpace)
- Database storage: ≈ 3PB (MYDB)

Hence The Chilean DAC would at a minimum have:

- Computing:2,40 cores (≈ 1.8 TFLOPs)
- File storage: ≈ 400TB (VOSpace)
- Database storage: ≈ 300TB (MYDB)

In addition there would be disk to hold the RAW data and Catalogs.

Access to he the DAC will of course require networks. Here we have done better than planned with 200 Gbps links

2.3 Data

The raw data will be stored on disk at the base facility.

TABLE 1: Primary data products and their retention methods. Virtual here means generated on request.

Data product	How retained
Raw Images	Files
Coadded Images	Files
Difference Images	Files
Objects	Relational Tables
Sources	Relational Tables
Forced Sources	Relational Tables
DIA objects	Relational Tables
DIA sources	Relational Tables
SS Objects	Relational Tables
Alerts	Relational Tables
External Catalogs	Relational Tables
Postage Stamp Images	Virtual
Processed Visit Images	Virtual
Observatory Metadata	Relational Tables
Observatory Metadata	Files

3 Conclusion

We will put a useful scientific service in the base facility in Chile. This should enable local science activities and also provide seamless access to further resources at NCSA.

4 References

5 Acronyms

Acronym	Description
DAC	Data Access Center
LSST	Large Synoptic Survey Telescope
NCSA	National Center for Supercomputing Applications

SS	Subsystem Scientist
TFLOP	Tera FLOP
US	United States