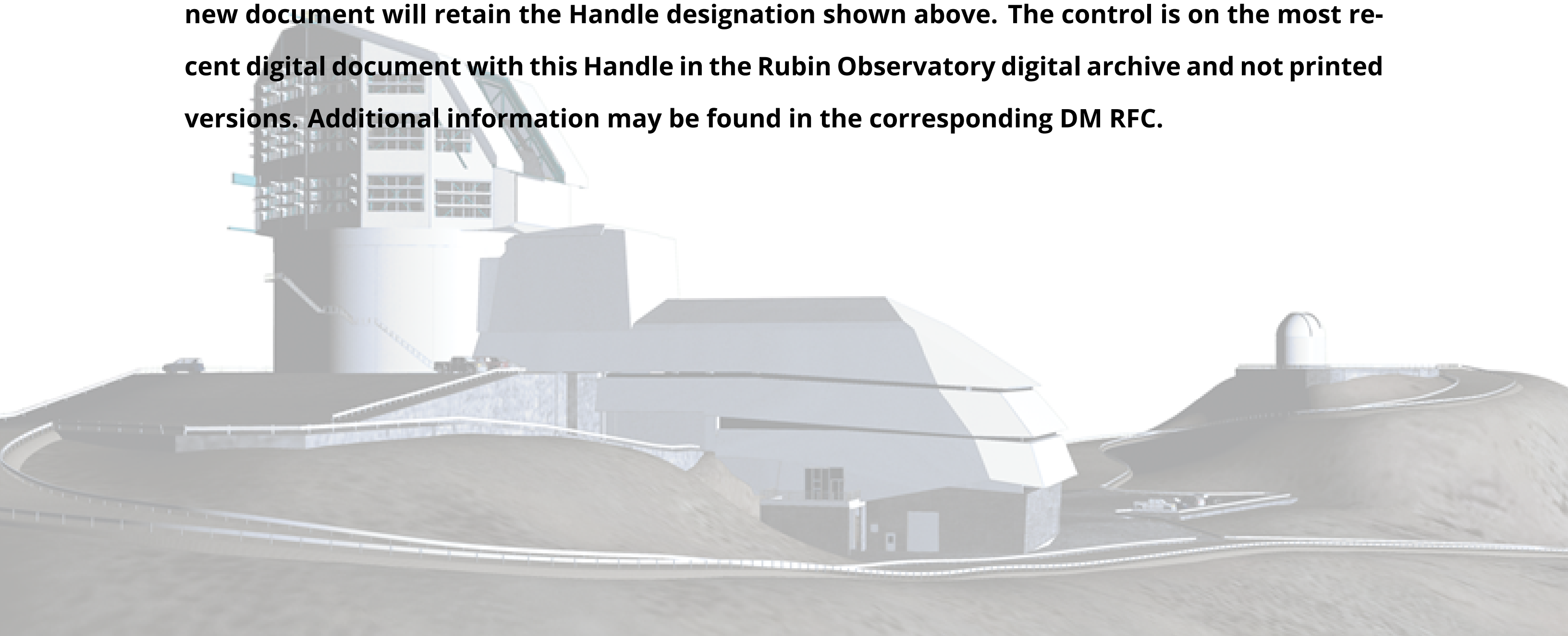
Vera C. Rubin Observatory
Data Management

Chilean Data Access Center

William O'Mullane

LDM-572

Latest Revision: 2021-04-29



Abstract

This document describes the Chilean Data Access Center.

Change Record

Version	Date	Description	Owner name
	2017-08-20	Initial version.	W. O'Mullane, D. Petravick
1.0	2017-12-12	Approved in RFC-419.	W. O'Mullane, T. Jenness
2.0	2021-04-29	Update inline with OPS plan. Approved in RFC-769	W. O'Mullane

Document source location: `https://github.com/lsst/LDM-572`

1 Introduction

The Memorandum of Understanding with Chile [Document-11622] roughly defines a Data Access Center (DAC) at the Base Facility in Chile. In this document we would like to begin to specify more precisely what 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 our 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 SLAC. We will therefore deploy the Rubin Science Platform (RSP) in Chile as the interface to the system, just as we do for the US DAC. The LSP is now well documented with the vision given in LSE-319, with more formal requirements in LDM-554 and the design in LDM-542.

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 ("Aspects") of the Science Platform: Portal, Data Access Services (DAX), and Notebooks (JupyterLab) will be deployed in Chile in the same manner as at NCSA. The baseline design is to use Kubernetes to deploy Docker containers, an approach we are currently using in our prototype systems. The Platform gives access, through the Notebook Aspect, to several versions of the LSST Software Stack allowing processing of image data and catalogs. 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 installed batch system.

The Science Platform provides a "user workspace" which allows users to upload images and other files, and to create databases, as well as store temporary results. While users of the Chilean DAC will have access to the US DAC, *user workspace data* is bound to the site where it was created.[1]

[1] The user workspace at the two DACs will be accessible via the Internet via VOSpace and the DAX Web services, but with reduced performance compared to their local use within each DAC.

All the regular security and other services 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 Chilean Grid integrated with the LSST system [LSE-279] – thus we should be able to authenticate users using Grid credentials, and furthermore we should be able to access Grid resources from the DAC using the same credentials.

2.2 Hardware and infrastructure

The Chilean MOU provides for a DAC which is 10% of the size of the US DAC. The Data Release 2[2] user computing resources at the US DAC are currently intended to comprise [DMTN-135]:

- Computing: 2,400 cores (≈ 18 TFLOPs)
- File storage: ≈ 4 PB (user file workspace: VOSpace or equivalent)
- Database storage: ≈ 3 PB (user database workspace)

Hence the Chilean DAC would at a minimum have:

- Computing: 240 cores (≈ 1.8 TFLOPs)
- File storage: ≈ 400 TB (user file workspace: VOSpace or equivalent)
- Database storage: ≈ 300 TB (user database workspace)

This raises an immediate problem - 10% of the data base is not very useful unless we bring only the Object catalog. We shall discuss further in Section 3.

In addition see Section 2.3. There will be additional computing capacity as required to support the operation of the Data Access services themselves as well as the Portal.

[2] Data Release 2 is used as a reference because the compute and storage resources required to support it are covered by the MREFC construction budget.

2.2.1 Networks

Access to the DAC will of course require networks and we have done better than planned with a 100 Gbps primary and 40 Gbps secondary links. A priority scheme is being put in place to ensure availability of this network for Chilean academic use, through our $7M USD investment in REUNA. The general notion is to provide 10 Gbps to the community in Chile from the backup link when it is available (i.e. primary is functioning). Should Rubin need th backup link the 10 Gbps would be as available with our traffic getting priority.

REUNA (the Chilean Research and Educational Network, similar to Internet2 in the US) is our primary network operator in Chile and connects networks on the AURA observatory site in the Elqui Valley with the AURA Campus in La Serena, and on to Santiago where there is a connection point with the LSST international networks to the United States [LSE-78]. REUNA will provide connections in Santiago for non-LSST segments and LSST segments to La Serena. Thus, priority access is available to the LSST Chilean Data Access center in La Serena from any institution that has access to REUNA networks.

REUNA also connects the LSST segments to the National Laboratory for High Performance Computing Center (NLHPC) at the University of Chile in Santiago. Both REUNA and AURA are formal members of the NLHPC. NLHPC has constructed a high-bandwidth ring in Santiago, managed by REUNA. The LSST segments will be connected with the Santiago NLHPC ring, enabling high-throughput data transfers as well as high-quality connectivity. AURA and REUNA will work with NLHPC to ensure a shared authorization and authentication service to provide for integrated access and services. These connectivity elements will enable grid-based access from any REUNA-connected institution to the NLHPC.

Leveraging AURA's membership in Internet2, and REUNA as an Internet2 peer in Chile, LSST will transit Internet2 traffic from Chilean astronomy institutions connected to REUNA to an Internet2 peering point in North America. Chilean astronomy Internet2 traffic on the REUNA networks will be carried on the LSST international segments to Florida. From there, the traffic will be directed to a US-based Internet2 peering point. The availability and bandwidth on the LSST international segments will be determined by treating the Chilean Internet2 traffic as one of several priority groups. When other higher priority traffic is present, the Chilean astronomy Internet2 traffic will remain in a queue. When there is no higher priority traffic, the Chilean astronomy Internet2 traffic will be transited. In this way, the Chilean Internet2 traffic will be transferred on an "as available" basis.

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

Data product	Mode of retention
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	Key-Value Store (TBC)
External Catalogs	Relational Tables
Postage Stamp Images	Virtual
Processed Visit Images	Virtual
Observatory Metadata	Relational Tables
Observatory Metadata	Files

We shall measure the network performance and under nominal conditions we will ensure that Chilean academic priorities are met on the network. REUNA will manage the north/south link hence we should be able to achieve this easily. We must ensure with REUNA that traffic coming into USA also meets the NSF acceptable use policy for the NSF funded US network – there is no problem with academic traffic.

2.3 Data

At least the LSST raw data and possibly data products from Data Releases, will be stored on disk at the Base Facility. The types of data and modes of access are indicated in Table 1. The table includes "virtual data products" which are recreated on demand rather than retained in persisted form. The storage and other resources required to support the released data products are in addition to the user workspace resources described above in Section 2.2.

3 Proposal for the Chilean DAC

In the operations proposal [LDO-31] we included funding for the Chilean DAC of the order $13M over the operations.

This is based on providing storage and compute above the level discussed in Section 2.2, sufficient to store raws and CoAdds. The total storage and compute is shown in Table 2. The cost is based on Rome processors.

Table 2: This table summarizes the compute and storage estimates usde to cost the Chilean DAC.

Year	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	
Chilean DAC Fast Storage (TB)	335	570	825	1064	1274	1484	1694	1904	2114	2324	
Chilean DAC Latent Storage (TB)	16399	29570	41786	46842	51899	56956	62013	67070	72127	77183	
Cores (Chilean DAC)	103	187	280	373	466	560	653	746	840	933	

This does not include storage and compute for Qserv, queries to Qserv would be serviced via the USDF. There is enough storage for raw images, LSSTCam CoAdds and some other images or catalogs but not everything. The baseline, which may be discussed, is to store the raw images in the DAC. The best use of the storage should be further discussed - i.e. some sort of caching of image files may be more useful than trying to have all of any specific data set. The object store backed butler would in effect make access to images in USDF fairly seamless. The sizes of the image data sets are given in DMTN-135 but a few key numbers are reproduced here in Table 3.

Table 3: This table summarizes some key data voluems used in DMTN-135. All Sizes TB.

Year	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	
LSSTCam Raw Images	6982	11798	16614	21430	26246	31062	35878	40694	45510	50326	
LSSTCam Output Images	3933	10676	16857	23599	30342	37084	43827	50570	57312	64055	
LSSTCam Output Coadd Images	8636	16364	23182	23182	23182	23182	23182	23182	23182	23182	
LSSTCam Output Parquet	9302	25248	47839	71758	95678	119597	143516	167436	191355	215275	
DIff DAC Latent to Raw+Coadd	781	1408	1990	2230	2471	2712	2953	3194	3435	3675	

There was also some discussion of a Chilean Data Cloud project. We should see if/how that fits in here.

4 Conclusion

We will put a useful scientific archive and data analysis service in the Base Facility in Chile. This should enable local science activities and also provide seamless access to further resources at the USDF. Discussion with the local community should be initiated near the start of operations to decide on the best way to set this up.

5 References

[1] **[LDO-31]**, Blum, R., et al., 2020, *LSST Operations Proposal* , LDO-31, URL https://ls.st/LDO-31

[2] **[DMTN-135]**, Butler, M., Lim, K.T., O'Mullane, W., 2019, *DM sizing model and purchase plan for the remainder of construction.*, DMTN-135, URL http://DMTN-135.lsst.io

[3] **[LDM-542]**, Dubois-Felsmann, G., Lim, K.T., Wu, X., et al., 2017, *LSST Science Platform Design*, LDM-542, URL https://ls.st/LDM-542

[4] **[LDM-554]**, Dubois-Felsmann, G., Ciardi, D., Mueller, F., Economou, F., 2018, *Science Platform Requirements*, LDM-554, URL https://ls.st/LDM-554

[5] **[LSE-319]**, Jurić, M., Ciardi, D., Dubois-Felsmann, G., 2017, *LSST Science Platform Vision Document*, LSE-319, URL https://ls.st/LSE-319

[6] **[LSE-78]**, Lambert, R., Kantor, J., Huffer, M., et al., 2017, *LSST Observatory Network Design*, LSE-78, URL https://ls.st/LSE-78

[7] **[Document-11622]**, Smith, W., Vera, V.P., 2011, *Supplementary and Clarifying Agreement between the Universidad de Chile and AURA covering the use of the LSST on Cerro Pachon*, Document-11622, URL https://ls.st/Document-11622

[8] **[LSE-279]**, Withers, A., 2017, *Concept of Operations for Unified LSST Authentication and Authorization Services*, LSE-279, URL https://ls.st/LSE-279

6 Acronyms

Acronym	Description
AURA	Association of Universities for Research in Astronomy
DAC	Data Access Center

DAX	Data Access Services
DIA	Difference Image Analysis
DM	Data Management
LDM	LSST Data Management (Document Handle)
LSE	LSST Systems Engineering (Document Handle)
LSP	LSST Science Platform (now Rubin Science Platform)
LSST	Legacy Survey of Space and Time (formerly Large Synoptic Survey Telescope)
MOU	Memo Of Understanding
MREFC	Major Research Equipment and Facility Construction
NCSA	National Center for Supercomputing Applications
NLHPC	National Laboratory for High Performance Computing Center
NSF	National Science Foundation
OPS	Operations
PB	PetaByte
REUNA	Red Universitaria Nacional
RFC	Request For Comment
SS	Subsystem Scientist
TB	TeraByte
TBC	To Be Confirmed
US	United States
USD	United States dollar