LARGE SYNOPTIC SURVEY TELESCOPE

Large Synoptic Survey Telescope (LSST)
LSST Document

Proposed Policy for Independent Data Access Centers

William O'Mullane, Beth Willman, Melissa Graham, Leanne Guy

LPM-251

Latest Revision: 2019-02-25

Abstract

This document describes the proposed policies for groups that are independent from the LSST Project and Operating Facility and would like to stand up an independent Data Access Center (iDAC). Some iDAC may want to serve only a subset of the LSST data products: this document proposes three portion sizes, from full releases to a ”light” catalog without posteriors. Guidelines and requirements for iDACs in terms of data storage, computational resources, dedicated personnel, and user authentication are described, as well as a preliminary assessment of the cost impacts. **The purpose of this draft document is to serve as a *preliminary* resource for partner institutions wishing to assess the feasibility of hosting an iDAC, and is subject to change.**

Change Record

Version	Date	Description	Owner name
	2018-03-24	Initial version.	WOM
	2019-02-25	Added site topology schematic	LPG

Document source location: `https://github.com/lsst/LPM-251`

1 Introduction

All access to LSST proprietary data products will be through a Data Access Center (DAC). The United States's DAC will be hosted at the National Center for Supercomputing Applications (NCSA), where authorized LSST users with data rights and data access privileges will perform scientific queries and analysis on the full data releases using the LSST Science Platform (LSP). 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 Chilean DAC will be equivalent in functionality to the US DAC, but scaled-down in terms of the computational resources available for query and analysis given the smaller Chilean community [LDM-572].

All access to, and use of, the LSST proprietary data and data products is subject to the policies described in LPM-261.

This document proposes a set of guidelines and policies for partner institutions – in the US, Chile, or one of the International Contributors with signed Memoranda of Agreement – that are interested in hosting the LSST data, in whole or in part, for their affiliated members as an independent Data Access Center (iDAC). The following sections include the types of data products that could be hosted (Section 2), the requirements and responsibilities that would be expected of an iDAC hosting LSST proprietary data products (Section 3), and a description of the main costs *vs.* their science impacts (Section 4).

The contents of this draft are meant to provide a preliminary resource for partner institutions who may be assessing the feasibility of hosting an iDAC. The specific mechanisms and processes by which future iDACs will negotiate the bulk transfer of data, the installation of software, etc. is considered beyond the scope of this document.

To help get an idea of sizes Table 1 gives an overview.

2 Types of Data Products for iDACs

In quite good shape, Leanne to an be fleshed out with demonstrative science use cases and experience from other archives

Table 1: Size summary based on LDM-141

Table	Bytes/row	Rows (DR1 -> DR11)	DR1 (TB)	$\times$ Growth	DR10 (PB)
Object_Lite i	1840	$2.26^{10} -> 4.74^{10}$	42	2.1	0.08
Object_Extra	20393	$2.26^{10} -> 4.74^{10}$	461	2.1	0.9
Source	453	$4.51^{11} -> 9.01^{12}$	204	20.0	4.0
ForcedSrc	41	$1.20^{12} -> 5.01^{13}$	49	42	2.0
DiaObject	1405	$7.94^{08} -> 1.54^{10}$	1.1	19.4	0.002
DiaSource	417	$2.26^{09} -> 4.52^{10}$	0.9	20	0.002
DiaForcedSource	49	$1.50^{10} -> 3.01^{11}$	0.7	20	0.001
Year 1 raw images:$3PB$, tables:$\sim 1PB$, half for Object_Extra,$0.2PB$ Sources					
Year 10 raw images:$30PB$, tables:$\sim 7PB$,$4PB$ Sources,$2.0PB$ Forced ,$1PB$ Object_Extra					

The full collection of LSST data products, from images to catalogs, is described by the Data Products Definitions Document [LSE-163]. Below, three potential levels of data products that iDACs might consider hosting are described: the full data release with images, the data release catalogs, and a low-volume (“`lite`”) subset of the catalogs.

2.1 Full Release(s)

In this case the iDAC would be hosting all of the raw and processed images and catalogs, as described in [LSE-163]. Including the raw image data in an iDAC requires roughly 6 petabytes per year of survey, so this is a serious augmentation on resources in terms of both hardware and personnel. The processed data and associated calibrations bring the total data volume to 0.5 exabytes for a single data release. Some data volume could be saved by taking only a single calibrated image per band, but the total would still be 60 petabytes (with compression it may be possible to reduce this even further). Any iDAC considering hosting the full data release should also deploy the full LSST Science Platform LSE-319 in order to maximize science productivity and their return on investment in hosting an iDAC.

2.2 Catalog Server

Alternatively, an iDAC may find that hosting only the data release catalogs, and not the images, is sufficient for the scientific needs of its community. This will probably require the specific LSST database server [LDM-135] and specific machines, and the deployment of the database system and the associated subset of data access services (DAX; e.g., web APIs, Qserv, LDM-152). The full `Object` catalog, which contains one row per object with a volume of ≈ 20 kilobytes per row, is estimated to contain about 40×10^9 objects (even in the first full-sky data release). Adding to this the full `Source` and `Forced Source` catalogs, which contains one row per measurement in each of the ~ 80 visit images obtained per year, brings the total storage volume required up into the petabytes range, and will require a serious commitment of resources at the proposed iDAC. The evolution of catalog sizes over the 10-year LSST survey is depicted in Figure 1, from which it is evident that the catalog size for the final release is order 15 petabytes. For more details on the row counts see the Key Numbers Page[1].

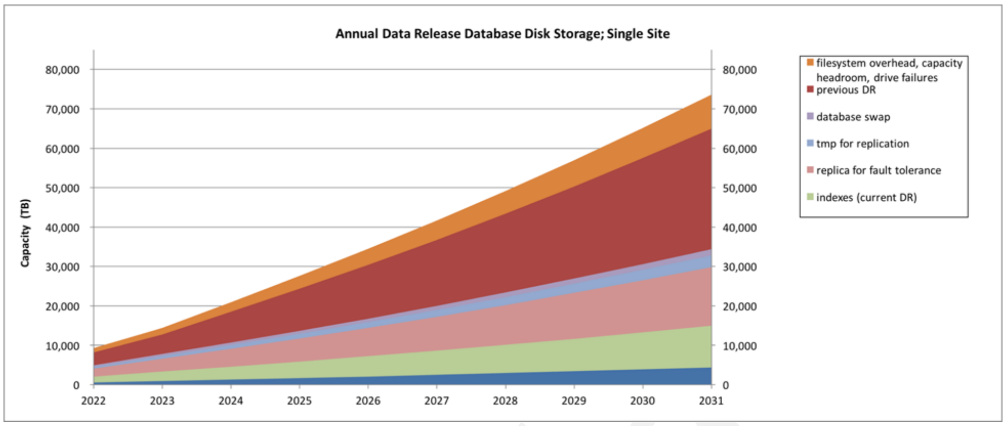

FIGURE 1: Catalog volume over time from LDM-144.

2.3 An "`Object Lite`" Catalog

Many – perhaps most – astronomers' science goals will be adequately served by a low-volume subset of the `Object` catalog's columns that do not include, for example, the full posteriors for the bulge+disk likelihood parameters. This `Object Lite` catalog would nominally contain 1840 bytes per row for the 40×10^9 objects, giving a size of $\approx 7.4 \times 10^{13}$ bytes (~ 74 terabytes). Even smaller, science-specific versions of `Object Lite` could be envisioned, with even less columns and/or separate star and galaxy catalogs. These would not be small enough to handle on a laptop, but might be served by a small departmental cluster (a mini-iDAC). Searching even a small `Object Lite` catalog would require some form of database, but many institutes would already have a system which may be capable of loading this data. In this case, LSST might only ship files with documentation and not provide administrative support for the system, but this would allow the `Object Lite` catalog to be widely available to all partner institution iDACs. Distribution options such as peer-to-peer networking to avoid download bandwidth limitations might be possible to implement in this case.

[1] https://confluence.lsstcorp.org/display/LKB/LSST+Key+Numbers

3 Requirements and Guidelines for iDACs

Since creating, delivering, and supporting the implementation of LSST data products in iDACs creates some cost to the LSST Project, iDACs will be expected to follow some basic requirements and guidelines that are described below. The actual costs of iDAC support and infrastructure development are considered separately in Section 4.

3.1 LSST sites topology

Leanne to describe the flow in the topology diagram.

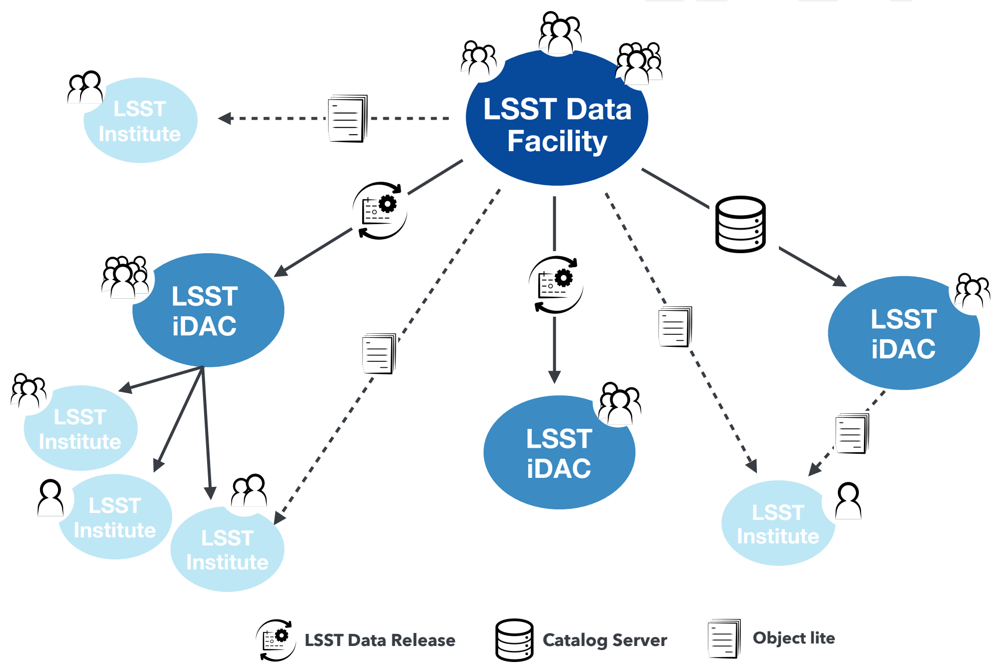

FIGURE 2: LSST Data Facility and Independent DAC network topology.

3.2 Data Storage

Any institution considering setting up an iDAC will need to show commitment on purchasing sufficient storage and CPU power to hold and serve the data. Sufficient storage ranges from 0.5 exabytes for the full data release(s) down to 100 terabytes for a catalog server, and potentially further down to 70 terabytes if the `Object Lite` option is offered. For the full catalog it is order 100 nodes to serve it up, and to serve images a DAC would need some additional servers; depending on load this may be order 10 additional nodes.

3.3 User Computational Resources

If the full set of data release products including images and catalogs are desired, it is highly recommended that the iDAC deploy the LSST Science Platform (LSP). The LSP serves as a portal to the data, and provides a user interface of web services and Jupyter notebooks for scientific queries and analysis, an open software framework for astronomy and parallel processing, and the associated computational, storage, and communications infrastructure needed to enable science. The LSP is described in full in LSE-319 and LDM-554. Depending on the assumed load, the LSP is relatively modest as it requires only ~ 2 servers to set up, and it is recommended to have 2 CPUs per simultaneous user (e.g., if the iDAC's desired capability is to serve 200 users, but only expect 50 to be active at a time, then 100 CPUs would be sufficient). From that starting point, the amount of next-to-the-data computational resources can be as large as the data center wishes to provide, and may make use of connecting to e.g., local super computer resources.

3.4 Dedicated Personnel

The significant hardware required by an iDAC is above the normal level for most astronomy departments, and would require dedicated technical personnel to set it up and keep it running. For an `Object Lite` catalog running on existing hardware, this might not be a significant increase in person power if the hardware is already serving on order $50–100$ terabytes. Still, it is recommended to assume $\gtrsim 0.25$ full-time equivalent (FTE) personnel hours for `Object Lite`, and perhaps closer to ~ 2 FTE for the full catalogs, which includes setting up and maintaining the service, and installing new data releases and software updates every year. For iDACs wishing to host the full data releases' images and catalogs and deploy the LSST Science Platform, it becomes necessary to employ $1–2$ storage engineers to mange the large amount of data, and possibly one more FTE to keep the Kubernetes (or equivalent) system updated with the latest software deploys. If the iDAC intends to support the science of many local users, support will become a specific issue which may not be covered by the usual institutional funding, and will require further effort. It is therefore recommended that any partner institution wishing to host a full-release iDAC provide a minimum personnel of 5 FTE to be considered viable.

3.5 Proprietary Data Access Policies

Defer for now until further guidance received.

All iDACs serving the proprietary LSST data products are subject to the policies in LPM-261. It might be necessary for iDACs to use the LSST authentication system to ensure secure access to the data for authorized users only, and this might require some specific IT work at the iDAC host site.

4 Cost Impacts

As previously mentioned, standing up and maintaining multiple iDACs comes at a significant cost impact to both the LSST Project and the partner institutions. Minimizing these costs – or at least maximizing the amount of science they enable – should be at the forefront of all considerations concerning partner iDACs, such as the following propositions.

4.1 Maximizing Profits with Science-Driven iDACs

There are two main cost impacts of iDACs being set up outside of the US and Chilean DACs: the positive impact is that some computational load may be taken off of these existing DACs, but the negative impact is the level of support required from the LSST Project in order to get them set up and running. This negative impact could be mitigated by ensuring that science productivity is maximized as a result of this extended effort. One way to do this might be to associate specific areas of science to a given iDAC, and encourage users working in that field to use that iDAC. This could create a customer base for the iDAC, bring together like-minded experts, and effectively distribute the computing load across a network of iDACs. This might also enhance internal funding arguments for investment resources by arguing for synergies with local science goals and attracting international users and official endorsement.

4.2 Data Transfer

Even with good networks the data transfer will not be trivial, and could be quite expensive. LSST is not currently set up to distribute data to multiple sites, i.e., there is no form of peer-to-peer sharing. The bandwidth at NCSA is adequate for receiving data and delivering `Alerts` to brokers during the night; perhaps some day time bandwidth could be used to transfer data to iDACs. A full data release of images and catalogs does not have to transferred within a given day; if the correct agreements are in place with an iDAC, a full release could be transferred slowly as it is produced, and then made available to the iDACs users in whole on the official release day.

4.2.1 Transfer cost use case

If we take the final number from the key numbers page [2] we could consider DR1 as about 6 PB (10% of the final size).

We would have at least two ways to transfer this : via the network, via physical devices.

A network transfer at 10Gbps of 6 PB would take $8 * 6 \times 10^{12}\ 10^{7} = 4.8 \times 10^{6}\ seconds$ or about 55 days[3]. Many institutes have 100 Gbps connections so this should be an upper limit and a transfer should be order one week. If we had a peer to peer network this may go down somewhat and we may be able to support it from NCSA.

Alternatively we could host the data on Amazon or Google and let people download it from there - they would have more capacity. Storage on the cloud for public data would be theoretically free - download (egress) would cost. Transfer to another cloud [4] or a Content Delivery Network (CDN)[5] end up costing about a cent a GB which for an open science project and at our volume should be negotiable. At one cent a transfer would cost $\sim \$0.01 * 6 \times 10^{12}\ 10^{6} = \$60K$.

For physical devices, today apparently we could get a device like Petarack `https://www.aberdeeninc.com/petarack/` for $300K. Theoretically we could get this cheaper though this is close to the drive price, Tape may also be a possibility especially if Sony/IBM commercialize high density tape with >300TB per cartridge[6]. A current 6TB cartridge is about $30, so enough tapes for 6PB would cost about 30K. If the density increased this could come down significantly. This could be be a partner data center cost as well as shipping it. Transfer of data on to this would be about the same as the network rate above so 7 days. SneakerNet [**?**] may still be cost

[2] `https://confluence.lsstcorp.org/display/LKB/LSST+Key+Numbers`
[3] day = 86400
[4] `https://cloud.google.com/storage/pricingi#network-pricing`
[5] `https://cloud.google.com/cdn/pricing`
[6] `https://newatlas.com/sony-ibm-magnetic-tape-density-record/50743/`

effective in the LSST era, which is predicted in the a paper.

4.3 Compute *vs.* Storage Resources

Data storage is a large cost to iDACs, and could be considered as an overhead relative to the amount of computational resources an iDAC can offer. If an iDAC is set up without a large compute capacity, the facility might be less useful to the science community than e.g., augmenting an existing DAC or iDAC to have more computational resources. It is conceivable that a partner institution may prefer to spend their money increasing the computational quotas available for a given collaboration or set of PIs, and it would be scientifically beneficial if this was possible at all DAC and iDACs. The notion of standard compute quotas and resource allocation committees to adjudicate on large proposals for substantial increases to computational allocations are described in LPM-261. Another way to approach a solution to this issue might be to have a *Cloud*-based iDAC where a user or PI could buy nodes on the provider cloud to access the holdings put there by LSST.

5 References

[1] **[LDM-141]**, Becla, J., Lim, K.T., 2013, *Data Management Storage Sizing and I/O Model*, LDM-141, URL `https://ls.st/LDM-141`

[2] **[LDM-135]**, Becla, J., Wang, D., Monkewitz, S., et al., 2017, *Data Management Database Design*, LDM-135, URL `https://ls.st/LDM-135`

[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] **[LDM-144]**, Freemon, M., Pietrowicz, S., Alt, J., 2016, *Site Specific Infrastructure Estimation Model*, LDM-144, URL `https://ls.st/LDM-144`

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

[7] **[LSE-163]**, Jurić, M., et al., 2017, *LSST Data Products Definition Document*, LSE-163, URL https://ls.st/LSE-163

[8] **[LDM-152]**, Lim, K.T., Dubois-Felsmann, G., Johnson, M., Jurić, M., Petravick, D., 2017, *Data Management Middleware Design*, LDM-152, URL https://ls.st/LDM-152

[9] **[LDM-572]**, O'Mullane, W., Petravick, D., 2017, *Chilean Data Access Center*, LDM-572, URL https://ls.st/LDM-572

[10] **[LPM-261]**, Willman, B., Graham, M., O'Mullane, W., Petravick, D., 2018, *Access Policy for LSST Data and Data Access Center*, LPM-261, URL https://ls.st/LPM-261

6 Acronyms

Acronym	Description
CDN	Content Delivery Network
CPU	Central Processing Unit
DAC	Data Access Center
DAX	Data access services
FTE	Full Time Equivalent
GB	GigaByte
IBM	International Business Machines
IT	Integration Test
LDM	LSST Data Management (Document Handle)
LPM	LSST Project Management (Document Handle)
LSE	LSST Systems Engineering (Document Handle)
LSP	LSST Science Platform
LSST	Large Synoptic Survey Telescope
NCSA	National Center for Supercomputing Applications
PB	PetaByte
PI	Principle Investigator
Qserv	Proprietary LSST Database system
US	United States