



LARGE SYNOPTIC SURVEY TELESCOPE

## Large Synoptic Survey Telescope (LSST) Butler working group charge

William O'Mullane, Tim Jenness

LDM-563

Latest Revision: 2017-08-02

**Draft Revision NOT YET Approved** – This LSST document has been approved as a Content-Controlled Document by the LSST DM Change Control Board. If this document is changed or superseded, the new document will retain the Handle designation shown above. The control is on the most recent digital document with this Handle in the LSST digital archive and not printed versions. Additional information may be found in the corresponding DM RFC. – **Draft Revision NOT YET Approved**

### Abstract

This is the charge for the Butler Working Group to be convened in August 2017.

## Change Record

Version	Date	Description	Owner name
1.0	2017-08-02	Initial version.	W. O'Mullane

*Document source location:* <https://github.com/lsst/LDM-563>

Draft

## 1 Scope

This Working Group is to look into the Butler layer of the Data Management System [LDM-148, LDM-463]. It should start immediately (August 4, 2017) and finish its remit by 29th September 2017.

## 2 Responsibilities

The Working Group (WG) has the following responsibilities:

- Take input from a broad range of interested parties on the butler abstraction layer, and produce a document of Use Cases.
- Pin down the requirements in a written document [LDM-556].
- Rank butler requirements in terms of priority in time as well as which are nice to have features as opposed to "needed").
- Define mode of work in DM to achieve implementation of the Butler.
- The WG Chair shall convene meetings on a regular basis.
- The WG will make a recommendation as to whether a new Butler should be written, or whether the existing Butler code can be extended to meet the new requirements.
- The WG will draft a document proposing a design and implementation.

### Comment for draft

I am somewhat concerned that we need a document proposing a design and implementation by the end of September.

## 3 Specific tasks

### 3.1 Draft requirements document(s)

LDM-556 defines all the middleware requirements, and is generated from the MagicDraw

SysML model. A section exists in LDM-556 that will be populated with Butler requirements. These requirements should carry priorities as used in LSE-61.

#### Comment for draft

In LSE-61 we only prioritize by delivery time. We do not have a secondary axis to prioritize by importance. If we want two axes of prioritization I need to get the SysML stereotype updated.

## 3.2 Plan to go forward

The path forward should be identified before the Working Group finishes. This should include advice to the PM on the distribution of work among institutes (if any) and indication of individuals responsible for different components.

## 3.3 Specific topics to be considered

1. **Operational considerations** The group must consider how to take care of operational constraints at NCSA.
2. **Data access** One major goal of Butler is to provide abstraction to data access – that could be with multiple implementations. How do developers access local data while allowing Condor jobs to access data on GPFS?
3. **Data format abstraction** How do we hide from the pipelines user whether data are stored in HDF5 or FITS, or whether associated metadata is coming from a database or from a file?
4. **Data discovery** This includes calibration logic and mappers – the distinction between simple data access (see 2) and associating data items with each other should be made clear.
5. **Expansion of partial data IDs** This must be well explained and pinned down for operations.
6. **Third-party instrument support** How can we make it easier for others to use the stack
7. **Subset data staging**

## 4 Membership

Membership on the order of seven people is optimal. The proposed membership is:

- Tim Jenness (**Chair**),
- John Swinbank,
- Maria Patterson and/or Colin Slater (UW),
- Jim Bosch and/or Pim Schellart (Princeton),
- Michelle Gower,
- Brian van Klaveren.

In addition K-T Lim is available as adviser.

### Comment for draft

There is no-one here from SQuaRE or SUIT. Both those groups have been strong proponents of butler requirements in the past. For example, SQuaSH needs to be able to retrieve data from different runs for comparisons. I feel we should also ask Dominique Boutigny whether he would like to be an external "non-voting" member. They have had a lot of experience with the current butler when doing processing at IN2P3.

## 5 Reporting

Th WG Chair shall report directly to the DM Project manager weekly.

## 6 References

- [1] **[LDM-463]**, Becla, J., Pease, N., 2017, *Data Access Design*, LDM-463, URL <https://ls.st/LDM-463>
- [2] **[LSE-61]**, Dubois-Felsmann, G., 2016, *LSST Data Management Subsystem Requirements*, LSE-61, URL <https://ls.st/LSE-61>
- [3] **[LDM-556]**, Dubois-Felsmann, G., 2017, *Data Management Middleware Requirements*, LDM-556, URL <https://ls.st/LDM-556>
- [4] **[LDM-148]**, Lim, K.T., Bosch, J., Dubois-Felsmann, G., et al., 2017, *Data Management System Design*, LDM-148, URL <https://ls.st/LDM-148>