Rubin Observatory

Vera C. Rubin Observatory
Data Management

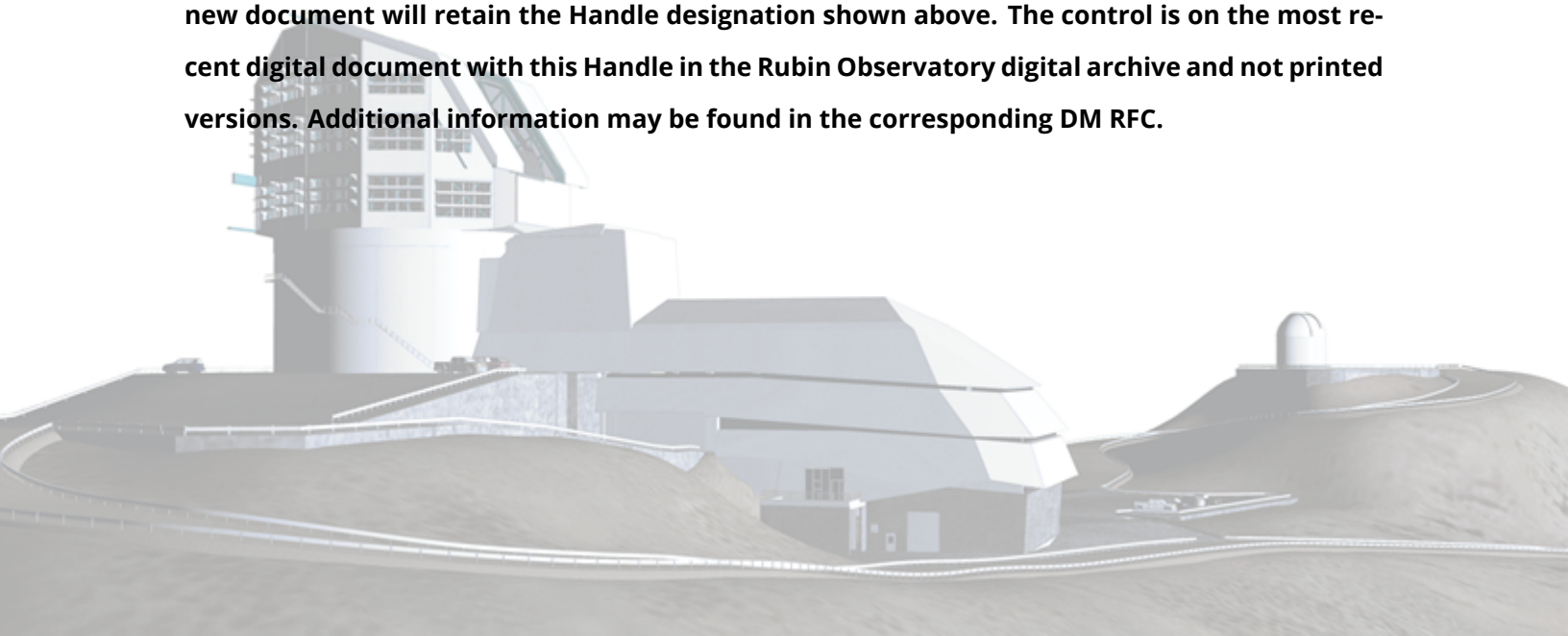
Proposed DM OPS Rehearsals

Margaret Johnson and Robert A. Gruendl

LDM-643

Latest Revision: 2020-08-19

This Rubin Observatory document has been approved as a Content-Controlled Document by the Rubin Observatory 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 If this document is changed or superseded, the new document will retain the Handle designation shown above.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 Rubin Observatory digital archive and not printed versions. Additional information may be found in the corresponding DM RFC.



Rubin Observatory

Abstract

For discussion about better specifying Operations Rehearsals.

The contents of this document are subject to configuration control by the Rubin Observatory DM Change Control Board.

Rubin Observatory

Change Record

Version	Date	Description	Owner name
	2018-07-16	Initial version.	Robert Gruendl
	2019-01-08	Revisions based on initial comments.	Robert Gruendl
	2019-07-10	Updates after Ops Rehearsal #1. Preliminary specification of Rehearsals #2 & #3.	Robert Gruendl
	2020-01-15	Updates for OPS Rehearal #2.	Robert Gruendl
	2020-06-26	Updates for OPS Rehearal #2 for ComCam at Base	Robert Gruendl
1.0	2020-08-19	First version via RFC-708	WOM

Document curator: Robert Gruendl

Document source location: `https://github.com/lsst/LDM-643`

Rubin Observatory

Contents

1 Introduction **1**

2 LDM 503-09: Operations Rehearsal #1 for Commissioning **2**

2.1 An Updated Goal: 2

2.2 Pre-Requisites: 3

2.2.1 Pre-Requisites: Roles: 3

2.2.2 Pre-Requisites: Services & Service Components: 4

2.2.3 Pre-Requisites: Work: 4

2.3 Rehearsal Outline: 5

2.4 Assess: 6

2.5 Addendum: 6

3 LDM 503-11: Operations Rehearsal #2 for Commissioning **7**

3.1 An Updated Goal: 7

3.2 Pre-Requisites: 8

3.2.1 Pre-Requisites: Roles: 8

3.2.2 Pre-Requisites: Services & Service Components: 9

3.2.3 Pre-Requisites: Work: 10

3.3 Rehearsal Outline: 10

3.4 Assess: 11

4 LDM 503-12: Operations Rehearsal #3 for Commissioning **13**

4.1 An Updated Goal: 13

Rubin Observatory

Proposed DM OPS Rehearsals

1 Introduction

As LSST DM moves from construction through commissioning and into operations a number of rehearsals have been proposed to help prepare for the execution of the survey. Specific rehearsals are outlined in [LDM-503] but in the more comprehensive cases, i.e. the Operations (Ops) Rehearsals (LDM 503-09, LDM 503-11, and LDM 503-12), the contents of those document alone, do not sufficiently outline the scope, content, action and interaction that are being rehearsed. From the software side, [LDM-564] summarizes the DM software features that should be available and helpfully identifies those software releases in the context of the rehearsals. However, the Ops Rehearsals are not simply periods to test hardware and software systems, they are opportunities to develop and understand operations processes and to observe the interactions of hardware, software and personnel.

This document attempts to outline the Ops Rehearsals in greater detail for the following reasons:

- Depending on the impact, missing or late software features and hardware systems may either require mitigation (e.g., shims, fake data, etc...) or might even be grounds for postponement.
- The purpose of an Ops Rehearsal is not to debug a freshly deployed system, but rather to understand whether that system does what is needed.
- The effort to carry out rehearsals will require coordination of personnel and facilities.

Note: This remains a work in progress. The current draft attempts to describe the process for the first Ops rehearsal. The subsequent rehearsals have only been outline to the extent to properly understand their scope.

The contents of this document are subject to configuration control by the Rubin Observatory DM Change Control Board.

2 LDM 503-09: Operations Rehearsal #1 for Commissioning

Nominal Date: April 2019

Original Description:

- Choose TBD weeks during commissioning to simulate.
- Pick which parts of plan we could rehearse.
- The commissioning team (via Chuck Claver) suggests Instrument Signal Removal should be the focus of this (or the next rehearsal).

2.1 An Updated Goal:

The primary goal is to simulate nominal operations, both daytime and nighttime, for a 2–3 day period including the daily meeting(s) that would occur among the SciOps and Data Facility staff. These activities will be accompanied by simulated observations obtained in a "sampling" mode in order to exercise:

1. the transfer, archiving and ingestion of raw data
2. offline processing of calibrations and science data
3. curation of the resulting data products

At the time of this initial rehearsal, we do not expect a functioning observatory system, instead:

- Sampling mode has been used to describe early LSST commissioning observations where observations occur based on the needs of the commissioning team. Such observations would typically include some basic set of calibrations (e.g. biases and flats) followed by nighttime observations that might be used to test and quantify system performance.
- A basic set of raw data will be transferred to a mountaintop computer which will then in turn mimic observations by sending those images from the summit to NCSA via the long-haul networks.
- The contents of the dataset will be minimal (no larger than the calibration and nightly observations that might be expected from ComCam). The current plan is to draw from a

suitable set of test-stand data and then present these as though they were coming from the telescope.

- On arrival at the LDF the observations will be ingested into the current data-backbone which can in turn be used to feed the data through a batch production service to produce "calibrations" and "reduced science products."
- In the context of this rehearsal, the sophistication (or correctness) of the pipeline(s) is not paramount. What is important is that the raw and resulting products are tracked and can be superficially examined by LDF and SciOps team members. The degree of realism would depend on both the data being sent and availability of working pipeline tasks.

2.2 Pre-Requisites:

There are three broad categories of pre-requisites that are needed:

1. Persons must be identified to fill roles within the rehearsals.
2. Services (or facsimiles) need to exist that will be used/tested throughout the rehearsal.
3. Elements that would not otherwise be available in the pre-operations LSST project will be prepared.

2.2.1 Pre-Requisites: Roles:

Persons need to be identified that will staff various roles in this rehearsal. These roles are drawn from those in operations which come from three groups: Observatory Operations (ObsOps), Science Operations (SciOps), and the LSST Data Facility (LDF).

- Coordinator: This person acts as an independent executor of the rehearsal. They would be responsible for executing outside actions that drive the simulation (e.g., initiating a script that would start data flowing from summit to LDF).
- ObsOps, Observing Specialist
- SciOps, QA Scientist
- SciOps, Verification and Validation Scientist

- LDF, Operator
- LDF, QA Scientist
- LDF, Admin

2.2.2 Pre-Requisites: Services & Service Components:

The broader pieces of the DM system that need to operate for this rehearsal are:

- A service must operate at the mountaintop that will send data. This can be as simple as a shell script that draws from a list of files and transfers them to NCSA with some cadence.
- Nominally the long-haul networks need to be available at the time of this rehearsal. (Note: at the nominal time of this rehearsal we can only expect transfer rates (BASE to LDF) of order 10 MB/s. Therefore, ~500 raft-scale images should require ~8 hrs. In addition, outages due to movement of equipment at the base may occur. A copy of test data should be kept at LDF to mitigate data transfer problems/outages during the rehearsal.)
- A data backbone endpoint to receive and ingest incoming files must exist.
- A mechanism must exist to distribute jobs to a compute resource to process the "new" data–Batch Production.
- A workflow system to configure and launch jobs must exist.
- Pipeline(s) to processes the data must be in place.
- A minimally functional science platform where raw and processed data products can be examined by staff must exist.

Additional monitoring services for the networks, batch production, compute resource(s), and data backbone are desirable.

2.2.3 Pre-Requisites: Work:

Work that must be completed prior to Ops Rehearsal but which is not part of the DM development:

- Generate a mock data set. This must have the ability to be ingested with either Gen2 or Gen3 Butler. It is not necessary that the generated data products be curated for a long period.
- Create a shim service that sends data from summit to LDF.
- Specify appropriate pipeline(s) that will be run during the rehearsal.
- Test that services in the preceding section can adequately function for the purposes of this rehearsal.
- Allocate compute and storage resources and specify location of stored products.
- Location to record information about incidents, problem backlog, processing and QA summaries (for the initial test this could be as simple a set of confluence pages).

2.3 Rehearsal Outline:

During normal operations the time observing occurs depends on local nighttime in Chile. This is not necessary for the rehearsal and so data delivery and can be shifted to occur in a normal working day. Prior to the execution of the rehearsal the work outlined in Section 3.2.3 must be completed and tested.

A basic outline of the processes that would occur during this rehearsal follows:

1. (ALL: ObsOps+SciOps+LDF) afternoon stand-up operations meeting
2. (ObsOps) mock transmit nightly calibration exposures to LDF for ingestion
3. (LDF) generate nightly master calibrations
4. (SciOps) select configuration and calibrations
5. (ObsOps) mock transmit nightly science images and ingest
6. (LDF) run science pipeline (.e.g. ISR) in offline/batch mode
7. (LDF) generate processing reports for discussion in stand-ups
8. (SciOps) examine input and output data from nightly observations and processing
9. (SciOps) generate quality reports for discussion in stand-ups

10. (ALL) monitor progress of nightly "campaigns," characterize and assess, make records of failures, diagnose issues, generate problem backlog
11. (ALL) create mock nightly reports
12. repeat (a total of 3 times)

While there are "realistic" components within the outline, much of focus by the actors should be on the processes. How is this going to look in day-to-day operations? If there are problems, what happens? Who gets a call and when? What information needs to be available between a geographically distributed team (and when)? Are the lines of communication between those groups sufficient?

2.4 Assess:

Among the activities in the Rehearsal Outline it is expected that some might influence the long-term development within DM. An example is exercising tools and services (e.g. the LSP) with a mind toward operational needs. Another example is to inform the processes and metrics needed to make decisions about configuration and calibration selection in the context of both production success and production failure.

Example questions that can be asked during the assessment phase are:

- Was the rehearsal successful? How long did it take? What anomalies/failure modes were identified, and how did the team cope?
- What fixes are needed, and on what timescale (e.g., next ops rehearsal, or we are go for commissioning)?
- What improvements in procedures, documentation, frameworks, systems, and algorithms were identified and at what priority?
- How is time and effort budgeted to plan and execute priority changes and improvements? How will the next rehearsal be planned?

2.5 Addendum:

Operations Rehearsal #1 occurred in May 2019. A short note, DMTN-119, gives a summary report of its execution.

3 LDM 503-11: Operations Rehearsal #2 for Commissioning

Nominal Date: July 2020

Original Description:
More complete commissioning rehearsal:

- How do the scientists look at data?
- How do they provide feedback to the telescope?
- How do we create calibrations?
- How do we update calibrations?

3.1 An Updated Goal:

The primary goal is to rehearse for commissioning operations prior to the ComCam verification and validation era (including the mini-surveys). Unlike Ops Rehearsal #1, we will be emulating daytime operations, for a 3–5 days, would include daily meetings, exercise data movement and processing. Additionally this rehearsal could include: application of software changes, simulated outages, or non-standard (unprocessable) engineering observations. As summit operations are restored, alternative or extended follow-on rehearsal activities should be considered utilizng AuxTel/LATISS or ComCam installed at the summmit and to supplement these exercises.

In the current time frame of this rehearsal, we do not expect a functioning telescope + camera. Instead:

- Despite the COVID-19 shutdown of summit activities, the project has managed to make ComCam available from the Base Facility. We will attempt to exercise as many elements of the emergent system as possible. Primary among these are automatic data transfer to the LDF (and if possible, to Base compute resources through the OODS).
- The contents of the dataset will roughly match those expected during ComCam verification activities, except that they will be confined to the calibration sequences possible from the temporary installation at the Base.

- On arrival at the LDF the observations will be ingested into the current data-backbone which can in turn be used to feed the Calibrations Products Pipelines (CPP) to produce "calibrations" and "reduced calibration products." As a stretch goal, a rudimentary prompt processing system can be attempted, meaning processing that would be triggered by arrival of new observations, rather than directly commanded by a user.
- Similar to the Ops Rehearsal #1, the sophistication (or correctness) of the pipelines are not paramount. What is important is that the raw and resulting data products are tracked and can be superficially examined by LDF and SciOps team members. The degree of realism would depend on both the data being sent and availability of working pipeline tasks.
- Once ComCam is available on the test stand (at the summit) and/or available on the telescope, follow-on rehearsal(s) should be considered to further exercise capabilities at raft-scale and better prepare for ComCam during commissioning. Depending on the camera location, processing would focus on again exercising CPP, or begin exercising pipeline reductions amenable to early on-sky data.

3.2 Pre-Requisites:

There are three broad categories of pre-requisites that are needed:

1. Persons must be identified to fill roles within the rehearsals.
2. Services (or facsimiles) need to exist that will be used/tested throughout the rehearsal.
3. Elements that would not otherwise be available in the pre-operations LSST project will be prepared.

3.2.1 Pre-Requisites: Roles:

Persons need to be identified that will staff various roles in this rehearsal. These roles are drawn from those in operations which come from three groups: Observatory Operations (ObsOps), Science Operations (SciOps), and the LSST Data Facility (LDF).

- Coordinator: This person acts as an independent executor of the rehearsal. They would be responsible for executing outside actions that drive the simulation (e.g., initiating a script that would start data flowing from summit to LDF).

- ObsOps, Site Astronomer
- ObsOps, Observing Specialist
- SciOps, QA Scientist
- SciOps, Verification and Validation Scientist
- LDF, Operator
- LDF, QA Scientist
- LDF, Admin

3.2.2 Pre-Requisites: Services & Service Components:

The broader pieces of the DM system that need to operate for this rehearsal are:

- Data transfer service must operate at the mountaintop for AuxTel that will send data. The method most recently used when LATISS was on the test-stand in Tucson, rsync, is sufficient for this task, and can act as a backup if forwarding through the camera DAQ is not yet available/reliable.
- Nominally the long-haul networks need to be available at the time of this rehearsal. (Note: at the time of this rehearsal we can only expect transfer rates (BASE to LDF) of order 40 GB/s. Therefore, raft-scale observations should require only a few seconds to transmit.)
- A data backbone endpoint to receive and ingest incoming files must exist.
- An automated mechanism to ingest into a Butler, on-arrival at an endpoint (OODS at BASE, DBB at LDF) must exist. Currently a mechanism for this exists assuming a Gen2 Butler at the LDF.
- Processing can occur either though an automated mechanism, i.e. jobs start after data ingestion or through a batch-production mode. Batch production (assuming Gen2 Butler) has already been demonstrated, if Gen3 Butler is used or on arrival processing is deemed necessary then shims may be necessary.
- Pipeline(s) to processes the data must be in place.
- A functional science platform (LSP) where raw and processed data products can be examined by staff is needed and already exists.

Exercising monitoring services for the networks, batch production, compute resource(s), and data backbone can be treated as a stretch goal.

3.2.3 Pre-Requisites: Work:

Work that must be completed prior to Ops Rehearsal but which is not part of the DM development:

- Effort to ensure that shims are available to facilitate on-arrival and/or batch processing.
- Specify appropriate pipeline(s) that will be run during the rehearsal.
- Use early ComCam test-stand data to test that pipelines are appropriately configured for the rehearsal.
- Make compute and storage resources are available, including a reservation for compute resoureces to ensure timely processing can occur.
- Create a central location (e.g. Confluence page) for the rehearsal to record information about incidents, problem backlog, processing and QA summaries.

3.3 Rehearsal Outline:

Since this proposal revolves around test-stand observations with ComCam, the activites can be centered around daytime (i.e., shifted in time compared to the normal sequence of event during Operations). This means the communications chains among the Observing Specialist (nighttime), Site Astonomer (daytime), and LDF Operator will be exercised, but in such a way that the fit within a normal workday. Currently automated feedback from processing (at LDF) to the observing team has not been provisioned within the system.

A basic outline of the processes that would occur during this rehearsal are given below, starting from the time that afternoon calibrations are acquired on the summit.

1. # Morning
2. (ObsOps) Acquire “afternoon” calibrations.
3. (LDF) Calibrations Products Pipeline runs.
4. (SciOps) select configuration and calibrations.

5. # Noon
6. (ALL: ObsOps+SciOps+LDF) 'Afternoon stand-up operations meeting.
7. (ALL: ObsOps+SciOps+LDF) Closeout previous night's observing report
8. (LDF) generate processing reports for discussion in stand-ups
9. (SciOps) examine input and output data from nightly observations and processing,
10. (SciOps) generate quality reports for discussion in stand-ups
11. repeat (between 3-5 times)

Similar to OPS Rehearsal #1, there are "realistic" components within the outline, but much of focus by the actors should be on the processes, meetings, and communications. At a minimum the actors should be considering how well this will work in day-to-day operations? on weekends? How problem are addressed? Who gets a call and when? What information needs to be available between a geographically distributed team (and when)? Are the lines of communication between those groups sufficient?

An example set of calibration sequences, assuming a 3 day rehearsal with ComCam at the Base could be:

- Day 1: Normal Calibration Sequence: N×Bias, N×Flat, N× Dark.
- Day 2: Change/move illumination source: repeat Day 1.
- Day 3: Repeat Day 1, but interrupt change illumination during Flat Sequence, and repeat a New flat sequence.

3.4 Assess:

Among the activities in the Rehearsal Outline it is expected that some might influence the long-term development of the processes that will occur during operations but also might provide feedback in the DM development effort. An example is exercising tools and services (e.g. the LSP) with a mind toward operational needs. Another example is to inform the processes and metrics needed to make decisions about configuration and calibration selection in the context of both production success and production failure.

Example questions that can be asked during the assessment phase are:

- Was the rehearsal successful? How long did it take? What anomalies/failure modes were identified, and how did the team cope?
- What fixes are needed, and on what timescale (e.g., next ops rehearsal, or we are go for commissioning)?
- What improvements in procedures, documentation, frameworks, systems, and algorithms were identified and at what priority?
- How is time and effort budgeted to plan and execute priority changes and improvements? How will the next rehearsal be planned?

4 LDM 503-12: Operations Rehearsal #3 for Commissioning

Nominal Date: August 2021

Original Description:
Dress rehearsal: commissioning starts in April so by this stage we should be ready to do everything needed.

4.1 An Updated Goal:

Here the primary goal is to rehearse for commissioning operations prior to LSSTCam start of integration and test (i.e. while LSSTCam is on the summit but not yet integrated on the telescope). Similar to Ops Rehearsal #2, we would emulate both daytime and nighttime, for a 3–5 days, would include daily meetings, exercise data movement and processing. Additionally this rehearsal could include: application of software changes, simulated problems, or non-standard (unprocessable) engineering observations.

- LSSTCam should be at the summit in the clean room on its test stand. LSSTCam would be exercised with its Camera Control System to obtain test-stand images and send them through the DAQ for archiving and batch processing. This could be supplemented with on-sky data from ComCam to exercise pipeline processing.
- The contents of the data would roughly match those expected during LSSTCam verification activities but the use of on-sky data from ComCam would not be supplemented (to "simulate" data volume) but real-time processing could be exercised.
- On arrival at the LDF the observations will be ingested into the data-backbone which can in turn be used to feed the data through a batch production service to produce calibrations, reduced science products, and quality assessments.
- Similar to the other Ops Rehearsal #1, the sophistication (or correctness) of the pipelines are not paramount. What is important is that the raw and resulting data products are tracked and can be examined by LDF and SciOps team members. The degree of realism would depend on both the data being sent and availability of working pipeline tasks.

Rubin Observatory

References

[LDM-564], O'Mullane, W., Economou, F., Jenness, T., Loftus, A., 2018, *Data Management Software Releases for Verification/Integration*, LDM-564, URL https://ls.st/LDM-564

[LDM-503], O'Mullane, W., Swinbank, J., Jurić, M., Economou, F., 2018, *Data Management Test Plan*, LDM-503, URL https://ls.st/LDM-503