Rubin Observatory

**Vera C. Rubin Observatory
Data Management**
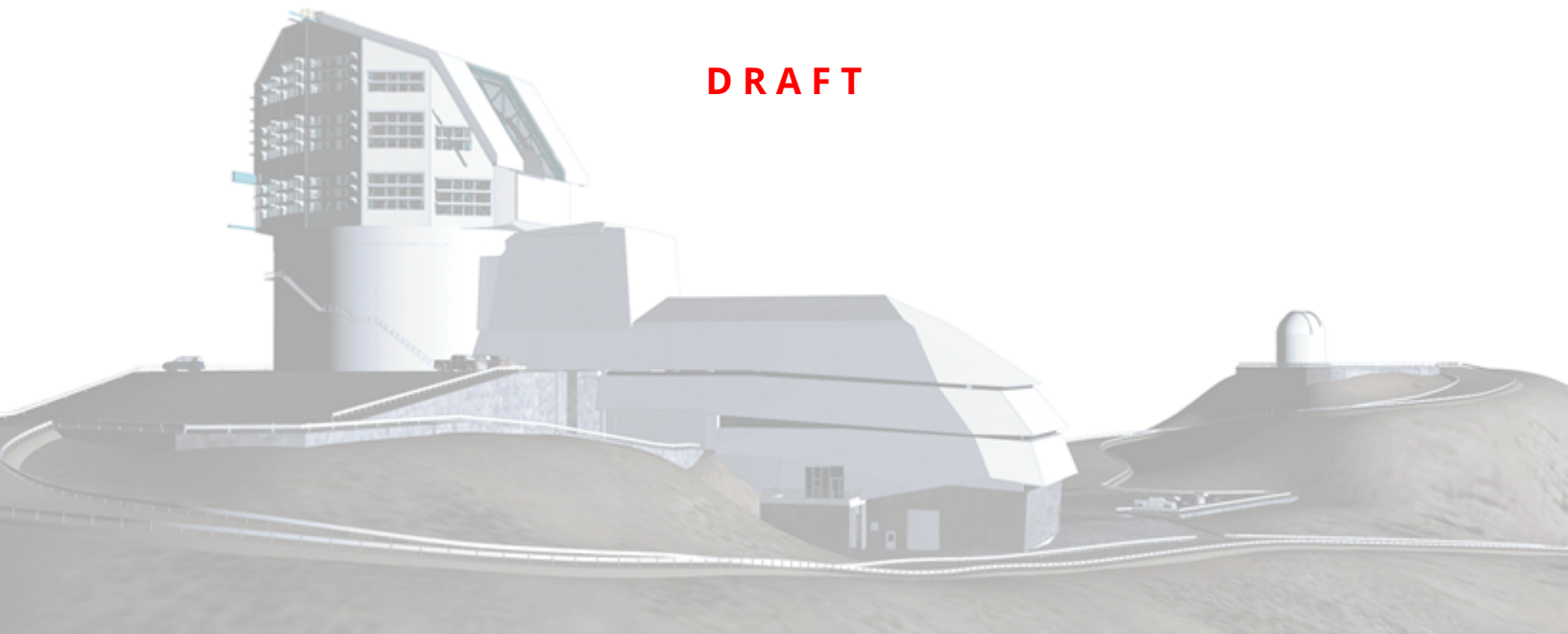
LDM-503-GEN3-01: Gen 3 Butler ready for internal DM use Test Plan and Report

Robert Gruendl

DMTR-271

Latest Revision: 2020-12-01

D R A F T

Abstract

This is the test plan and report for **Gen 3 Butler ready for internal DM use** (LDM-503-GEN3-01), an LSST milestone pertaining to the Data Management Subsystem.

Change Record

Version	Date	Description	Owner name
	2020-10-30	First draft	Robert Gruendl

Document curator: Robert Gruendl

Document source location: `https://github.com/lsst-dm/DMTR-271`

Version from source repository: ebb2418

Rubin Observatory

Contents

Rubin Observatory

Rubin Observatory

LDM-503-GEN3-01: Gen 3 Butler ready for internal DM use Test Plan and Report

1   Introduction

1.1   Objectives

The goal of this test is to demonstrate that the Gen3 Butler software project has sufficiently matured that subsequent DM development should begin focusing on adoption of Gen3 Butler software are repositories throughout the DM software project (i.e. that deprecation of Gen2 Butler usage within the project can begin).

1.2   System Overview

The Gen3 refactoring of the Butler is central to evolution of the overall DM software design and has repercussions throughout the rest of the DM project.   This test plan is designed to verify that minimal requirements have been met and the DM project can now begin the process of integrating the Gen3 Butler within the pipelines and analysis tools.   Those minimal requirements are that:

1. possible to ingest all raw dataset types currently hosted at NCSA (for supported instruments)
2. cp_pipe equivalent under Gen3 is available
3. developers can run a pipeline with a single-node using pipetask
4. an RC2-like processing is possible
5. a 3-tract RC2 test run can be performed in a reasonable time (under a week)
6. Calibration Product Pipelines (CPP) can be run at scales to support above investigations
7. Batch Processing System (BPS) is available to support testing at larger scales

In addition, at the time these tests occur the Gen3 Butler schema be considered stable enough that changes no longer occur on a weekly basis (i.e force re-ingestion/migration of existing repositories are no longer a weekly occurrence). Changes requiring wholesale reingestion/migration may still be required but will occur in a regimented manner.

Applicable Documents:
LDM-592: Data Access Use Cases
LDM-556: Data Management Middleware Requirements
LDM-639: Data Management Acceptance Test Specification

1.3   Document Overview

This document was generated from Jira, obtaining the relevant information from the LVV-P77 Jira Test Plan and related Test Cycles (LVV-C160 LVV-C162).

Section 1 provides an overview of the test campaign, the system under test (Software Products), the applicable documentation, and explains how this document is organized. Section 2 provides additional information about the test plan, like for example the configuration used for this test or related documentation. Section 3 describes the necessary roles and lists the individuals assigned to them.

Section 4 provides a summary of the test results, including an overview in Table 2, an overall assessment statement and suggestions for possible improvements. Section 5 provides detailed results for each step in each test case.

The current status of test plan LVV-P77 in Jira is **Draft** .

1.4   References

[1] **[LDM-556]**, Dubois-Felsmann, G., Jenness, T., Bosch, J., et al., 2018, *Data Management Middleware Requirements*, LDM-556, URL `https://ls.st/LDM-556`

[2] **[LDM-639]**, Guy, L., 2018, *DM Acceptance Test Specification*, LDM-639, URL `https://ls.st/LDM-639`

[3] **[LDM-592]**, Jenness, T., Bosch, J., Gower, M., et al., 2017, *Data Access Use Cases*, LDM-592,

Rubin Observatory

URL `https://ls.st/LDM-592`

2   Test Plan Details

2.1   Data Collection

Observing is not required for this test campaign.

2.2   Verification Environment

These tests assume a stable weekly stack which supports Gen3 running of the above, that
services that automatically ingest new data can support on-going ingestion to Gen3 reposito-
ries (i.e. DBB shared spaces and OODS support serving data through Gen3), and that batch
processing services can support pipeline execution of Gen3 products.

2.3   Related Documentation

No additional documentation provided.

2.4   PMCS Activity

Primavera milestones related to the test campaign:

- LDM-503-GEN3-01

3   Personnel

The personnel involved in the test campaign is shown in the following table.

	T. Plan LVV-P77 owner:	**Robert Gruendl**	
	T. Cycle LVV-C160 owner:	**Robert Gruendl**	
Test Cases	**Assigned to**	**Executed by**	**Additional Test Personnel**
LVV-T1985	Leanne Guy		
LVV-T1984	Leanne Guy		
LVV-T1982	Leanne Guy		
LVV-T1987	Leanne Guy		
LVV-T1986	Leanne Guy		
LVV-T1983	Leanne Guy		
	T. Cycle LVV-C162 owner:	**Undefined**	
Test Cases	**Assigned to**	**Executed by**	**Additional Test Personnel**
LVV-T2007	Gabriele Co-moretto		

4   Test Campaign Overview

4.1   Summary

T. Plan LVV-P77:	**LDM-503-GEN3-01: Gen 3 Butler ready for internal DM use**			Draft
T. Cycle LVV-C160:	**LDM-503-GEN3-01: Gen 3 Butler ready for internal DM use**			Not Executed
Test Cases	**Ver.**	**Status**	**Comment**	**Issues**
LVV-T1985	1	Not Executed		
LVV-T1984	1	Not Executed		
LVV-T1982	1	Not Executed		
LVV-T1987	1	Not Executed		
LVV-T1986	1	Not Executed		
LVV-T1983	1	Not Executed		
T. Cycle LVV-C162:	**LDM-503-GEN3-01: Gen 3 Ingest ComCam raw dataset**			Not Executed
Test Cases	**Ver.**	**Status**	**Comment**	**Issues**
LVV-T2007	1	Not Executed		

Table 2: Test Campaign Summary

4.2   Overall Assessment

Not yet available.

4.3   Recommended Improvements

Not yet available.

Rubin Observatory

5   Detailed Test Results

5.1   Test Cycle LVV-C160

Open test cycle *LDM-503-GEN3-01: Gen 3 Butler ready for internal DM use* in Jira.

Test Cycle name: LDM-503-GEN3-01: Gen 3 Butler ready for internal DM use
Status: Not Executed

Tests that Gen3 software maturity now supports development across the DM project.

[GCM]
I see here 2 main groups of test:
i- ingestion of different raw dataset:
       in case we have one single test case for ingestion of data in Gen3 (T1985), we need to
have one test cycle for each type of raw data ingested
ii- various pipeline processing
      all these test cases can be part of a single test cycle. In case T1985 is divided in many test
cases, one per dataset, those tests can also be included here.
       These test cases seem to be generic Science Pipelines test cases, ae we sure they have
not already been defined and executed?

5.1.1   Software Version/Baseline

Not provided.

5.1.2   Configuration

Gen3 Butler repositories with test data are available within DBB spaces.   Weekly DM stack
has Gen3 and BPS elements present for tests.

Rubin Observatory

5.1.3   Test Cases in LVV-C160 Test Cycle

5.1.3.1   LVV-T1985 - Ingest all the raw datasets that are currently hosted at NCSA for supported instruments.

Version **1**. Open *LVV-T1985*   test case in Jira.

[GCM]
The objective of this test, considering it related to the Daf Butler software product, should be to prove that the raw data ingest functionality is available and working.
This implies this test should be independent of which raw data is ingested, making in this way the test procedure independent and reusable.
The type of data to ingest should be defined case by case in the Test Cycle where this test is executed.
I suggest therefore to change the name of this test to: "Verify daf_butler raw data ingest".
There should be the following steps:
- verify that the Gen3 repository is available
- ingest the raw data defined in the test cycle (to be divided in multiple steps in case it makes sense)
- verify that the raw data is available in the repository as expected
In the case that the second step is different, depending on the raw dataset, at this point it is necessary to have a different test case for each of them, and multiple test cycles are not needed.

To note that, the original name "Ingest all the raw datasets that are currently hosted at NCSA..." looks more to be a milestone, than a test case definition.

 Preconditions:
In order to run this test, a Gen3 daf butler should be deployed and ready to use, with access to the filesystems where the raw data to ingest is stored.

Execution status: **Not Executed**

Final comment:

Detailed steps results:

Step 1	Step Execution Status: **Not Executed**
	Description
	Expected Result
	Actual Result

5.1.3.2 LVV-T1984 - Provide a recipe and example for a cp_pipe equivalent

Version **1**. Open *LVV-T1984* test case in Jira.

Preconditions:

Execution status: **Not Executed**

Final comment:

Detailed steps results:

Step 1	Step Execution Status: **Not Executed**
	Description
	Expected Result
	Actual Result

5.1.3.3   LVV-T1982 - Run a pipeline on a single node using pipetask.

Version **1**. Open *LVV-T1982*   test case in Jira.

To show that individual users have the ability to run either locally (w/ sqlite) or generally (w/ Postgres) using Gen3 Butler infrastructure.

 Preconditions:
This test requires that Gen3 Butler infrastructure and underlying pipets have been integrated. It further requires (in spirit) that gen3 schema stability has been reached to facilitate comparison of pipeline results with further stack development can be compared.

Execution status: **Not Executed**

Final comment:

Detailed steps results:

Step 1	Step Execution Status: **Not Executed**
	Description
	Expected Result
	Actual Result

5.1.3.4   LVV-T1987 - Run Calibration Products Processing (CPP)

Version **1**. Open *LVV-T1987*   test case in Jira.

Tests should be run at scales corresponding to appropriate scales (e.g DC2, RC2)

Rubin Observatory

Preconditions:

Execution status: **Not Executed**

Final comment:

Detailed steps results:

Step 1	Step Execution Status: **Not Executed**
	Description
	Expected Result
	Actual Result

5.1.3.5   LVV-T1986 - Run DC2-like processing (not necessarily full-scale)

Version **1**. Open *LVV-T1986*   test case in Jira.

Preconditions:

Execution status: **Not Executed**

Final comment:

Detailed steps results:

Step 1	Step Execution Status: **Not Executed**
	Description

Rubin Observatory

	Expected Result

	Actual Result

5.1.3.6   LVV-T1983 - Run a 3 tract dataset (RC2) in under a week on NCSA Condor system using BPS

Version **1**. Open *LVV-T1983*   test case in Jira.

 Preconditions:

Execution status: **Not Executed**

Final comment:

Detailed steps results:

Step 1	Step Execution Status: **Not Executed**
	Description
	Expected Result
	Actual Result

5.2   Test Cycle LVV-C162

Open test cycle *LDM-503-GEN3-01: Gen 3 Ingest ComCam raw dataset* in Jira.

Test Cycle name: LDM-503-GEN3-01: Gen 3 Ingest ComCam raw dataset
Status: Not Executed

Rubin Observatory

In the context of the milestone LDM-503-GEN3-01, Gen 3 Butler readiness, this test cycle is defining the configuration and the dataset for running a generic **Raw Data Ingestion Into Gen3 Butler** test case, using specifically ComCam raw data.

5.2.1   Software Version/Baseline

Not provided.

5.2.2   Configuration

Not provided.

5.2.3   Test Cases in LVV-C162 Test Cycle

5.2.3.1   LVV-T2007 - Raw Data Ingestion Into Gen3 Butler

Version **1**. Open *LVV-T2007*   test case in Jira.

With this test case we prove that the ingestion of raw data is Gen3 Butler is working as expected.

We are not specifying any raw dataset type or version, nor we are reffing to a specific version of Gen3 Butler software.

The raw dataset type and version shall be specified in the Test Cycle where this test case is included. It is assumed that the ingestion procedure in a Gen3 repository do not depend from the raw dataset type.

The Gen3 Butler version shall be specified in the Test Plan, which defines the test campaign and corresponds to the test milestone to be fulfilled.

Rubin Observatory

Preconditions:

A Gen3 Butler instance has been created and is ready to use with sufficient resources

Execution status: **Not Executed**

Final comment:

Detailed steps results:

Step 1	Step Execution Status: **Not Executed**
	Description

Verify that the repository where to ingest the data is available (create it if it is not)

	Expected Result
	Actual Result

Step 2	Step Execution Status: **Not Executed**
	Description

Ingest the raw dataset following the instruction in the example code

	Expected Result
	Actual Result

Step 3	Step Execution Status: **Not Executed**
	Description

Check that the raw dataset is available in the Gen3 repository

	Expected Result
	Actual Result

A Acronyms used in this document

Acronym	Description
BPS	Batch Production Service
CPP	Calibration Production Processing
ComCam	The commissioning camera is a single-raft, 9-CCD camera that will be installed in LSST during commissioning, before the final camera is ready.
DBB	Data Backbone
DC2	Data Challenge 2 (DESC)
DM	Data Management
DMTR	DM Test Report
LDM	LSST Data Management (Document Handle)
LSST	Legacy Survey of Space and Time (formerly Large Synoptic Survey Telescope)
NCSA	National Center for Supercomputing Applications
OODS	Observatory Operations Data Service
PMCS	Project Management Controls System
VE	vendor estimate

B   Traceability

Test Case	VE Key	VE Summary
LVV-T1982		
LVV-T1983		
LVV-T1984		
LVV-T1985		
LVV-T1986		
LVV-T1987		
LVV-T2007		