Large Synoptic Survey Telescope (LSST)
Measurement of DM Key Performance Metrics

David Nidever, Frossie Economou,
Donald Petravick, K-T Lim

LDM-XXX

Latest Revision: Jun 6, 2016

This LSST document has been approved as a Content-Controlled Document by the LSST DM Technical Control Team. If this document is changed or superseded, the new document will retain the Handle designation shown above.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 LSST DM TCT minutes.
## Change Record

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Description</th>
<th>Owner name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>5/6/2016</td>
<td>Initial version</td>
<td>David Nidever</td>
</tr>
<tr>
<td>1.1</td>
<td>5/17/2016</td>
<td>Further Specification of NCSA tasks.</td>
<td>Donald Petravick</td>
</tr>
<tr>
<td>1.2</td>
<td>6/6/2016</td>
<td>Edits for submission to System Engineering</td>
<td>Frossie Economou</td>
</tr>
</tbody>
</table>
Table of Contents

CHANGE RECORD ........................................................................................................................................ I

1 INTRODUCTION ........................................................................................................................................ 1

2 PHOTOMETRIC METRICS .......................................................................................................................... 1
  2.1 PHOTOMETRIC REPEATABILITY ........................................................................................................ 1
    2.1.1 Calibration Processing Performance ............................................................................................ 1
    2.1.2 Photometric Repeatability Performance ....................................................................................... 2
  2.2 PHOTOMETRIC SPATIAL UNIFORMITY ............................................................................................... 3
  2.3 COLOR ZERO-POINT ACCURACY ......................................................................................................... 4
  2.4 ABSOLUTE PHOTOMETRIC ACCURACY ............................................................................................ 5

3 ASTROMETRIC METRICS .......................................................................................................................... 6
  3.1 RELATIVE ASTROMETRY .................................................................................................................... 6
  3.2 ABSOLUTE ASTROMETRY .................................................................................................................. 7

4 SHAPE METRICS ...................................................................................................................................... 8
  4.1 RESIDUAL PSF ELLIPTICITY CORRECTIONS ....................................................................................... 8

5 TRANSIENT & MOVING OBJECT METRICS .......................................................................................... 9
  5.1 MOVING OBJECT LINKAGE EFFICIENCY ............................................................................................. 9
  5.2 SPURIOUSNESS METRIC EFFICIENCY – TRANSIENTS ................................................................. 10
  5.3 SPURIOUSNESS METRIC EFFICIENCY – MOPS ............................................................................. 11

6 COMPUTATIONAL PERFORMANCE METRICS .................................................................................... 12
  6.1 COMPUTATIONAL BUDGETS .............................................................................................................. 12
  6.2 DATABASE METRICS ......................................................................................................................... 14
  6.3 WEB UI METRICS ............................................................................................................................. 15

7 SYSTEM RELIABILITY METRICS ......................................................................................................... 15
  7.1 SCIENCE VISIT ALERT GENERATION RELIABILITY ....................................................................... 16
  7.2 SCIENCE IMAGE ARCHIVING RELIABILITY ..................................................................................... 16
The Measurement and Verification of DM Key Performance Metrics

1 Introduction

The Key Performance Metrics (KPMs) are the main criteria by which the performance of the Data Management (DM) software system are evaluated quantitatively. Many of these metrics are initially discussed in the Science Requirements Document (SRD; LPM-17) but detailed further in the LSST System Requirements (LSR; LSE-29) and the Observatory System Specifications (OSS; LSE-30). They can be grouped into six main categories: photometric, astrometric, shape, transient, computational and system stability metrics.

It is important to note at the outset that these metrics are meant to test to performance of the software and not the data itself. Therefore, high quality datasets need to be used for these tests otherwise they will only reveal the flaws in the data and not the software.

In the following, RSS stands for Root of Sum of Squares or adding all of the error terms in quadrature.

IT IS CURRENTLY NOT CLEAR WHAT TYPES OF PHOTOMETRIC MAGNITUDES TO USE FOR THE PHOTOMETRIC KPMs. SINGLE-FRAME MEASUREMENT, COADDS, FORCED-PHOTOMETRY, MULTI-FIT, PSF, OR APERTURE?

2 Photometric Metrics

There are several metrics on the precision (i.e. repeatability) and accuracy of the photometry.

2.1 Photometric Repeatability

2.1.1 Calibration Processing Performance
Reference Document: LSE-30; OSS-REQ-0275

Discussion: Specification: The LSST System shall photometrically calibrate raw image data such that the data processing contributes no more than the allocations of procCalAlloc for repeatability. These allocations include effects from calibration algorithms, errors and noise in producing the necessary calibration data products, as well as errors and uncertainties in any reference catalogs used in the calibration process.
The maximum allowed RSS contribution to the overall photometric repeatability of bright isolated point sources caused by errors introduced in the data processing pipelines.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.003</td>
<td>ABmag</td>
<td>procCalRep</td>
<td>DLP-307</td>
</tr>
</tbody>
</table>

**Verification Owner:** SQuaRE

**Verification Technique:** Two visits (exposures) of the same field with the same telescope/filter will be processed through the DM science pipelines including calibration. The sources will then be matched between the two visits and the RMS of the photometric differences of bright stars computed.

**Verification Code:** Software exists to perform this calculation using data processed through the first step of the science pipeline (single frame processing with processCcd) called validate_drp ([https://github.com/lsst/validate_drp](https://github.com/lsst/validate_drp)). However, validate_drp does not include the calibration processing which are covered by this KPM. Therefore, a modified version of this software will need to be written that covers a larger portion of the DM data processing.

**Verification Implementation:** TBD

**Verification Data:** The KPM can be testing using precursor datasets. The recommended dataset is the Hyper Suprime-Cam (HSC) engineering data of the COSMOS field.

### 2.1.2 Photometric Repeatability Performance

**Reference Document:** LSE-29; LSR-REQ-0093

**Discussion:** The specifications for photometric repeatability applies to the cataloged LSST magnitudes, $m^\text{std}(\text{catalog})$ (see SRD eq. 8), for appropriately chosen main sequence stars (e.g. non-variable stars color-selected from the main stellar locus).

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>The RMS photometric repeatability of bright non-saturated unresolved point sources in the $g$, $r$, and $i$ filters.</td>
<td>5</td>
<td>milli-Mag</td>
<td>PA1gri</td>
<td>DLP-315</td>
</tr>
<tr>
<td>The RMS photometric repeatability of bright non-saturated unresolved point sources in the $u$, $z$, and $y$ filters.</td>
<td>7.5</td>
<td>milli-Mag</td>
<td>PA1uzy</td>
<td>DLP-316</td>
</tr>
</tbody>
</table>

**Verification Owner:** SQuaRE
**Verification Technique:** The technique is essentially the same as that used to compute the “procCalRep” KPM.

**Verification Code:** The same software as used for “procCalRep” KPM can be used here.

**Verification Implementation:** TBD

**Verification Data:** The same data as used for “procCalRep” KPM can be used here.

### 2.2 Photometric Spatial Uniformity

Reference Document: LSE-29; LSR-REQ-0093. See also SRD pg. 21.

**Discussion:** The distribution width (rms) of the internal photometric zero-point error (the system stability across the sky) will not exceed PA3 millimag. Use same source selection as for PA1gri.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS width of internal photometric zero-point (precision of system uniformity across the sky) for u-band.</td>
<td>20</td>
<td>milli-Mag</td>
<td>PA3u</td>
<td>DLP-317</td>
</tr>
<tr>
<td>RMS width of internal photometric zero-point (precision of system uniformity across the sky) in the g-band.</td>
<td>10</td>
<td>milli-Mag</td>
<td>PA3g</td>
<td>DLP-318</td>
</tr>
<tr>
<td>RMS width of internal photometric zero-point (precision of system uniformity across the sky) in the y-band.</td>
<td>10</td>
<td>milli-Mag</td>
<td>PA3y</td>
<td>DLP-319</td>
</tr>
</tbody>
</table>

**Verification Owner:** SQuaRE

**Verification Technique:** Checking the spatial uniformity of ground-based photometric zero-points is challenging because there is no hyper-accurate photometric dataset that covers the entire southern sky (at least to the level that LSST needs). Therefore, the recommendation is to use space-based data from the Gaia mission that should be accurate across the sky to the few milli-mag level. Since Gaia uses a different photometric system from LSST, photometric transformation equations from Gaia’s G-band to the LSST bands will need to be derived (one set per filter). The Gaia G-band photometry, properly transformed, will then be compared to ground-based photometry in various parts of the sky. To improve the comparison only bright, blue, non-variable main-sequence stars will be used. The median difference (zero-point error) in each patch of sky will be computed and then the RMS computed over many patches of sky. Note that if the Gaia dataset are used to calibrate the LSST or pre-cursor data, then a separate test dataset will need to be used such as the HST white dwarf standard stars.

**Verification Code:** Software will need to be written to ingest the Gaia data and to determine and apply the transformation equations from the ground-base photometric system to the Gaia G-band. Next,
software will be needed to select the necessary sample of bright, blue, non-variable main-sequence stars in various regions across the sky, compare the photometry to the Gaia and calculate the final metric of the RMS of the differences in each band.

**Verification Implementation:** TBD

**Verification Data:** The Gaia G-band dataset (available in 2017) or the HST white dwarf standard star data will be used as the reference. The test data should be comprised of many fields spread across the sky using a modern multi-chip imager (e.g. HSC, DECam, CFHT). A subset of the HSC or DES survey data could be used for this purpose.

### 2.3 Color Zero-point Accuracy

**Reference Document:** LSE-29; LSR-REQ-0093, also see SRD pg. 22

**Discussion:** The absolute band-to-band zero-point transformations (color zero-points, e.g. for constructing the spectral energy distribution, SED) for main-sequence stars must be known with an accuracy of PA5 millimag. These requirements are primarily driven by the desired accuracy of photometric redshift estimates. Note that an overall stable gray error in the absolute calibration of the system does not have an impact on the above requirements. Such an error is specified by PA6. The same source selection as for PA1gri should be used.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy of absolute band-to-band color zero-point for all colors constructed from any filter pair, excluding the u-band.</td>
<td>5</td>
<td>milli-Mag</td>
<td>PAS</td>
<td>DLP-320</td>
</tr>
<tr>
<td>Accuracy of absolute band-to-band color zero-point for colors constructed using the u-band.</td>
<td>10</td>
<td>milli-Mag</td>
<td>PA5u</td>
<td>DLP-321</td>
</tr>
</tbody>
</table>

**Verification Owner:** SQuaRE

**Verification Technique:** This requirement is on the color of stars. It can be tested by using spectrophotometric standard stars which have well-determined space-based spectra. These spectra will need to be convolved with the ground-based filter throughput curves to produce the synthetic magnitudes and colors. The synthetic colors of a number of standard stars will be compared to the calibrated ground-based colors and the median difference measured.

**Verification Code:** The calibrated ground-based colors for the calibrators will be compared to the synthetic colors of the standard stars. Simple software will be needed to “convolve” the spectrophotometric spectrum with the filter response curves to create the synthetic photometry.
Verification Implementation: TBD

Verification Data: The HST standard star flux calibrators can be used as the reference data (http://www.stsci.edu/hst/observatory/crds/calspec.html). The test data should be comprised of many fields spread across the sky (in the regions where the calibrators are located) using a modern multi-chip imager (e.g. SDSS, HSC, DECam, CFHT).

2.4 Absolute Photometric Accuracy

Reference Document: LSE-29; LSR-REQ-0093. See also SRD, pg. 22.

Discussion: The LSST photometric system must transform to a physical scale (e.g. AB magnitude scale) with an accuracy of PA6 millimag. The requirements are driven by the accuracy of absolute determination of quantities such as luminosity and asteroid size for objects with well determined distances. Note that the internal band-to-band transformations are required to be much more accurate as they may be calibrated and controlled by other means, and are not sensitive to errors in overall flux scale of photometric calibrators. Use same source selection as for PA1gri.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy of the transformation of the internal LSST</td>
<td>10</td>
<td>milli-Mag</td>
<td>PA6</td>
<td>DLP-322</td>
</tr>
<tr>
<td>photometry to a physical scale (e.g. AB magnitudes).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Verification Owner: SQuaRE

Verification Technique: To accurately perform absolute flux calibration of ground-based photometry is quite challenging. Normally, spectrophotometric standard stars, such as white dwarfs, are used for this purpose. There is currently an effort underway, led by A. Saha and C. Stubbs, to calibrate many hot DA white dwarfs (to the few millimag level using HST photometry, Gemini spectroscopy and precise spectral models) for the purpose of absolute flux calibration for surveys like LSST. These standards will likely set the absolute flux calibration of the LSST photometric system after an ubercal calibration has been applied to the data. To test the accuracy of this calibration one would want to use a subset of the calibrators to perform the calibration and use the rest as the reference to test against.

Verification Code: The software needed is fairly straightforward. The ground-based photometry for the calibrators not used in the calibration will be compared to the flux calibrated spectrum of the star “convolved” with the filter response function (which itself must be quite accurate).

Verification Implementation: TBD

Verification Data: Until the data for the new DA white dwarf calibrators become available the existing HST standard star flux calibrators can be used as the reference.
The contents of this document are subject to configuration control by the LSST DM Technical Control Team.

(http://www.stsci.edu/hst/observatory/crds/calspec.html). The test data should be comprised of many fields spread across the sky (in the regions where the calibrators are located) using a modern multi-chip imager (e.g. SDSS, HSC, DECam, CFHT). As this requirement is also a constraint on the relative ubercalibration across the sky, it would be useful to use a large dataset like SDSS, Pan-STARRS1, DES or the HSC survey to assess this metric.

### 3 Astrometric Metrics

There are relative and absolute requirements on the astrometry produced by the LSST DM stack.

#### 3.1 Relative Astrometry


**Discussion:** This KPM concerns the astrometric quality of single visit exposures. The rms of the astrometric distribution for stellar pairs with separation of D arcmin (repeatability) will not exceed AMx milliarcsec (median distribution for a large number of sources). The three selected characteristic distances reflect the size of an individual sensor, a raft, and the camera. The required median astrometric precision is driven by the desire to achieve a proper motion accuracy of 0.2 mas/yr and parallax accuracy of 1.0 mas over the course of the survey. These two requirements correspond to relative astrometric precision for a single image of 10 mas (per coordinate).

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median relative astrometric measurement error on 5 arcminute scales.</td>
<td>&lt;10</td>
<td>milli-Arcsec</td>
<td>AM1</td>
<td>DLP-310</td>
</tr>
<tr>
<td>Median relative astrometric measurement error on 20 arcminute scales.</td>
<td>&lt;10</td>
<td>milli-Arcsec</td>
<td>AM2</td>
<td>DLP-311</td>
</tr>
<tr>
<td>Median relative astrometric measurement error on 200 arcminute scales.</td>
<td>&lt;15</td>
<td>milli-Arcsec</td>
<td>AM3</td>
<td>DLP-312</td>
</tr>
<tr>
<td>RMS difference between separations measured in the r-band and those measured in any other filter.</td>
<td>10</td>
<td>milli-Mag</td>
<td>AB1</td>
<td>DLP-313</td>
</tr>
</tbody>
</table>

**Verification Owner:** SQuaRE

**Verification Technique:** LSST needs to have good relative astrometry not just locally but over various scales on the sky. This can be checked by selecting pairs of stars, measuring their distances many times (over many visits), and determining how much the angular distance varies around its mean value. This exercise is repeated for pairs of stars separated on different angular scales. For AB1, the angular distances of pairs of stars are compared between one band and the r-band and the RMS of the differences determined.
**Verification Code:** Software will be needed to match sources across different visits. Second, stars should be selected at random across the field and their “partner” star with a separation close to the angular scale that is being tested (5, 20, 200 arcmin). Then, the angular distance for each pair should be measured for each visit and the RMS determined about the mean. Finally, a median is taken of the distribution of RMS values. For AB1, the angular distance is compared between one band and the r-band and the RMS of the differences computed (only one visit is needed for each band for this test).

**Verification Implementation:** TBD

**Verification Data:** The test data should be comprised of many visits in a small region of the sky using a modern multi-chip imager (e.g. HSC, DECam, CFHT). The HSC engineering data on the COSMOS field, or a subset of the HSC or DES survey data could be used for this purpose.

### 3.2 Absolute Astrometry

**Reference Document:** LSE-29; LSR-REQ-0094. See also SRD, pg.24.

**Discussion:** The LSST astrometric system must transform to an external system (e.g. ICRF extension) with the median accuracy of AA1 milliarcsec. The accuracy of absolute astrometry is driven by the linkage and orbital computations for solar system objects. A somewhat weaker constraint is also placed by the need for positional cross-correlation with external catalogs. Note that the delivered absolute astrometric accuracy may be fundamentally limited by the accuracy of astrometric standard catalogs.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median error in absolute position for each axis, RA &amp; DEC, shall be less than AA1.</td>
<td>50</td>
<td>milli-Arcsec</td>
<td>AA1</td>
<td>DLP-309</td>
</tr>
</tbody>
</table>

**Verification Owner:** SQuaRE

**Verification Technique:** The astrometric positions of stable sources (e.g. bright quasars) will be compared between astrometrically-calibrated LSST precursor datasets and an astrometrically accurate catalog such as Gaia.

**Verification Code:** The software will select stable sources from a catalog of quasars. The positions will be compared to a standard catalog, such as Gaia, and the median difference of the positions computed.

**Verification Implementation:** TBD

**Verification Data:** The test data should be cover a large portion of the sky to cover enough quasars. The SDSS, DES or HSC survey datasets could be used for this purpose.
4 Shape Metrics
The measurements of shapes of galaxies are important for weak lensing and constraining dark energy.

4.1 Residual PSF Ellipticity Corrections
Reference Document: LSE-29; LSR-REQ-0097. See also SRD, pg. 30-32.

Discussion: Using the full survey data, the E1 and E2 (see SRD for definitions) distributions averaged over an arbitrary FOV shall have medians less than TE1 for theta ~ 1 arcmin, and less than TE3 for theta < 5 arcmin. No more than TEF % of images shall have these medians for E1 and E2 larger than TE2 for theta ~ 1 arcmin, or larger than TE4 for theta < 5 arcmin.

The requirements specified here require the full survey data set to exist before they can be met. Thus these are intended to ensure that the LSST system design enables that these requirements can be met after the 10-year survey. Prior to survey start, they will be verified to the extent possible using simulations incorporating the as-built telescope and camera performance characteristics.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median residual PSF ellipticity correlations averaged over an arbitrary field of view for separations less than 1 arcmin shall be no greater than</td>
<td>2.0e-5</td>
<td>--</td>
<td>TE1</td>
<td>DLP-290</td>
</tr>
<tr>
<td>Median residual PSF ellipticity correlations averaged over an arbitrary field of view for separations between 1 and 5 arcmin shall be no greater than</td>
<td>1.0e-7</td>
<td>--</td>
<td>TE2</td>
<td>DLP-308</td>
</tr>
</tbody>
</table>

Verification Owner: SQuaRE

Verification Technique: The mean correlation of PSF ellipticity residuals needs to be measured between pairs of bright stars. The ellipticity is measured for bright stars and the low-order dependence of the ellipticity on the position in the focal plane is removed (creating ellipticity residuals). The auto and cross-correlation functions of the ellipticity residuals averaging over pairs of stars at a given angular separation are then computed. An important component of the ellipticities are the contributions from the optics. Therefore, this needs to be tested on simulated LSST data.

Verification Code: Elliptical PSF models need to be fit to the bright stars and PCA fit performed on the spatial dependence of the ellipticity values to produce the ellipticity residuals. Pairs of stars on certain angular scales should be selected and the auto and cross-correlation functions of the ellipticity residuals measured.

Verification Implementation: TBD
**Verification Data:** Simulated LSST data of a small patch of sky will be required to test this KPM. The Twinkles simulation data might be sufficient for this test.

## 5 Transient & Moving Object Metrics

The detection and characterization of transients, including solar system objects, is an important component of LSST.

### 5.1 Moving Object Linkage Efficiency

**Reference Document:** LSE-30; OSS-REQ-0159

**Discussion:** Valid identification and orbits shall be determined for at least a fraction of Solar System objects which are detected orbit observations (2) times in orbit observation interval (3) days at a level of orbit observation threshold (5) sigma or more above the single frame background.

Valid identification means that detections of the same Solar System objects have been correctly associated as such. The verification method for this requirement will be comparison with simulated inputs.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum fraction of Solar System objects meeting reference criteria for which valid orbits shall be determined.</td>
<td>95%</td>
<td>Percent</td>
<td>orbitCompleteness</td>
<td>DLP-323</td>
</tr>
</tbody>
</table>

**Verification Owner:** SQuaRE, with help from University of Washington initially

**Verification Technique:** The LSST reduction and MOPS software will be run on simulated images that include solar system objects and the completeness of the recovered objects measured.

**Verification Code:** The simulated data for several nights will need to be processed through the Level 1, image differencing, pipeline to detect the transients. These detections will then be fed to the MOPS software to perform the linkages. Separate software will then compare these linkages to the actual simulated orbits.

**Verification Implementation:** TBD

**Verification Data:** Simulated LSST images with solar system objects for several nights are required to perform this test. These do not exist yet and should be requested from the simulation team. A subset of the data needed to test the MOPS spuriousness metric can be used here.
5.2 Spuriousness Metric Efficiency – Transients

Reference Document: LSE-30; OSS-REQ-0353

Discussion: There shall exist a spuriousness threshold $T$ for which the completeness and purity of selected difference sources are higher than transCompletenessMin and transPurityMin, respectively, at the SNR detection threshold transSampleSNR (6). This requirement is to be interpreted as an average over the entire survey.

This specification captures representative completeness and purity rates supportive of time-domain science cases. Note that these are rates determined only using the spuriousness metric cut; it is likely the end-users will perform further classification steps to increase the purity of their samples, depending on their particular science case.

This specification will be tested using simulations, by insertion and recovery of artificial sources, and comparisons to ground truth where known (i.e., asteroids, known variable stars, known variable quasars, etc).

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum average completeness for transient science.</td>
<td>90</td>
<td>Percent</td>
<td>transCompletenessMin</td>
<td>DLP-324</td>
</tr>
<tr>
<td>Minimum average purity for transient science.</td>
<td>95</td>
<td>Percent</td>
<td>transPurityMin</td>
<td>DLP-325</td>
</tr>
</tbody>
</table>

Verification Owner: SQuaRE, with help from University of Washington initially

Verification Technique: The LSST reduction and image differencing software will be run on simulated images with transients (including solar system objects) and the completeness and purity of the recovered transients measured.

Verification Code: The simulated data for several nights will need to be processed through the Level 1, image differencing, pipeline to detect the transients. These detections will then be compared to the input transients to ascertain the completeness and purity.

Verification Implementation: TBD

Verification Data: Simulated data are needed to test this requirement. As this metric is meant to be an average over the entire survey, the simulated data should be for a representative patch of sky (or many of them) and include visits for the entire survey with realistic weather and must include the various types of transients. The Twinkles data might be approaching what is needed for this, but it is likely that more types of transients are needed and multiple fields on the sky.
5.3 Spuriousness Metric Efficiency - MOPS

Reference Document: LSE-30; OSS-REQ-0354

**Discussion**: There shall exist a spuriousness threshold $T$ for which the completeness and purity of difference sources are higher than $mops$CompletenessMin and $mops$PurityMin, respectively, at the SNR detection threshold $orbitObservationThreshold$ (5). This requirement is intended to be interpreted as an average for any one month of observing.

This specification captures representative completeness and purity rates needed to enable successful identification and linking of observed Solar System objects. In particular, the need to have a Solar System object repeatedly detected $orbitObservation$ (2) times in $orbitObservationInterval$ (3) days strongly prefers high completeness, even at the expense of purity.

This specification will be tested using simulations, by insertion and recovery of artificial sources, and comparisons to ground truth where known (i.e., asteroids, known variable stars, known variable quasars, etc).

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum average completeness for Solar System object discovery</td>
<td>99</td>
<td>Percent</td>
<td>$mops$CompletenessMin</td>
<td>DLP-326</td>
</tr>
<tr>
<td>Minimum average purity for Solar System object discovery</td>
<td>50</td>
<td>Percent</td>
<td>$mops$PurityMin</td>
<td>DLP-327</td>
</tr>
</tbody>
</table>

**Verification Owner**: SQuaRE, with help from University of Washington initially

**Verification Technique**: The verification of this metric proceeds very similarly to the Moving Object Linkage Efficiency metric. The LSST reduction and MOPS software will be run on simulated images that include solar system objects and the completeness and purity of the recovered objects over month-long time spans are measured.

**Verification Code**: The simulated data for many months (maybe ~1 year) will need to be processed through the Level 1, image differencing, pipeline to detect the transients. These detections will then be fed to the MOPS software to perform the linkages and compute orbits. Separate software will then compare the solar system objects to those injected to determine the completeness and purity.

**Verification Implementation**: TBD
Verification Data: Simulated full focal-plane LSST images including solar system objects for many months (maybe ~1 year) are required to perform this test. These do not exist yet and should be requested from the simulation team.

6  Computational Performance Metrics

There are various requirements on the computational performance of the DM stack and hardware.

6.1  Computational Budgets

Reference Document: LDM-138|Output|C4, F4 and LSE-81|G263. Also see SRD (pg.33-34) and LDM-140.

Discussion: Data on likely optical transients will be released with a latency of at most OTT1 (1) seconds. This includes the time to transfer the data from Chile to the US, the computational time to process the data and any overheads (e.g. I/O). The AP computational budget is the amount of computational processing power the software needs to meet the OTT1 with a reasonable amount of overhead. In a similar vein, the DRP computational budget is the amount of computational processing power the software needs to process the data in 1 year (6 months for the first year) with reasonable overheads.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>The minimum latency for releasing data on optical transients. Performance of all AP pipelines, middleware and infrastructure.</td>
<td>60</td>
<td>seconds</td>
<td>OTT1</td>
<td>DLP-328</td>
</tr>
<tr>
<td>Alert Production computational budget. Performance of all AP pipelines and middleware.</td>
<td>39</td>
<td>TFLOPS</td>
<td>AP computational budget</td>
<td>DLP-329</td>
</tr>
<tr>
<td>Data Release Production computational budget (DR1). Performance of all DRP+AP pipelines.</td>
<td>108</td>
<td>TFLOPS</td>
<td>DRP computational budget (DR1)</td>
<td>DLP-314</td>
</tr>
</tbody>
</table>

Verification Owner: NCSA

Verification Technique:

Computational budget: The AP and DRP computational budgets are straightforward to measure, due to the specification being floating point operations. A representative sample of data, further described in the verification data section, will be processed through the specified pipelines and the total number of floating point operations counted. For all pipelines, floating point operations need only be counted within the scientific codes. This is because the orchestration and other support systems used to support
pipeline code execution in the production system is not floating-point intensive. For data release processing, this number is then processed to obtain the number of teraflops needed to produce the first data release, which is 6 months of data processed in 4.5 months. For Alert processing this number indicate that sufficient floating point operations are available to meet the computational duration budget of 47 seconds. The Alert Production computational budget of 39 Tflops provides for two "strings" of computing. For Alert processing, the test is to show that 39/2 = 19.5 Tflops can process standard visits as indicated in the Verification data section, below, in 47 seconds.

The OTT1 test measures the time from digitization of the last pixel of the last exposure in a standard 60 second visit until the time of the presentation of the last VO event resulting from that visit to an event broker, or 10,000th event, which ever comes first. The test includes all data handling as well as computation time, including allowances for the bandwidth and latency in the networks between the camera system and the NCSA.

**Verification Code:** For the AP and DRP codes the AP and DRP will be pipelines as procured by the LSST configuration management system, compiled as for production use.

For the OTT1 tests, all software and system elements from digitization of the last pixel of the last visit until the last alert from that visit is presented to an event broker’s “author” component. Here “author” means the software (or network protocol) used to communicate with a broker, and where that software imposes no constraint on the OTT1 measurement. Simulations/emulations of the camera system and OCS will be used.

**Verification Implementation:**

Verification of the number of floating point operations for AP computational budget and Data Release Production computational budget will occur on computers that are representative of the floating point architecture and compiler architecture that will be procured for the production system. Floating point operations can be acquired from hardware registers. Floating point operations will be the sum of all floating point operations of any width, e.g. 32-bit floating point operations, 64-bit floating point operations.

Test OTT1 can be conducted at NCSA, as the project possesses a network simulator that can simulate the network between the base center and NCSA. The test presumes the supplied OCS simulator and CCS simulators supplied to NCSA are performant, or that a calibration compensating for any performance differences can be applied. Calibrations for the number of floating point operations available in test computer compared to computers that will be purchased for the beginning of operations will be applied, if needed.

**Verification Data:** Simulated full-frame LSST data will be needed for both the AP and DRP processing. For AP, the test data will be drawn from standard visits of crowded fields in the most sensitive filters.
For DRP enough visits to realistically model the expected distribution of

- Crowded and sparse fields,
- The distribution of exposures across filters
- Including only standard visits.
- Corresponding supporting data, including calibration data.

Coadds for DR1 will be needed for a small number of fields.

### 6.2 Database Metrics

**Reference Document:** LSE-81|G270, G283, G284. Also see LDM-135.

**Discussion:** These are metrics on the performance of the Qserv database service. See LDM-135 section 4.3 for more information on “low volume” and “high volume” queries.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qserv query rate</td>
<td>150</td>
<td>Simult. queries</td>
<td>query rate</td>
<td>DLP-659</td>
</tr>
<tr>
<td>Qserv low volume query response time (Type A-D)</td>
<td>10</td>
<td>seconds</td>
<td>LV query response time</td>
<td>DLP-651 – DLP-654</td>
</tr>
<tr>
<td>Qserv low volume Object/Source joins response time</td>
<td>12</td>
<td>hours</td>
<td>LV Object/Source joins response time</td>
<td>DLP-657</td>
</tr>
<tr>
<td>Qserv high volume object scan response time</td>
<td>1</td>
<td>hours</td>
<td>HV object scan response time</td>
<td>DLP-655</td>
</tr>
<tr>
<td>Qserv high volume ObjectExtra scan response time</td>
<td>8</td>
<td>hours</td>
<td>HV ObjectExtra scan response time</td>
<td>DLP-656</td>
</tr>
<tr>
<td>Qserv high volume Object/ForcedSource join response time</td>
<td>12</td>
<td>hours</td>
<td>HV Object/ForcedSource join response time</td>
<td>DLP-658</td>
</tr>
</tbody>
</table>

**Verification Owner:** NCSA

**Verification Technique:** The Qserv database will be loaded with a large precursor dataset. Sample queries will be run at a high rate and the response time measured. The results will need to be scaled up to the size of the eventual LSST database.

**Verification Code:** Software will need to written to auto-generate a large number of queries of various types, submit the queries to Qserv, receive the returned data and measure the response and delay time.
Verification Implementation: TBD

Verification Data: For this test a precursor dataset is needed that is large enough to allow the results to be appropriately scaled up to the LSST database size. The Pan-STARRS1, SDSS or DES survey datasets should be good for this purpose.

6.3 Web UI Metrics

Reference Document: JIRA tickets.

Discussion: These are requirements on the speed of the web user interface (UI).

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web UI portal can support concurrent users.</td>
<td>100</td>
<td>users</td>
<td>Web UI portal concurrent user supported</td>
<td>DLP-330</td>
</tr>
<tr>
<td>Web UI query results display response time after the search results have been received from the data provider.</td>
<td>1</td>
<td>seconds</td>
<td>Web UI query response time</td>
<td>DLP-331</td>
</tr>
<tr>
<td>Web UI image preparation (convert to PNG or other image format) time on the server side.</td>
<td>0.5</td>
<td>seconds</td>
<td>Web UI image preparation time</td>
<td>DLP-332</td>
</tr>
<tr>
<td>Web UI image rendering time on user’s display.</td>
<td>0.5</td>
<td>seconds</td>
<td>Web UI image display rendering time</td>
<td>DLP-333</td>
</tr>
</tbody>
</table>

Verification Owner: SQuaRE

Verification Technique: Precursor datasets will be loaded into a database and queried via the web UI and the response and rendering times measured (on the server or user side as appropriate). Software will need to be used to emulate many users for the first metric.

Verification Code: The server side and user side software will need to be able to measure the rendering and response times to test these metrics. Software to emulate ~100 users, including simulated queries and the capture of the response, will be needed used or developed for this test.

Verification Implementation: TBD

Verification Data: Precursor datasets such as WISE or Pan-STARRS1 can be used for this test.

7 System Reliability Metrics

There are additional requirements on the reliability of the entire DM system.
7.1 Science Visit Alert Generation Reliability
Reference Document: LSE-30; OSS-REQ-0112

**Discussion**: No more than sciVisitAlertFailure % of science visits read out in the camera [and specified to be analyzed by Data Management] shall fail to be subjected to alert generation and distribution, integrated over all stages of data handling from data acquisition through transmission of the alerts across the project boundary. No more than sciVisitAlertDelay % of science visits read out in the camera [and specified to be analyzed by Data Management] shall have their alert generation and distribution completed later than the SRD specification for alert latency (OTT1).

The "specified to be analyzed" language allows for the possibility that under some operating modes, such as diagnostics, images deliberately not be analyzed for alerts.

This requirement applies to visits, and not to individual alerts, because a specification that, e.g., "no more than 1% of alerts shall fail to be generated" gets tangled with questions of the scientific performance of the actual alert detection. This requirement is a performance specification on the DM system, taking the alert-detection algorithm as a given.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum fraction of science visits for which alerts are generated, but delivered later than the OTT1 latency specification.</td>
<td>1</td>
<td>Percent</td>
<td>sciVisitAlertDelay</td>
<td>--</td>
</tr>
<tr>
<td>Maximum fraction of visits for which alerts are not generated or distributed.</td>
<td>0.1</td>
<td>Percent</td>
<td>sciVisitAlertFailure</td>
<td>--</td>
</tr>
</tbody>
</table>

**Verification Owner**: SQuaRE, with help from University of Washington initially

**Verification Technique**: This requirement will need to be tested with a “full-up” test of the DM system. This might not be possible until commissioning with the full camera.

**Verification Code**: This metric can be evaluated by using the internal database of when visits were taken and their alerts generated.

**Verification Implementation**: TBD

**Verification Data**: This metric will need to be tested with the “full-up” DM system, likely during commissioning with the full camera.

7.2 Science Image Archiving Reliability
Reference Document: LSE-30; OSS-REQ-0111
**Discussion:** No more than sciImageLoss % of science, wavefront, and guider images read out in the camera and specified to be acquired by Data Management shall be permanently lost or corrupted, integrated over all stages of data handling through archiving and availability of the image for access in the archive, or have its essential image acquisition metadata permanently lost or disassociated with the image.

Corruption of images shall mean changes of any image pixel values, with the exception of application of a compression algorithm. Any such compression algorithm must be shown to have no effect on any SRD requirement.

The "specified to be acquired" language allows for the possibility that under some operating modes, such as diagnostics, images might be acquired by the camera and deliberately not archived.

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
<th>Name</th>
<th>JIRA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum fraction of read-out raw images permitted to be permanently lost or</td>
<td>1</td>
<td>Percent</td>
<td>sciImageLoss</td>
<td>--</td>
</tr>
<tr>
<td>corrupted downstream, including loss due to the loss or corruption of essential associated metadata.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Verification Owner:** NCSA

**Verification Technique:** This requirement will need to be tested with a “full-up” test of the DM system. This might not be possible until commissioning with the full camera.

**Verification Code:** This metric can be evaluated by using the metadata database, the omission of any exposure numbers in the database, and the quality assurance metrics on the corruption of data.

**Verification Implementation:** TBD

**Verification Data:** This metric will need to be tested with the “full-up” DM system, likely during commissioning with the full camera.