LARGE SYNOPTIC SURVEY TELESCOPE

**Large Synoptic Survey Telescope (LSST)
Data Management**

LSST Alerts: Key Numbers

M. L. Graham, E. Bellm, L. Guy, et al.

DMTN-102

Latest Revision: 2018-12-06

Abstract

Abstract text.

Change Record

Version	Date	Description	Owner name
1	2018-12-06	Unreleased.	Melissa Graham

Document source location: `https://github.com/lsst-dm/dmtn-102`

Contents

LSST Alerts: Key Numbers

1 Introduction

This document is currently in development and contains a lot of detailed information for the purposes of internal discussion; to be reduced/clarified later. Text in red needs some discussion by DM-SST.

The LSST Data Management System's (DMS) Alert Production (AP) pipeline will process new data as it is obtained by the telescope. Difference Imaging Analysis (DIA) will be performed, and all sources with a signal-to-noise ratio > 5 in positive or negative flux will be considered "detected", recorded in the source catalogs, and generate an alert. Each alert is a packet containing LSST data about the source such as coordinates, photometry, and image cutouts. For a full description of detected sources and alert packet contents, see LSE-163. The LSST alert stream will be delivered to several community-developed brokers, and also accessible to users via the LSST Science Platform's mini-broker; plans and policies for alert distribution are provided in LDM-612.

The purpose of this document is to quantitatively inform broker developers, and the broader scientific community planning to use alerts, on the key numbers regarding alert generation, distribution, and access via the LSST Mini-Broker. The goals of this document are threefold: (1) to provide all of the key numbers regarding alert generation in one place; (2) to include any and all basis information, assumptions, and derivations that contributed to the key number; and (3) to be clear about whether each key number represents an estimate, a requirement, or a boundary.

In this work we use 8 bits per byte (B), and $1024\,\mathrm{B}$ per KB, $1024\,\mathrm{KB}$ per MB, and so forth.

A list of resources being used in the preparation of this document.

- *LSST: From Science Drivers to Reference Design and Anticipated Data Products*, Ivezic et al. (2008)

- LSST Science Requirements Document (SRD), LPM-17.

- LSST System Requirements (LSR), LSE-29.

- Observatory System Specifications (OSS) document, LSE-30.

- Data Management System Requirements (DMSR) document, LSE-61.

- Science Requirements and System Specifications Spreadsheet (SR&SSS), LSE-81.

- Data Products Definitions Document (DPDD), LSE-163

- Plans and Policies for LSST Alert Distribution, LDM-612

- Data Management Science Pipelines Design, LDM-151
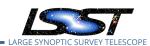
2   Alert Stream

The concept and existence of the LSST alert stream is baked into the highest-level document: the SRD specifies that information about the detections of transient, variable, and moving objects be released as a data stream within 1 minute.

2.1   Alert Production Timescale

It is a formal requirement that the Data Management System distribute alerts to community brokers within 60 seconds of the end of image readout.

Formal Requirements – Regarding the DMS's ability to generate alerts, the SRD states that *"data on likely optical transients … will be released with a latency of at most* $0TT1$ *minutes"*, where $0TT1$ is $1/2/0.5$ minutes (design/minimum specifications and stretch goal; LPM-17). The design specification value of $0TT1$ flows down to a formal requirement in the LSR that specifies *"LSST shall meet the following specification for reporting of data on optical transients detected in single-visit data [in]* $0TT1$*"*, and defines $0TT1$ as *"the latency of reporting optical transients following the completion of readout of the last image of a visit"* (LSR-REQ-0101, LSE-29). The LSR also makes a formal requirement that the DMS's *"algorithm shall be applied and the alert transmitted within the specified latency for at least a fraction* $0TR1$ *of instances … [and] remaining transients … be identified and recorded at the next processing opportunity"*, where $0TR1$ is 98% (LSR-REQ-0025, LSE-29; see § 2.4). The LSR requirement also flows down to more specific formal requirement in the OSS and DMSR, which both state that *"Alerts shall be made available within time* $0TT1$ *from the conclusion of readout of the raw exposures used to generate each alert to the distribution of the alert to community distribution mechanisms"* (OSS-REQ-0127, LSE-30; DMS-REQ-0004, LSE-61).

Tension in the Formal Requirements Definitions – The SRD states that OTT1 is a requirement on the *releasing* of an alert; the LSR states that OTT1 is a requirement on the *reporting* of an alert (LSR-REQ-0101), and then later that it is the time for *transmitting* the alert to the broker (LSR-REQ-0025). It is only the OSS and DMSR that specifies that OTT1 ends at the time of distribution *to the broker*. Note that we often colloquially say *OTT1* starts *at the time of shutter close*, but it actually starts 2 seconds later *at the end of readout* (or > 2 seconds if slower reads are adopted).

Requirement Breach Protocol – For the OTR1 $= 2\%$ of alerts allowed to be transmitted with $>$ OTT1 seconds, how will they be identified as delayed? I.e., with a flag in the alert packet?
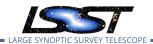
2.2   Number of Alerts per Visit

It is a formal requirement that the Data Management System can sustain the generation of an average of at least 10000 alerts per visit.

Formal Requirements – Regarding the DMS's ability to generate alerts, the SRD states that *"The system should be capable of reporting such data for at least* transN *candidate transients per field of view and visit,"* where transN is $10^4/10^3/10^5$ (design/minimum specifications and stretch goal; LPM-17). The design specification value of transN flows down to a formal requirement in the LSR, which describes transN as *"the minimum number of optical transients for which data can be reported per visit"*, with a note that *"it is unclear whether the SRD specification of* transN *refers to the number of alerts that can be generated for a single visit (i.e. an instantaneous limit), or the number per visit averaged over time"* (LSR-REQ-0101, LSE-29). The OSS does clarify: *"The LSST Data Management system shall be sized to accommodate an average value of at least* nAlertVisitAvg *alerts generated per standard visit while meeting all its other requirements,"* where nAlertVisitAvg is 10^4 (OSS-REQ-0193 in LSE-30). No corresponding requirement is flowed down to the DMSR, but the SR&SSS also uses the same minimum average value, and furthermore uses 40000 as the *peak* number of alerts per visit, with a note that it was chosen as a number in between the SRD's minimum of 10^4 and the stretch goal of 10^5 [LSE-81].

Tension in the Formal Requirements Definitions – OTT1 appears in the DMSR, but neither transN nor nAlertVisitAvg do; is this a problem?

Requirement Breach Protocol – Is there a new upper limit on the peak volume? In cases where

there are > 40000 alerts generated by a visit, does DM expect they issued with a delay?

2.3   Fraction of Visits with Delayed/Failed Alert Distribution

It is a formal requirement that $< 0.1\%$ of all science visits experience a total failure in alert generation and distribution. It is a formal requirement that $< 1\%$ of all science visits have any fraction of their alerts experience a distribution delay > 60 seconds.

Formal Requirements – The SRD and LSR do not say anything on this topic. The OSS specifies that *"no more than* `sciVisitAlertFailure` *% of science visits … shall fail to be subjected to alert generation and distribution"*, where `sciVisitAlertFailure` $= 0.1\%$, and that *"no more than* `sciVisitAlertDelay` *% of science visits … shall have their alert generation and distribution completed later than [*`OTT1`*]"*, where `sciVisitAlertDelay` $= 1\%$ (OSS-REQ-0112; LSE-30). The OSS furthermore makes the distinction that even if only *some* of the alerts for a given visit are distributed later than `OTT1`, it counts towards *sciVisitAlertDelay*. The DMSR makes no statements about the fraction of visits with failed/delayed alert distribution.

Tension in the Formal Requirements Definitions – The OSS requirements do not use `OTR1` to define a failed/delayed visit (see § 2.4); instead, if a visit has any number of alerts delayed beyond `OTT1` it is considered as counting towards `sciVisitAlertDelay`. Should this be changed?
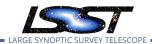
Requirement Breach Protocol – Are alerts from a delayed visit flagged in some way?

2.4   Fraction of Alerts per Visit with Delayed Distribution

It is a formal requirement that 98% of all alerts for a given visit are issued within 60 seconds.

Formal Requirements – The SRD does not say anything on this topic. As mentioned in § 2.1, the LSR defines `OTR1` as the *"fraction of detectable alerts for which an alert is actually transmitted within latency* `OTT1`*"*, where `OTR1` $= 98\%$ (LSR-REQ-0025; LSE-29). The OSS does not state any requirements on the fraction of failed alerts per visit (but see § 2.3). The DMSR makes no statements about the fraction of alerts per visit with delayed/failed distribution.

Tension in the Formal Requirements Definitions – LSR-REQ-0025 and `OTR1` do not appear to have flowed down to the OSS or DMSR, and `OTR1` does not seem to actually be used for any-

thing (unless it's flowed down to science verification?). Also, OTR1 specifies the fraction of alerts that can be delayed beyond OTT1, but there is no specification for an acceptable fraction of alerts which completely fail to be distributed (i.e., like sciVisitAlertFailure, but for alerts per visit); should there be? Is it 0?
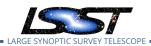
Requirement Breach Protocol – With regards to OTR1 = 98%, LSR-REQ-0025 states that *" the remaining transients so detectable must still be identified and recorded at the next processing opportunity"*, but this is not this flowed down to DMSR, and it is unclear what *"the next processing opportunity"* means.

2.5   Fraction of False Positives per Visit

There is no formal requirement on the fraction of false positives per visit, only that threshold values be supplied, which can be used to filter the detections into a subsample with known upper limits of completeness and purity.

Formal Requirements – The SRD makes no statements about alert stream purity or completeness, but does quote that the *"minimum signal-to-noise ratio in difference image for reporting detection of a transient object"* has a design specification of transSNR = 5 [LPM-17]. The LSR contains essentially the same definition for transSNR, *"the signal-to-noise ratio in single-visit difference images above which all optical transients are to be reported"* (LSR-REQ-0101; LSE-29). There is no minimum specification or stretch goal associated with transSNR. The OSS requires that *"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. *This requirement is to be interpreted as an average over the entire survey"* (OSS-REQ-0353; LSE-30). In other words, the DMS must be able to provide the value for a spuriousness threshold T, below which all difference sources detected with a signal-to-noise ratio transSampleSNR = 6, over the entire LSST survey, have transCompletenessMin = 90% and transPurityMin = 95%. This is not the same as a formal requirement on the fraction of false positives per visit in the alert stream, but the spuriousness threshold T will allow users to filter their stream to a fiducial completeness and purity. The DMSR does not appear to have any requirements on the fraction of false positives.

The SR&SSS estimates that the number of alerts per visit due to false positives will be 5050 **alerts, or** $\sim 50\%$ **of all alerts [LSE-81].** Based on a footnote on page 19 of LDM-151, which states that *"50% false positive rate is given in the OSS (when discussing Solar System Object*

requirements) and impacts the sizing model for the alert stream", it appears that this stems from the OSS specification that *"There shall exist a spuriousness threshold* T *for which the completeness and purity of difference sources are higher than* mopsCompletenessMin *and* mopsPurityMin*, respectively, at the SNR detection threshold* orbitObservationThreshold*. This requirement is intended to be interpreted as an average for any one month of observing"*, where orbitObservationThreshold $= 5$, mopsCompletenessMin $= 99\%$, and mopsPurityMin $= 50\%$ (OSS-REQ-0354; LSE-30).

2.6   Alert Packet Size

There are no formal requirements regarding the alert packet size. The size of an individual alert packet is estimated to be $\sim 82\,\mathrm{KB}$ (without schema or compression).

Sizing Estimate – Alert packet contents will include all of the LSST science data for the triggering detection, including a ~ 12 month historical record of detections, plus image stamp cutouts. Alert contents are described in more detail in Section 3.5 of LSE-163. Simulated alert packets based on the Apache Avro format average $\sim 82/126\,\mathrm{KB}$ with and without the schema, respectively. The application of gzip compression can further reduce these sizes to $\sim 68/65\,\mathrm{KB}$ (JIRA ticket DM-16280). Cutout stamps included in the alert will be at least 30×30 pixels and contain flux (32 bit/pix), variance (32 bit/pix), and mask (16 bit/pix) extensions for both the template and difference image, plus a header of metadata [LSE-163]. The stamps alone will contribute $\gtrsim 18\,\mathrm{KB}$ to the total size of the uncompressed alert packet.

What if brokers plan to do all the associations, and want an alert packet "lite" with no historical information and/or no stamps? Is that an option?

It might be useful to quote the distribution of alert packet sizes by doing a larger simulation.

2.7   Alert Stream Data Rate

There are no formal requirements regarding the alert stream data rate. The time-averaged data rate of the alert stream is estimated to be $\sim 27\,\mathrm{MB/sec}$ (for alert packets without schema or compression).

Sizing Estimate – The size of a single LSST alert will be $\sim 82\,\mathrm{KB}$ (including image stamps but not schema nor compression). Using an average of ~ 10000 alerts released per ~ 30 second

image, this leads to a *continuous* alert stream data rate of $\sim 27 \mathrm{~MB~s}^{-1}$.

Quoting a Peak Flow Value – The 10000 alerts per visit probably aren't going to be released evenly over the 60 seconds so we should quote the maximum expected stream rate. E.g., LDM-612 says *"about 7 brokers can receive the full stream if 5 seconds of the 60 second latency budget is devoted to outbound data transfer"*, but does that mean brokers should expect bursts of an alert stream for 5 out of every 60 seconds – which would be $\sim 150 \mathrm{~MB~s}^{-1}$ for 5 seconds? Is than an acceptable upper limit?

Quoting a Data Rate for Filtered Streams – If pre-filtering will be offered, it might be useful to quote stream rates for the expected common filtering. For example, if a broker wants to limit its incoming flow to an alert stream with `transCompletenessMin` $= 90\%$ and `transPurityMin` $= 95\%$ (i.e., apply `spuriousness` threshold `T`; § 2.5), then given the purity of the `transSNR` > 5 stream is only $\sim 50\%$ (§ 2.5 and LSE-81) this might decrease the stream by $\sim 40\%$.
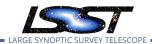
2.8   Number of Selected Brokers

There are no formal requirements on the number of brokers. The DM team estimates that resources will allow for the delivery of the alert stream to 4-7 brokers.

Formal Requirements – Neither the SRD, LSR, OSS, nor the DMSR place any formal requirements on – or even mention – the number of brokers that the alert stream should be delivered to. MLG: I thought *somewhere* a minimum of at least 4 was stated as a formal requirement? Neither the DPDD nor the SR&SSS mention an estimate for the number of brokers.

Sizing Estimate – As described in LDM-612, *"An allocation of $10 \mathrm{~Gbps}$ is baselined for alert stream transfer from the [LSST Data Facility (LDF)]. With an estimated packet size of $\sim 82 \mathrm{~KB}$ and 10000 alerts per visit, about 7 brokers can receive the full stream if 5 seconds of the 60 second latency budget is devoted to outbound data transfer."* Citation for the allocation of 10 Gbps?

2.9   Alert Database Volume

There are no formal requirements on the alerts database volume. The estimated upper limit is $\lesssim 3 \mathrm{~PB}$ (without schema or compression).

Sizing Estimate – An upper estimate is derived by starting with a maximum of ~ 1000 visits per night, and ~ 10000 alerts per visit, which amounts to ~ 10 million alerts per night. At ~ 82 KB per alert (§ 2.6), that leads to a total of ~ 782 GB per night. An extreme upper limit is 365 nights per year for 10 years, which would amount to ~ 2.7 PB *at the very most*. Therefore we quote an upper limit on the alerts database as $\lesssim 3$ PB. Compression could drastically lower this, as could reformatting: every alert contains a ~ 12 month historical record and links to the most recent DIAObject and DR Object catalogs. The set of alerts for the same transient/variable would contain significant redundant information which could be reformatted (i.e., removed and compiled).
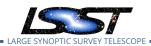
2.10 Number of New Transients per Visit

Estimate unclear. Need for quoting it here also unclear (except, it is called out in the spreadsheet).

MLG thoughts – This would refer to new, unassociated `DIASources` which are not false. The SR&SSS quotes the *"average number of alerts per visit due to new transients"* as 100, with no reference (Science Book maybe?). Here we could quote, for example, the average number of unassociated DIASources in either a new WFD extragalactic image or a new Galactic Plane image of a field that has not been observed in 3 days. Or the predicted number of unassociated DIASources which are actually new/first detections of moving objects. Brainstorm other useful quantities for brokers.

3 Mini-Broker

It is a formal requirement that the LSST provide a simple alerts filtering service for users, which is hereafter referred to as the LSST Mini-Broker.

Formal Requirements – The SRD specifies that *"users will have an option of a query-like pre-filtering of [the alert] data stream in order to select likely candidates for specific transient type"* and that *"several pre-defined filters optimized for traditionally popular transients, such as supernovae and microlensed sources, will also be available"* [LPM-17]. Neither the LSR nor the OSS have a formal requirement on this capability, as it is a product of the DMS. The DMSR has a formal requirement that *"a basic, limited capacity, alert filtering service shall be provided that can be given user defined filters to reduce the alert stream to manageable levels"*, and that this service

include *"a predefined set of simple filters"* (DMS-REQ-0342, -0348; LSE-61).

The term "user" in the DMSR is assumed to mean individual with data rights? Even when discussing public data products?

3.1   Number of Simultaneous Mini-Broker Users

It is a formal requirement that the mini-broker support a minimum of 100 simultaneous users.

Formal Requirements – The DMSR specifies that the LSST *"alert filtering service shall support* `numBrokerUsers` *simultaneous users"*, where `numBrokerUsers` = 100 (DMS-REQ-0343; LSE-61).

MLG: I can't find the background numbers that drive this limit of 100.

3.2   Number of Alerts per Visit Returned per User-Defined Filter

It is a formal requirement that the mini-broker return 20 alerts per visit per user.

Formal Requirements – The DMSR specifies that within the LSST alert filtering service *"each user [shall be] allocated a bandwidth capable of receiving the equivalent of* `numBrokerAlerts` *alerts per visit"*, where `numBrokerAlerts` = 20 (DMS-REQ-0343; LSE-61).

MLG: I can't find the background numbers that drive this limit of 20.

Requirement Drivers – Note that in a footnote of LDM-612, it says that the *"requirement on the number of simultaneously connected users and number of passed alerts is largely driven by outbound bandwidth limitations from the DAC at NCSA. We are investigating approaches that would support larger numbers of active filters"* (page 12; LDM-612). I can't find any references to mini-broker capacities in the SR&SSS.
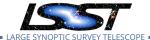
3.3   Alerts Database Query Latency

It is a formal requirement that *"All published transient alerts ... shall be available for query"* (OSS-REQ-0185; LSE-30). It is unclear whether `0TT1` = 1 minute or `L1PublicT` = 24 hours applies to the availability of alerts in the database. MLG: I think it should be the latter but am still looking
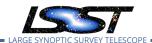
for the proof in the docs.

3.4   Summary

References

[LDM-612], Bellm, E., co authors, 2018, *Plans and Policies for LSST Alert Distribution*, LDM-612, URL https://ls.st/LDM-612

[LSE-29], Claver, C.F., The LSST Systems Engineering Integrated Project Team, 2017, *LSST System Requirements (LSR)*, LSE-29, URL https://ls.st/LSE-29

[LSE-30], Claver, C.F., The LSST Systems Engineering Integrated Project Team, 2018, *Observatory System Specifications (OSS)*, LSE-30, URL https://ls.st/LSE-30

[LSE-81], Dubois-Felsmann, G., 2013, *LSST Science and Project Sizing Inputs*, LSE-81, URL https://ls.st/LSE-81

[LSE-61], Dubois-Felsmann, G., Jenness, T., 2018, *LSST Data Management Subsystem Requirements*, LSE-61, URL https://ls.st/LSE-61

[LPM-17], Ivezić, Ž., The LSST Science Collaboration, 2018, *LSST Science Requirements Document*, LPM-17, URL https://ls.st/LPM-17

Ivezic, Z., et al., 2008, ArXiv e-prints (arXiv:0805.2366), ADS Link

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

[LDM-151], Swinbank, J.D., et al., 2017, *Data Management Science Pipelines Design*, LDM-151, URL https://ls.st/LDM-151