Rubin Observatory

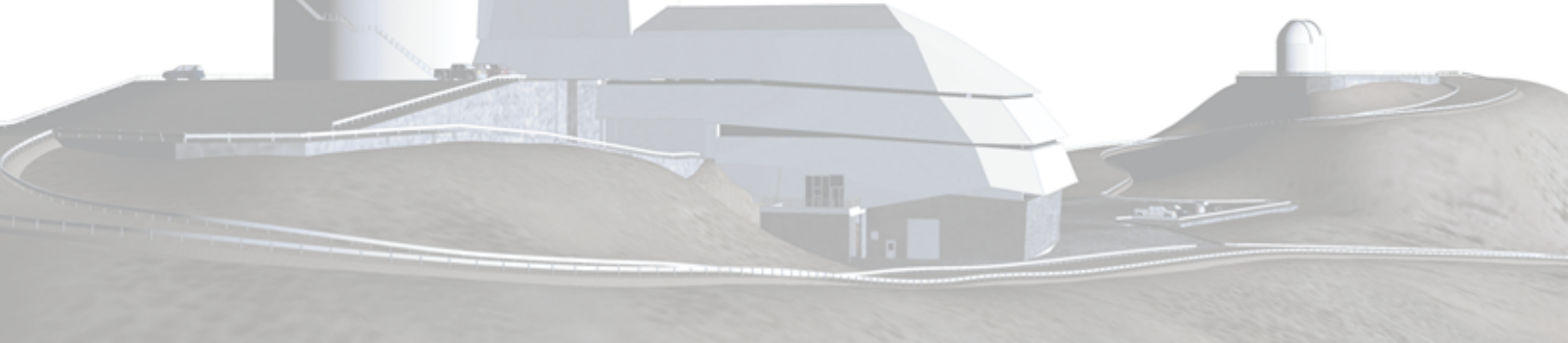
Vera C. Rubin Observatory
Systems Engineering

Data Management Data Backbone Services Requirements

Michelle Gower, Michelle Butler, K-T Lim

LDM-635

Latest Revision: 2019-06-14



Rubin Observatory

Change Record

Version	Date	Description	Owner name
	2018-06-21	Initial draft.	Spec: M. Gower, Model: T. Jenness
	2019-06-14	Add introduction and use folders	M. Gower; T. Jenness
1.0	2019-07-16	First release, approved in RFC-604	M. Gower; T. Jenness

Document source location: MagicDraw SysML
Version from source repository: 240

Rubin Observatory

Contents

Rubin Observatory

Data Management Data Backbone Services Requirements

The LSST DM System needs to maintain a set of data storage resources distributed between the Base Center and NCSA to hold the record of the survey and to be used by Prompt Products, Data Release Products, and Data Access Center (DAC) services. The Data Backbone (DBB) is the component that will provide storage, management, and replication for all LSST data products.

The scope of the Data Backbone includes both files and data maintained by relational and other database engines. While the Data Backbone provides read-only data service to the US and Chilean DACs, it will not host data stores where DAC users create state. For example, DAC resources such as Qserv and User Database Workspaces are provisioned in the context of a DAC, not the Data Backbone.

This document presents requirements for the Data Backbone. Currently these initial requirements are derived directly from LSE-61 requirements. The requirements will increase in number and detail as requirements flow down from the other components/services that use the Data Backbone.

1 Core Requirements

1.1 Pipeline Availability

ID: DMS-DBB-REQ-0001

Specification: Except in cases of major disaster, the Data Backbone shall not have an unscheduled outage which in turn causes an unscheduled outage of any active DMS production pipelines extending over a period greater than **productionMaxDowntime**.

Description	Value	Unit	Name
Maximum allowable outage of active DM production.	24	hour	productionMaxDowntime

Derived from Requirements:

DMS-REQ-0008: Pipeline Availability

1.2 Raw Science Image Metadata

ID: DMS-DBB-REQ-0002

Specification: For each raw science image, the Data Backbone shall store image metadata including at least the information detailed in DMS-REQ-0068

Derived from Requirements:

DMS-REQ-0068: Raw Science Image Metadata

1.3 Difference Exposure Attributes

ID: DMS-DBB-REQ-0003

Specification: For each Difference Exposure, the Data Backbone Services shall store at least the information detailed in DMS-REQ-0074

Derived from Requirements:

DMS-REQ-0074: Difference Exposure Attributes

1.4 Maintain Archive Publicly Accessible

ID: DMS-DBB-REQ-0004

Specification: The Data Backbone Services shall maintain and preserve the catalog files for all releases in a publicly accessible state for the entire operational life of the LSST observatory.

Discussion: The scientific intent is satisfied by keeping data products from the current DRP release and the one prior available with low-latency, in a form readily queryable by the public. Earlier releases may be available from deep-store with potentially high latency, for bulk download by users. The queryable catalogs (i.e., loaded into a database) may only be in QServ in which case this is not part of the Data Backbone.

Derived from Requirements:

DMS-REQ-0077: Maintain Archive Publicly Accessible

1.5 Solar System Objects Available Within Specified Time

ID: DMS-DBB-REQ-0005

Specification: The Data Backbone Services shall replicate tables containing detected moving objects and associated metadata to all endpoints to enable public access in the DMS science data archive within time **L1PublicT** of their generation by the DMS.

Description	Value	Unit	Name
Maximum time from the acquisition of science data to the public release of associated Level 1 Data Products (except alerts)	24	hour	L1PublicT

Derived from Requirements:

DMS-REQ-0089: Solar System Objects Available Within Specified Time

1.6 Keep Historical Alert Archive

ID: DMS-DBB-REQ-0006

Specification: The Data Backbone Services shall preserve and replicate in an accessible state an alert archive with all issued alerts for a historical record and for false alert analysis.

Discussion: Part of L1 Data Products which have availability requirements at both endpoints. May not be in consolidated database.

Derived from Requirements:

DMS-REQ-0094: Keep Historical Alert Archive

1.7 Public access to Engineering and Facility Database Archive data

ID: DMS-DBB-REQ-0007

Specification: The Data Backbone Services shall permanently archive and replicate the data provided by the EFD ETL Service to be available at both endpoints for public access within **L1PublicT** hours of their original generation by the OCS.

Description	Value	Unit	Name
Maximum time from the acquisition of science data to the public release of associated Level 1 Data Products (except alerts)	24	hour	L1PublicT

Derived from Requirements:

DMS-REQ-0102: Provide Engineering & Facility Database Archive

1.8 Level 3 Access to Metadata and Provenance

ID: DMS-DBB-REQ-0008

Specification: The Data Backbone Services shall store and make queryable metadata and provenance that allow Level 3 processing tasks ensure that they are getting self-consistent inputs from the Data Backbone.

Discussion: This requirement does not mean that the LSST Data Backbone is responsible

for saving the outputs of the Level 3 processing tasks nor the metadata and provenance for those outputs.

Derived from Requirements:

DMS-REQ-0120: Level 3 Data Product Self Consistency

1.9 Maintaining and validatting exported data

ID: DMS-DBB-REQ-0009

Specification: The Data Backbone shall facilitate Level 3 image processing that may take place at external facilities outside the DACs by facilitating the export of catalogs and the provision of tools for maintaining and validating exported data.

Discussion: Maintaining is doing things like checking to see whether anything has been accidentally removed from my exported collection and re-exporting it if it was. Checksums are sufficient for validating, although digital signatures would be better.

Derived from Requirements:

DMS-REQ-0122: Access to catalogs for external Level 3 processing

1.10 Federation with external catalogs

ID: DMS-DBB-REQ-0010

Specification: For project approved externally provided catalogs, the Data Backbone shall load and replicate the catalog for accessibility at all endpoints.

Discussion: The project will provide specification for how external data must be provided in order to make loading possible. These catalogs can be joined with other catalogs in the Data Backbone as well as LSP user catalogs that are loaded into their consolidated database quota.

Derived from Requirements:

DMS-REQ-0124: Federation with external catalogs

1.11 Level 3 image processing facilitation

ID: DMS-DBB-REQ-0011

Specification: The Data Backbone shall facilitate Level 3 image processing that may take place at external facilities outside the DACs by facilitating the export of image datasets and the provision of tools for maintaining and validating exported data.

Discussion: Maintaining is doing things like checking to see whether anything has been accidentally removed from my exported collection and re-exporting it if it was. Checksums are sufficient for validating, although digital signatures would be better.

Derived from Requirements:

DMS-REQ-0126: Access to images for external Level 3 processing

1.12 Calibration Data Products and Image Provenance

ID: DMS-DBB-REQ-0012

Specification: The Data Backbone Services shall archive and replicate calibration data products, specified in DMS-REQ-130, including metadata, provenance, and calibration database tables.

Derived from Requirements:

DMS-REQ-0130: Calibration Data Products
DMS-REQ-0132: Calibration Image Provenance

1.13 Daily Calibration Products Update Availability

ID: DMS-DBB-REQ-0013

Specification: The Data Backbone Services shall be capable of supporting daily calibration product updates to be generated from a group of **nCalExpProc** related exposures where outputs must be accessible from the Data Backbone Services within **calProcTime** of the end of the acquisition of images/data for that group

Discussion: calProcTime includes many things so the requirement needs to be broken down into its components to understand exact time requirements for the Data Backbone.

Description	Value	Unit	Name
Maximum number of calibration exposures that can be processed together within time **calProc-Time**.	25	integer	nCalExpProc
Time allowed to process **nCalExpProc** calibration exposures and have them available within the DMS.	1200	second	calProcTime

Derived from Requirements:

DMS-REQ-0131: Calibration Images Available Within Specified Time

1.14 Optimization of Cost, Reliability and Availability in Order

ID: DMS-DBB-REQ-0014

Specification: Within a fixed cost envelope for the Data Backbone Services, the allocation of storage facilities will optimize reliability over availability to end users.

Discussion: Reliability of services needed for production over availability/access for end users. (Bursts of end user activity can mean temporary higher priority resources for end users.)

Derived from Requirements:

DMS-REQ-0161: Optimization of Cost, Reliability and Availability in Order

1.15 Pipeline Throughput

ID: DMS-DBB-REQ-0015

Specification: The Data Backbone Services shall be capable of supporting the data processing pipelines in completing processing of a night's observing data prior to the start of the next observing night, assuming no system outages

Discussion: While the crosstalked raw files will not be retrieved from the Data Backbone, templates and master calibrations that are used in the Alert Processing are loaded from the DBB

Derived from Requirements:

DMS-REQ-0162: Pipeline Throughput

1.16 Re-processing Capacity

ID: DMS-DBB-REQ-0016

Specification: The Data Backbone Services shall be capable of supporting the execution of the DMS Data Release Production over all pre-existing survey data in a time no greater than **drProcessingPeriod**, without impacting observatory operations.

Description	Value	Unit	Name
Duration of the Data Release Production, including quality validation.	1	year	drProcessingPeriod

Derived from Requirements:

DMS-REQ-0163: Re-processing Capacity

1.17 Temporary Storage for Communications Links

ID: DMS-DBB-REQ-0017

Specification: The infrastructure for the Data Backbone Services will provide for temporary storage for a minimum of **tempStorageRelMTTR** of the mean time to repair of any communications network link at or before the source end of that link.

Discussion: Most of the storage in the Data Backbone Services is actually permanent storage. This requirement does cover any internal temporary storage used by the Data Backbone Services. This requirement does not cover any temporary storage used by other services for data prior to ingesting into the Data Backbone Services.

Description	Value	Unit	Name
Temporary storage required relative to network Mean Time to Repair.	200	percent	tempStorageRelMTTR

Derived from Requirements:

DMS-REQ-0164: Temporary Storage for Communications Links

1.18 Infrastructure Sizing for “catching up”

ID: DMS-DBB-REQ-0018

Specification: The infrastructure for the Data Backbone Services will be sized such that after outages, "catch up" processing of the temporarily stored raw image data may occur at the rate of one night’s observing data processed per day, without interrupting the current day’s observatory operations.

Derived from Requirements:

DMS-REQ-0165: Infrastructure Sizing for "catching up"

1.19 Fault Tolerance Features

ID: DMS-DBB-REQ-0019

Specification: The Data Backbone Services will incorporate infrastructure with fault-tolerance features to prevent loss of data in the event of hardware or software failure.

Derived from Requirements:

DMS-REQ-0166: Incorporate Fault-Tolerance

1.20 Incorporate Autonomics

ID: DMS-DBB-REQ-0020

Specification: The infrastructure for the Data Backbone Services will incorporate sufficient capability for self-diagnostics and recovery to provide for continuation of processing in the event of partial hardware or software failures.

Derived from Requirements:

DMS-REQ-0167: Incorporate Autonomics

1.21 Base Facility Infrastructure

ID: DMS-DBB-REQ-0021

Specification: The infrastructure of the Data Backbone Services at the Base Facility shall be sufficient to store a copy of designated data from the Archive Facility data holdings, and to provide the access capabilities to support Commissioning activities and the Chilean DAC.

Discussion: TBD: Determine details to guide infrastructure sizing such as number of DB connections, number of simultaneous file downloads, transfer throughput, This requirement does not include raw ingestion as still no requirements yet for whether raw files must be accessible at Base via DBB while NCSA endpoint is unavailable to initiate the ingestion.

Derived from Requirements:

DMS-REQ-0176: Base Facility Infrastructure

1.22 Archive Center

ID: DMS-DBB-REQ-0022

Specification: The infrastructure of the Data Backbone Services at the Archive Center shall support, simultaneously: nightly processing including image processing, detection, association, and moving object pipelines, and the generation of all time-critical data products, i.e. alerts; the data release production, including Level-2 data product creation, permanent storage for all data products (with provenance), including federated Level-3 products; and serve data for replication to data centers and end user sites.

Derived from Requirements:

DMS-REQ-0185: Archive Center

1.23 Archive Center Disaster Recovery

ID: DMS-DBB-REQ-0023

Specification: The Data Backbone Services shall provide disaster recovery support preventing loss of LSST data in the case of infrastructure or facility-threatening events. This support shall enable recovery of all LSST archived data from backed up sources, including project Data Access Centers.

Derived from Requirements:

DMS-REQ-0186: Archive Center Disaster Recovery

1.24 No Limit on Data Backbone Endpoints

ID: DMS-DBB-REQ-0024

Specification: The number of Data Backbone endpoints shall be limited only by available internal or external funding. No architectural constraints will be placed on the Data Backbone Services that prohibit the addition of Data Backbone endpoints at any time, subject to funding.

Derived from Requirements:

DMS-REQ-0197: No Limit on Data Access Centers

1.25 Exposure Catalog

ID: DMS-DBB-REQ-0025

Specification: The Data Backbone Services shall store an Exposure Catalog containing information for each exposure that includes information detailed in DMS-REQ-0266.

Discussion: This requirement includes the catalog file itself as well as the contents in a queryable table following lifecycle policies.

Derived from Requirements:

DMS-REQ-0266: Exposure Catalog

1.26 Catalog

ID: DMS-DBB-REQ-0026

Specification: The Data Backbone Services shall store a catalog of all Sources detected on Difference Exposures with SNR > transSNR. For each Difference Source (DIASource), the catalog will include information detailed in DMS-REQ-0269.

Discussion: This requirement includes the catalog file itself as well as the contents in a queryable table following lifecycle policies for the DIASource catalogs from the Alert Processing. The contents of the DIASource catalogs from Data Release Processing instead go into Qserv.

Derived from Requirements:

DMS-REQ-0269: DIASource Catalog

1.27 DIAObject Catalog

ID: DMS-DBB-REQ-0027

Specification: The Data Backbone Services shall store a catalog of all astrophysical objects identified through difference image analysis (DIAObjects). The DIAObject entries shall include metadata attributes including at least those detailed in DMS-REQ-0271.

Discussion: This requirement includes the catalog file itself as well as the contents in a queryable table following lifecycle policies for the DIAObject catalogs from the Alert Processing. The contents of the DIAObject catalogs from Data Release Processing instead go into Qserv.

Derived from Requirements:

DMS-REQ-0271: DIAObject Catalog

1.28 SSObject Catalog

ID: DMS-DBB-REQ-0028

Specification: The Data Backbone Services shall store a catalog of all Solar System Objects (SSObjects) that have been identified via Moving Object Processing. The SSObject catalog shall include for each entry attributes including at least those detailed in DMS-REQ-0273.

Discussion: This requirement includes the catalog file itself as well as the contents in a queryable table following lifecycle policies.

Derived from Requirements:

DMS-REQ-0273: SSObject Catalog

1.29 DIASource Precovery

ID: DMS-DBB-REQ-0029

Specification: For all DIASources not associated with either DIAObjects or SSObjects, the Data Backbone Services shall support the calculation of forced photometry at the location of the new source (precovery) on all Difference Exposures obtained in the prior **precoveryWindow**, and make the results publicly available within **L1PublicT**.

Description	Value	Unit	Name
Maximum look-back time for precovery measurments on prior Exposures.	30	day	precoveryWindow
Maximum time from the acquisition of science data to the public release of associated Level 1 Data Products (except alerts)	24	hour	L1PublicT

Derived from Requirements:

DMS-REQ-0287: DIASource Precovery

1.30 Query Repeatability

ID: DMS-DBB-REQ-0030

Specification: The Data Backbone Services shall ensure that the data is available such that any query executed at a particular point in time against any Data Backbone delivered database shall be repeatable at a later date, and produce results that are either identical or include additional results (owing to updates from Level-1 processing).

Discussion: Schemas may change as long as the front-end services adapt old queries to the new schema. The "delivered database" is not the "ForcedSource Catalog" for all time; it's the "ForcedSource Catalog for DR5" or the "ForcedSource Catalog for DR6" etc.

Derived from Requirements:

DMS-REQ-0291: Query Repeatability

1.31 Uniqueness of IDs Across Data Releases

ID: DMS-DBB-REQ-0031

Specification: To reduce the likelihood for confusion, all IDs shall be unique across databases and database versions, other than those corresponding to uniquely identifiable entities (i.e., IDs of exposures).

Discussion: For example, DR4 and DR5 (or any other) release will share no identical Object, Source, DIAObject or DIASource IDs. However, IDs shall be identical for the same data across Data Backbone endpoints.

Derived from Requirements:

DMS-REQ-0292: Uniqueness of IDs Across Data Releases

1.32 Selection of Datasets

ID: DMS-DBB-REQ-0032

Specification: A Dataset may consist of one or more pixel images, a set of records in a file or database, or any other grouping of data that are processed or produced as a logical unit. The Data Backbone Services shall store and replicate data needed to be able to identify data in the Data Backbone and retrieve complete, consistent datasets from the Data Backbone for processing.

Derived from Requirements:

DMS-REQ-0293: Selection of Datasets

1.33 Processing of Datasets

ID: DMS-DBB-REQ-0033

Specification: The Data Backbone Services shall support codes that process all requested datasets until either a successful result is recorded or a permanent failure is recognized. If the results of any dataset is processed, in part or in whole, more than once and stored in the Data Backbone, the Data Backbone Services will support the labeling such that only one of the wholly processed results will be found for further processing.

Discussion: The requirement is limited to within a single processing campaign. It is not just about what gets released but also about what gets used in subsequent processing.

Derived from Requirements:

DMS-REQ-0294: Processing of Datasets

1.34 Data Product Ingest

ID: DMS-DBB-REQ-0034

Specification: The Data Backbone Services shall provide software for designated staff or processes to ingest designated data products into the Data Backbone.

Derived from Requirements:

DMS-REQ-0299: Data Product Ingest

1.35 Raw Data Archiving Reliability

ID: DMS-DBB-REQ-0035

Specification: The Data Backbone Services shall archive all data, including science, wavefront, and guider images and associated metadata, that are presented for archiving by upstream systems, with a rate of permanent data loss or corruption not to exceed **dataLossMax**.

Discussion: Data obtained for diagnostic and other limited-use engineering purposes are specifically excluded from this requirement.

Description	Value	Unit	Name
Maximum fraction of raw images that are permitted to be permanently lost or corrupted, including the loss or corruption of essential associated metadata, once acquired by the DMS.	1.0e-5	float	dataLossMax

Derived from Requirements:

DMS-REQ-0309: Raw Data Archiving Reliability

1.36 Un-Archived Data Product Cache

ID: DMS-DBB-REQ-0036

Specification: The Data Backbone Services shall provide low-latency storage for un-archived Level 1 data products of at least **l1CacheLifetime** to enable efficient precovery and other Level-1 production measurements.

Discussion: Data must be available at both endpoints.

Description	Value	Unit	Name
Lifetime in the cache of un-archived Level-1 data products.	30	day	l1CacheLifetime

Derived from Requirements:

DMS-REQ-0310: Un-Archived Data Product Cache

1.37 Level 1 Data Product Access

ID: DMS-DBB-REQ-0037

Specification: The Data Backbone shall store and replicate a Level 1 Database for query by science users, updated as a result of Alert Production processing.

Derived from Requirements:

DMS-REQ-0312: Level 1 Data Product Access

1.38 Level 1 and 2 Catalog Access

ID: DMS-DBB-REQ-0038

Specification: The Data Backbone Services shall store and replicate designated catalogs from the most recent two Data Releases for query by science users, as well as versions of the most recent catalogs generated from Special Programs data.

Discussion: There is no requirement for older data releases to be queryable.

Some catalogs from the Data Releases may be only be stored in QServ (which is not in the Data Backbone)

Derived from Requirements:

DMS-REQ-0313: Level 1 & 2 Catalog Access

1.39 Compute Platform Heterogeneity

ID: DMS-DBB-REQ-0039

Specification: At any given LSST facility the Data Backbone Services shall be capable of operations on a heterogeneous cluster of machines. The hardware, operating system, and other machine parameters shall be limited to a project-approved set.

Discussion: The necessity of replacing hardware throughout the course of the survey essentially guarantees heterogeneity within a cluster.

Derived from Requirements:

DMS-REQ-0314: Compute Platform Heterogeneity

1.40 DIAForcedSource Catalog

ID: DMS-DBB-REQ-0040

Specification: The Data Backbone Serves shall store a DIAForcedSource Catalog which in-

cludes at least information detailed in DMS-REQ-0317.

Discussion: This requirement includes the catalog file itself as well as the contents in a queryable table following lifecycle policies for the DIAForcedSource catalogs from the Alert Processing. The contents of the DIAForcedSource catalogs from Data Release Processing instead go into Qserv.

Derived from Requirements:

DMS-REQ-0317: DIAForcedSource Catalog

1.41 Unscheduled Downtime

ID: DMS-DBB-REQ-0041

Specification: The Data Backbone Servcies shall be designed to facilitate unplanned repair activities that directly prevent the collection of survey data expected not to exceed **DMDowntime** days per year.

Discussion: The reference case would be a failure of communication or archiving that lasted longer than the capacity of the Summit buffer – i.e., an 11-day outage would exceed the nominal buffer capacity by one day and therefore use up the proposed allocation.

This requirement does not invoke the need to verify by reliability analysis. Verification is by analysis that identifies likely hardware failures and identifies mitigations to minimize downtime caused by those failures.

Description	Value	Unit	Name
Unplanned downtime per year.	1	day	DMDowntime

Derived from Requirements:

DMS-REQ-0318: Data Management Unscheduled Downtime

1.42 Special Programs Database

ID: DMS-DBB-REQ-0042

Specification: Data products for special programs shall be stored in databases/schemas that are distinct from those used to store standard Level 1 and Level 2 data products. It shall be possible for these databases to be federated with the Level 1 and Level 2 databases to allow cross-queries and joins.

Discussion: Separate database tables are prefered over the same database table with a label column (which could be more error-prone).

Derived from Requirements:

DMS-REQ-0322: Special Programs Database

1.43 Persisting Data Products

ID: DMS-DBB-REQ-0043

Specification: All per-band deep coadds and best seeing coadds shall be kept indefinitely and made available to users.

Derived from Requirements:

DMS-REQ-0334: Persisting Data Products

1.44 Targeted Coadds

ID: DMS-DBB-REQ-0044

Specification: The Data Backbone Services shall be sized such that all generated coadds that cover designated small sections of sky can be accessible with low-latency to support quality

assessment and targeted science.

Derived from Requirements:

DMS-REQ-0338: Targeted Coadds

1.45 Tracking Characterization Changes Between Data Releases

ID: DMS-DBB-REQ-0045

Specification: The Data Backbone Services shall be sized such that designated small, overlapping, samples of data from older releases shall be kept loaded in the database.

Discussion: This requirement covers subsets of data already in databases in the Data Backbone. It does not mean that data only loaded in Qserv will now be copied back to the Data Backbone's databases.

Derived from Requirements:

DMS-REQ-0339: Tracking Characterization Changes Between Data Releases

1.46 Constraints on Level 1 Special Program Products Generation

ID: DMS-DBB-REQ-0046

Specification: The Data Backbone Services shall support the publishing of Level 1 data products from Special Programs within the same performance requirements of the standard Level 1 system. In particular **L1PublicT** and **OTT1**.

Description	Value	Unit	Name
The latency of reporting optical transients following the completion of readout of the last image of a visit	1	minute	OTT1

Continued on next page

Description	Value	Unit	Name
Maximum time from the acquisition of science data to the public release of associated Level 1 Data Products (except alerts)	24	hour	L1PublicT

Derived from Requirements:

DMS-REQ-0344: Constraints on Level 1 Special Program Products Generation

1.47 Raw Data Availability

ID: DMS-DBB-REQ-0047

Specification: All raw data used to generate any public data product (raw exposures, calibration frames, telemetry, configuration metadata, etc.) shall be kept and made available for download.

Derived from Requirements:

DMS-REQ-0346: Data Availability

1.48 Access to Previous Data Releases

ID: DMS-DBB-REQ-0048

Specification: The Data Backbone Services shall provide access to the current Level 1 data, and a number of Data Releases at least the most recent **nDRMin** Data Releases, and multiple older Data Releases where the exact number is to be determined by Operations funding.

Discussion: This requirement does not include queryable catalogs only loaded into Qserv as Qserv is not part of the Data Backbone Services.

Description	Value	Unit	Name
Minimum number of recent data releases	2	integer	nDRMin

Derived from Requirements:

DMS-REQ-0363: Access to Previous Data Releases

1.49 Data Access Services

ID: DMS-DBB-REQ-0049

Specification: The Data Backbone Services shall be designed to permit, and their software implementation shall support, the service of at least **nDRTot** Data Releases accumulated over the **surveyYears**-year planned survey.

Discussion: It is an operations-era decision to choose the actual number of releases to be served, and to allocate hardware resources accordingly.

Description	Value	Unit	Name
Length of the survey in years	10	integer	surveyYears
Total number of data releases over the survey.	11	integer	nDRTot

Derived from Requirements:

DMS-REQ-0364: Data Access Services

1.50 Operations Subsets

ID: DMS-DBB-REQ-0050

Specification: The Data Backbone Services shall be designed to permit the service of operations-designated subsets of the full content of the “older Data Releases” referred to in DMS-REQ-

0363.

Derived from Requirements:

DMS-REQ-0365: Operations Subsets

1.51 Subsets Support

ID: DMS-DBB-REQ-0051

Specification: The Data Backbone Services shall be designed to support the service of operations-designated subsets of the content of the "older Data Releases" referred to in requirement DMS-REQ-0363 from high-latency media.

Discussion: This means that the "toolkit" of Data Backbone Services should include elements that, for instance, allow users to understand that certain queries (e.g., for data on tape) may take much longer than for current data releases, and to monitor the status of such queries.

Derived from Requirements:

DMS-REQ-0366: Subsets Support

1.52 Access Services Performance

ID: DMS-DBB-REQ-0052

Specification: The Data Backbone Services shall support Data Access Services to meet the performance requirements set forth in OSS-REQ-0180 and OSS-REQ-0181 for the most recent **nDRMin** Data Releases.

Description	Value	Unit	Name
Minimum number of recent data releases	2	integer	nDRMin

Derived from Requirements:

DMS-REQ-0367: Access Services Performance

1.53 Implementation Provisions

ID: DMS-DBB-REQ-0053

Specification: Nothing in the design and software implementation of the Data Backbone Services shall prevent the performance requirements set forth in OSS-REQ-0180 and OSS-REQ-0181 from being met for the "older Data Releases" referred to in DMS-REQ-0363, subject to the provision of sufficient computing and storage resources in the operations era.

Discussion: It is left to the operations project to set standards for the performance on older releases, but they should not be limited by design choices made in the construction era. That is, the system must be scalable to handle full-performance service of all Data Releases, should the operations project so choose. This situation does not arise until, at the release of Data Release (**nDRMin**+1), the operations project must decide on the level of service to be provided for Data Release 1.

Description	Value	Unit	Name
Minimum number of recent data releases	2	integer	nDRMin

Derived from Requirements:

DMS-REQ-0368: Implementation Provisions

1.54 Evolution

ID: DMS-DBB-REQ-0054

Specification: The Data Backbone Servcies shall be designed to accommodate evolution of the LSST data model from Data Release to Data Release.

Derived from Requirements:

DMS-REQ-0369: Evolution

1.55 Older Release Behavior

ID: DMS-DBB-REQ-0055

Specification: Apart from the flexibility provided by requirements DMS-REQ-0365, DMS-REQ-0366, DMS-REQ-0368, and DMS-REQ-0369, the qualitative behavior of the Data Backbone Services on the "older Data Releases" defined in DMS-REQ-0363 shall match that for the most recent **nDRMin** Data Releases.

Description	Value	Unit	Name
Minimum number of recent data releases	2	integer	nDRMin

Derived from Requirements:

DMS-REQ-0370: Older Release Behavior

1.56 Atomic Accessibility

ID: DMS-DBB-REQ-0056

Specification: The Data Backbone Services shall ensure data is accessible at a given endpoint in an atomic manner as not to allow conditions where ingestion or replication is partially complete at time of access.

Discussion: This requirement does not imply restrictions at a collection level. This requirement should help drive the discussions about timing regarding file replication vs database entries replication (especially Repository entries) when discussing technologies for replication. Designated operations staff and certain monitoring processes should not be prohibited from seeing data in partial state.

References

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