

Centre de Calcul
de l'Institut National de Physique Nucléaire
et de Physique des Particules

News on computing for LSST at FrDF

fabio hernandez, quentin le boulc'h, gabriele mainetti

doc.lsst.eu

DATA RELEASE SCHEDULE

- [RTN-011](#) “Rubin Observatory Plans for an Early Science Program”

Rubin Early Science – Data Release Scenario								
Data Product	Jun 2021	Jun 2022	Jun 2023	Jun 2025 – Jul 2025	Mar 2026 – May 2026	Sep 2026 – Jan 2027	Sep 2027 – Jan 2028	Sep 2028 – Nov 2028
	DP0.1	DP0.2	DP0.3	DP1	DP2	DR1	DR2	DR3
	DC2 Simulated Sky Survey	Reprocessed DC2 Survey	Solar System PPDB Simulation	ComCam Data	LSSTCam Science Validation Data	LSST First 6 Months Data	LSST Year 1 Data	LSST Year 2 Data
Raw Images	●	●	-	●	●	●	●	●
DRP Processed Visit Images and Source Catalogs	●	●	-	●	●	●	●	●
DRP Coadded Images and Object Catalogs	●	●	-	●	●	●	●	●
DRP Cell-based Coadded Images and ShearObject Catalog	-	-	-	-	●	●	●	●
DRP ForcedSource Catalogs	●	●	-	●	●	●	●	●
DRP Difference Images and DIA Catalogs	-	●	-	●	●	●	●	●
DRP SSP Catalogs	-	-	●	●	●	●	●	●

Table 1: Summary of the main data products expected in each data preview and early LSST data releases. A dark teal dot denotes confirmed data products whereas a gray dot denotes data products that currently remain a stretch goal.

Data Preview 1 (DP1) to be released on June 30th, 2025
See [announcement](#) in Community
Access to data via the Rubin Science Platform at <https://data.lsst.cloud>



US Data Facility

SLAC

HQ Site

UK Data Facility

France Data Facility

Brazil

Cerro Pachón

CHILE

Santiago

DATA PROCESSING CAMPAIGNS

- Coordinated by the Campaign Management team
execution at FrDF lead by Quentin Le Boulc'h
a run every two weeks with the latest available version of the LSST Science Pipelines
inputs: HSC and simulated DESC data
- Currently preparing a campaign to generate products similar to Data Preview 1
LSSTComCam data
identifying the right calibration data more difficult than initially anticipated ([DM-48746](#))

STORAGE

- Client for interacting with dCache completely rewritten

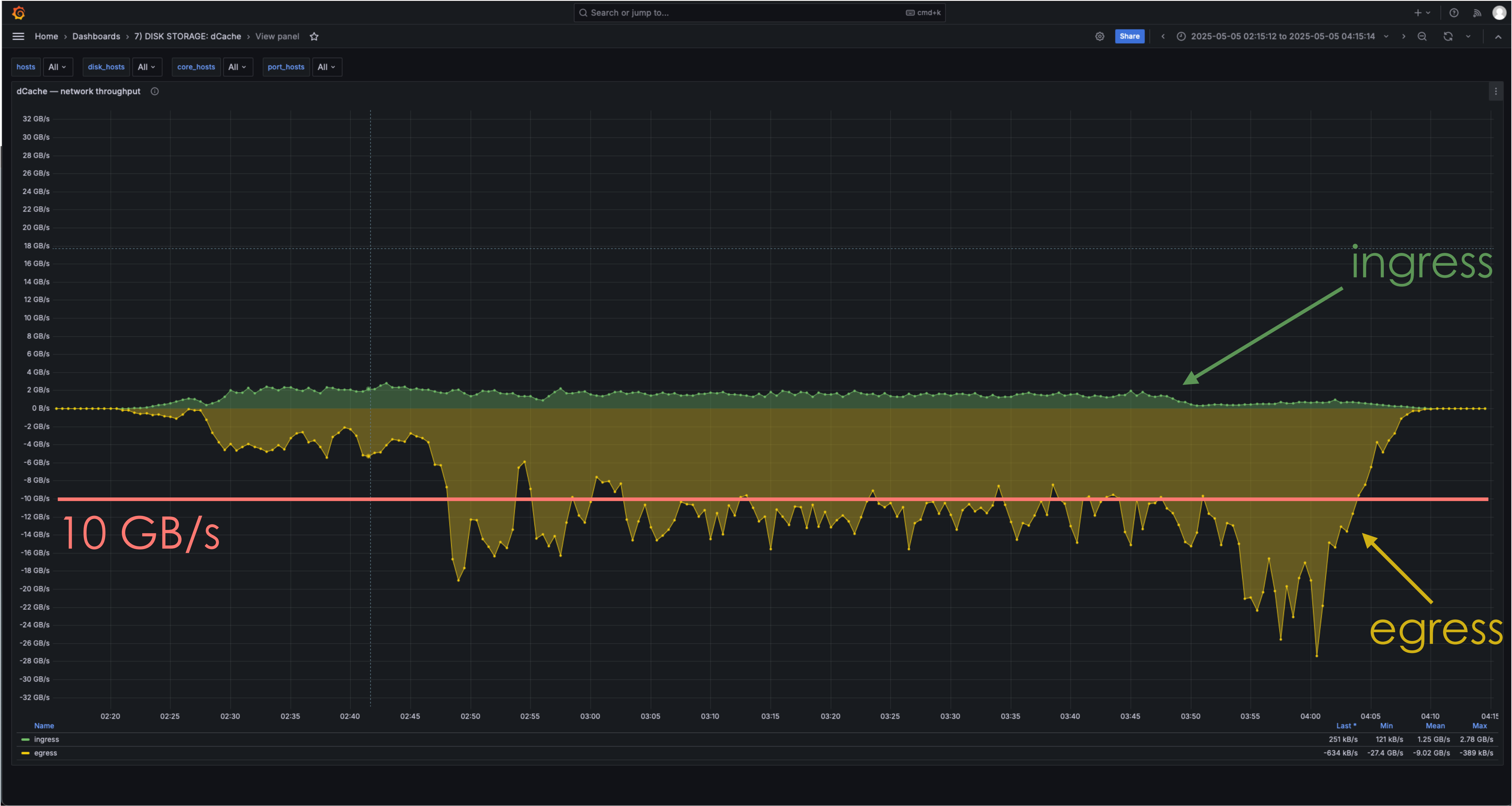
Butler uses this client to upload and download data, to manage metadata

takes into account the specifics of dCache behavior for improving performance and to work at the expected scale

integrated to the LSST Science Pipelines since release w_2025_20

JIRA: [DM-49784](#)

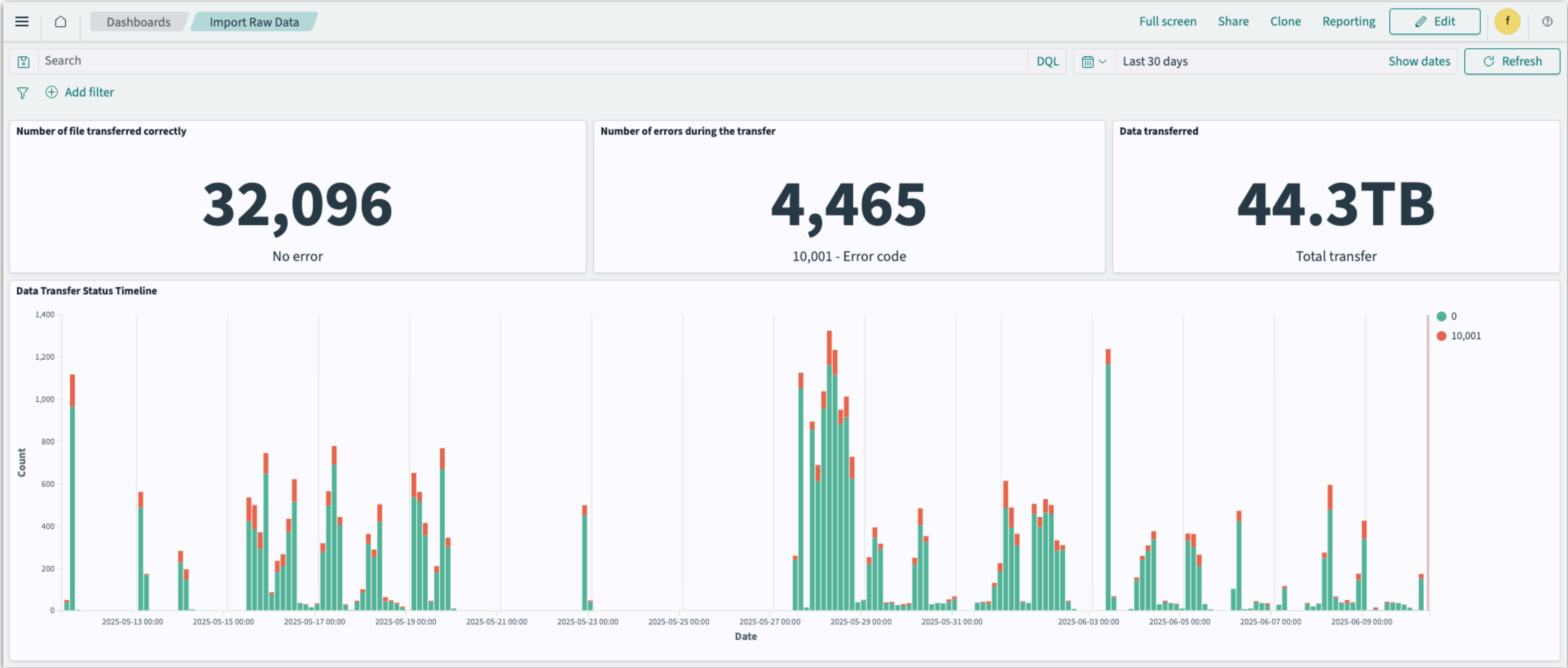
STORAGE (CONT.)



DATA REPLICATION

- Full set of LSSTComCam raw data replicated from USDF
- LSSTCam raw data replicated daily to FrDF since mid April
 - off-sky raw exposures replicated without delay, typically by the end of the observation day*
 - on-sky raw exposures replicated after the end of their embargo period (30 days during commissioning)*
 - aggregated: 18k exposures, 52 TB*
- Transfer efficiency to be improved
 - first-try transfer failure rate of approximately 15%*
 - failure scenario understood but root cause not yet identified*
 - details [DM-50535](#)*
- Tools for verification and reconciliation to be developed
 - details [DM-50982](#)*

DATA REPLICATION (CONT.)



Raw exposures replicated from USDF to FrDF over the period 2025-05-13 to 2025-06-09

ARCHIVAL OF RAW DATA

- FrDF committed to save on tape a copy of the entire raw data set
mainly for the purposes of recovery in case of disaster
- Configuration of dCache prepared and tested
to write on tape every raw exposure
to keep on disk only a copy of the exposures to be processed at FrDF
- Final details of the configuration of HPSS being worked out
- Automation and activation of the system will follow

IDENTITY AND ACCESS MANAGEMENT SYSTEM

- An evaluation instance of INDIGO IAM deployed
Rubin data rights holders, members of the LSST France community who need access to raw data to be enrolled in this system
IAM delivers access tokens to storage conveniently scoped to prevent accidental deletion of precious data
- Currently in the final stages of tests before opening for users to enrol

SCIENCE PLATFORM

- FrDF's Rubin Science Platform up to date
including the underlying operating system version
- Available catalogs
cosmoDC2_v1_1_4_image
dp01_dc2_catalogs
dp02_dc2_catalogs
dp02_dc2_catalogs_frdf
skysim5000_v1_1_1_parquet
- DP0.2 image data set

data-dev.lsst.eu/portal/app/?__action=layout.showDropDown&view=LSST%20DP0.2

Results DP0.2 Images DP0.2 Images SIAv2 DP0.2 Catalogs DP0.3 Catalogs SIAv2 Searches Upload Background Monitor Logout

LSST DP0.2 DC2 Tables

Table Collection (Schema): dp02_dc2_catalogs_frdf (tables: 8)
Data Preview 0.2 contains the image and catalog products of the Rubin Science Pipelines v23 processing of the DESC Data Challenge 2 simul...
Table Collection (Schema) count: 7

Tables: dp02_dc2_catalogs_frdf.Object
Properties of the astronomical objects detected and measured on the deep coadded images.
Table count: 8

View: UI assisted Edit ADQL

Enter Constraints ?

☒ Spatial ? no target provided

Shape Type: ☒ Cone Shape ☐ Polygon Shape

Coordinates or Object Name Try NED then Simbad

Examples: 62, -37 60.4 -35.1 4h11m59s -32d51m59s equ J2000 239.2 -47.6 gal NGC 1532 (NB: DC2 is a simulated sky, so names are not useful)

Radius: 10 arcseconds
Valid range between: 1" and 360000"

Position Columns: coord_ra, coord_dec (from the selected table on the right)

☐ Temporal ?

☐ Object ID Search ?

Output Column Selection and Constraints

Name char	constraints char	unit char	ucd char	description char	datatype char
<input type="checkbox"/> coord_dec		deg	pos.eq.dec;meta.main	Fiducial ICRS Declination of centri	double
<input type="checkbox"/> coord_ra		deg	pos.eq.ra;meta.main	Fiducial ICRS Right Ascension of cer	double
<input type="checkbox"/> deblend_nChild				Number of children this object has	int
<input type="checkbox"/> deblend_skipped				Deblender skipped this source	boolean
<input type="checkbox"/> detect_fromBlend				This source is deblended from a par	boolean
<input type="checkbox"/> detect_isDeblend				True if source has no children and is	boolean
<input type="checkbox"/> detect_isDeblend				True if source has no children and is	boolean
<input type="checkbox"/> detect_isolated				This source is not a part of a blend.	boolean
<input type="checkbox"/> detect_isPatchInn				True if source is in the inner region c	boolean
<input type="checkbox"/> detect_isPrimary				True if source has no children and is	boolean
<input type="checkbox"/> detect_isTractInne				True if source is in the inner region c	boolean
<input type="checkbox"/> footprintArea		pixel		Number of pixels in the sources deti	int
<input type="checkbox"/> g_ap03Flux		nJy		Flux within 3.0-pixel aperture. Force	double
<input type="checkbox"/> g_ap03Flux_flag				General Failure Flag. Forced on g-ba	boolean
<input type="checkbox"/> g_ap03FluxErr		nJy		Flux uncertainty within 3.0-pixel api	double
<input type="checkbox"/> g_ap06Flux		nJy		Flux within 6.0-pixel aperture. Force	double
<input type="checkbox"/> g_ap06Flux_flag				General Failure Flag. Forced on g-ba	boolean
<input type="checkbox"/> g_ap06FluxErr		nJy		Flux uncertainty within 6.0-pixel api	double
<input type="checkbox"/> g_ap09Flux		nJy		Flux within 9.0-pixel aperture. Force	double
<input type="checkbox"/> g_ap09Flux_flag				General Failure Flag. Forced on g-ba	boolean
<input type="checkbox"/> g_ap09FluxErr		nJy		Flux uncertainty within 9.0-pixel api	double
<input type="checkbox"/> g_ap12Flux		nJy		Flux within 12.0-pixel aperture. Forc	double
<input type="checkbox"/> g_ap12Flux_flag				General Failure Flag. Forced on g-ba	boolean
<input type="checkbox"/> g_ap12FluxErr		nJy		Flux uncertainty within 12.0-pixel af	double

Search Row Limit: 50000 Title: dp02_dc2_catalogs_frdf.O... Populate and edit ADQL

<https://data-dev.lsst.eu>

SIZING OF COMPUTING FOR ANALYSIS (DESC)

- Initial estimates of resources needed at CC-IN2P3 for **science analysis** in the framework of the DESC collaboration

study conducted by D. Boutigny with inputs from science coordinators

science use cases included in the study: 3x2pt + cluster analysis, simulations, synthetic source injection, supernovae studies

*goals: determining the **budget**, making **contribution statements** to the collaboration and ultimately **purchasing** and **provisioning** the equipment*

estimates include compute (mostly CPU) and disk storage

*needs of GPU equipment acknowledged but not yet fully understood:
inputs welcome*

EVENTS

- Rubin Data Facilities & Multisite Processing (Feb '25)
<https://indico.in2p3.fr/event/34450/>
- 19th International dCache Workshop (May '25)
<https://indico.desy.de/event/48191/>

QUESTIONS & COMMENTS