FASTDB Overview

"Fast Access to Survey Transients"

Rob Knop

DESC FASTDB Workshop

2025 July

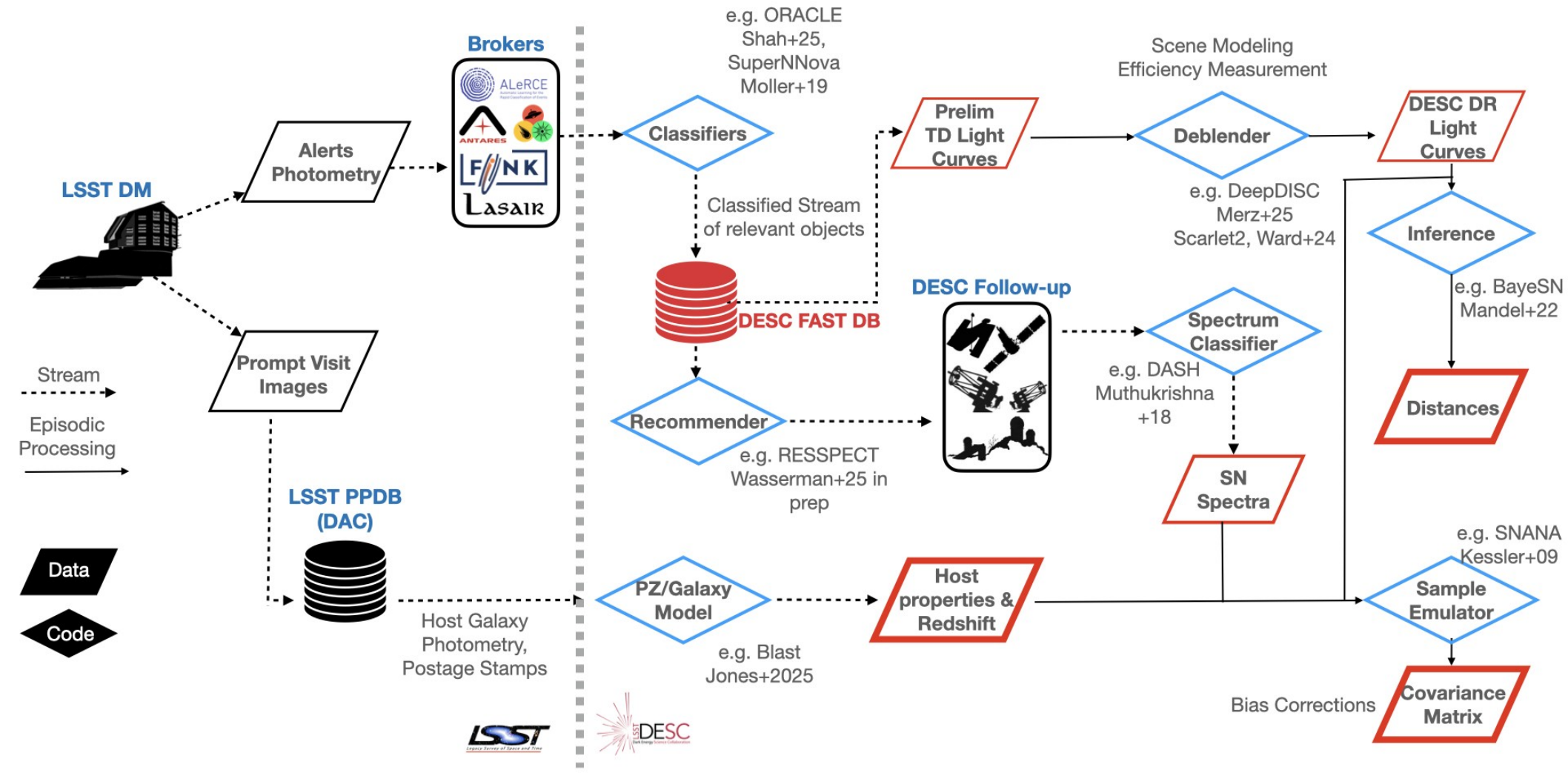
What is FASTDB?

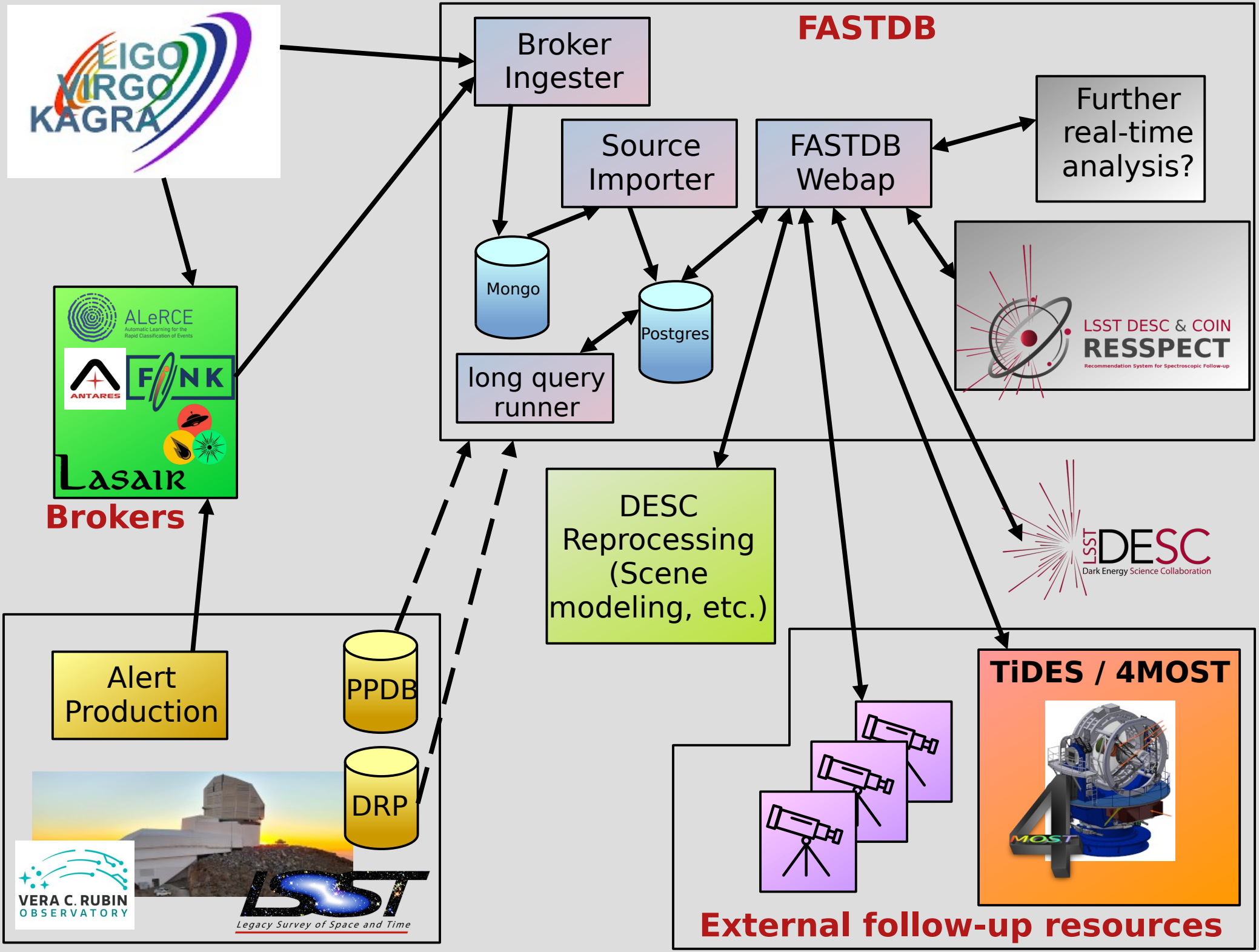
Two things...

1. (Original conception.) The definitive storage for DESC of transient data for everything needed to do a cosmological analysis. Lightcurves, redshifts, subsample selections.

2. The place where DESC keeps track of what it knows about transients from alerts filtered through brokers, and where it keeps track of what follow-up spectra it wants and has received.

A PIPELINE FOR SNIA COSMOLOGY




LIGO
VIRGO
KAGRA
FASTDB
Broker Ingester
Source Importer
FASTDB Webap
Further real-time analysis?
Mongo
Postgres
long query runner
LSST DESC & COIN
RESSPECT
Recommendation System for Spectroscopic Follow-up
ALeRCE
Automatic Learning for the Rapid Classification of Events
ANTARES
FINK
Lasair
Brokers
DESC Reprocessing (Scene modeling, etc.)
LSST DESC
Dark Energy Science Collaboration
Alert Production
PPDB
DRP
VERA C. RUBIN OBSERVATORY
LSST
Legacy Survey of Space and Time
TiDES / 4MOST
MOST
External follow-up resources

Design Constraint: Multiple Versions

“The” lightcurve of a transient is not a thing.

- The lightcurves we get from real-time LSST DIA searches via. alerts
- Forced-photometry lightcurves from alerts and PPDB.
- Lightcurves from LSST DR. (Note: object IDs will *not* match between versions!)
- Lightcurves from DESC DIA reprocessing.
- Lightcurves from DESC SM reprocessing.
- Lightcurves from DESC SM reprocessing when we improved the algorithm (or fixed bugs) and did it again.
- Lightcurves when we fixed SM reprocessing for just a few objects and need to patch them in
- Lightcurves from the *other* DESC SM reprocessing (etc.)

Expected Size

~ 2 million diaobjects of interest per year of the survey

~ 30 million diasources per year of the survey

~ 500 million diaforcedsources per year of the survey

~ 20 broker classifications per diasource (600 million per year of the survey)

~ 10^3-10^4 of transient spectra?

~ hundreds of thousands of host spectra?

- Multiply the diaobjects, diasources, and diaforcedsources by the number of releases (~10)?
- Add reprocessing for a subset.

ELAsTiCC2: 3 years, 4M diaobjects, 62M diasources, 987M diaforcedsources.

Anticipated Usage Patters — Querying

- Export parquet (?) files of “DESC data releases” with standard Ia cosmology data sets; users will use these files as is.
- DESC members asking for a big set of transient lightcurves based on some criteria for a particular analysis. (Dynamic queries.)
- People with follow-up resources asking for a list of currently active transients.
- Interactive users exploring individual candidates and lightcurves.
- Spectroscopic facilities asking for highest priority transient and/or host targets.
- Exports to SkyPortal (or another TOM).

Anticipated Usage Patters — Live Automatic

- Monitor broker alert feeds of filtered LSST alerts, with value-added broker information. (Classifications, host galaxy matching?)
- Sift through the massive quantity of broker alerts to glean out new objects, detections, and measurements (“diaobject”, “diasource”, and “diaforcedsource”).
- Provide RESSPECT with of active transients.
- Maintain RESSPECT’s lists of spectroscopic priorities.
- Maintain records of planned and taken spectra, plus information (z, classification) from cooperating follow-up facilities.
- Do backups. Failover/replication site?

Anticipated Usage Patters — Updating

- ~Yearly: import from LSST DR. (Import what? Not *all* diaobjects; how to filter?)
- Several times ~yearly: import DESC reanalyses of subsets of diaobjects.
- Tagging subsets.
- Manual (or semi-manual) import of follow-up data.
- Updating ability limited to very few users to avoid chaos.
- (Somehow!) Indicating trusted/preferred spectroscopic parameters (z, classification), trusted/preferred lightcurve classifications, trusted/preferred photo-z.