THE UNIVERSITY OF UTAH

Department of Physics & Astronomy
115 South 1400 East #201 Salt Lake City, Utah 84112-0830 (801) 581-6901 FAX (801) 581-4801

Summary of Review for Quicklook Framework
LIneA Development Team

On Thursday, May 25, we held a review of quicklook framework software system being developed by the team at Laboratório Interinstitucional de e-Astronomia (LIneA). The purpose of the review was to assess progress on the software system that will provide a view of the data quality to observers from the Dark Energy Spectroscopic Instrument (DESI). In addition, the review provided the opportunity for experienced observers from other surveys to provide feedback to the LIneA team on features that would be most useful for this software system. Those in attendance were: Klaus Honscheid (OSU), Luiz da Costa (LIneA), Kyle Dawson (Utah), Angelo Fausti-Neto (LIneA), Bob Kehoe (SMU), Elizabeth Buckley-Geer (FermiLab), Xavier Prochaska (UC-Santa Cruz), David Schlegel (LBNL), Stephen Bailey (LBNL), Connie Rockosi (UC-Santa Cruz), and the LIneA development team (Rafael, Cristiano).

In the time since the review, the group re-convened at the DESI collaboration meeting in Berkeley, in June 2017. Many of the comments and recommendations from the May review were addressed in person at the DESI collaboration meeting. In particular, the LIneA team met with the Southern Methodist University team who is responsible for the software that will process the DESI exposures. In that day of discussion, the LIneA team and the SMU team came to agreement on all of the data quality metrics that will be produced by the analysis software and displayed by the quicklook framework system.

The findings, comments, and recommendations from the May review panel follow below.

Findings

Angelo presented the material on the latest version of the quicklook framework. The development team runs in two-week sprints with a list of tasks to address, meeting regularly over the two-week interval. Cristiano leads development pertaining to processing while Rafael works on the interfaces. The code is currently at https://github.com/linea-it/qlf/.

The current system relies on a simple emulator for the Instrument Control System that sets the timing on reads to a raw data repository as if the files were posted during observations. A configuration file sets the timing, data locations, and output locations. For the ease of development, an SQLite database is being used. Each run, the existing database is destroyed and a new database is created from scratch, so as to avoid confusion between entries. Quicklook is called directly by the ICS emulator. Upon completion of quicklook for all frames for a given exposure, the files are read and ingested into the database. In the last week, it was installed and run locally by both Kyle Dawson and Bob Kehoe following the install instructions.

The interface has a panel that shows progress on the quicklook processing. This interface will eventually provide direct access to the output logs through a simple click. This automatically changes with a new exposure. On this panel, there are also placeholders to allow direct commands to be sent to the quicklook software (run, stop, reset).

There is another panel in the interface that shows QA monitoring. This page displays the results from the quicklook processing for each exposure. It is possible to scan between exposures and show the results from multiple quality assurance tests. The results from the S/N figure were shown in the presentation as an example. The high level panel displayed the scalar metrics and a pass/fail status. One level down displayed the S/N versus magnitude for each target class for a single camera. Another level down showed the location of each target on the DECaLS image viewer. Eventually, there will also be an option to see the spectra.

~3 min to process 1 camera
~15 min to process 1 exposure (30 cameras in parallel)
~30 min to process 2 exposures (30 cameras in parallel)
Quick Look execution (as is) is limited by I/O
  Large intermediate files written to the disk
~5 min to ingest QA outputs for 1 exposure
  Ingestion in sqlite3 is done in series
Should go faster with postgres database

Night Summary and observing history time series not yet implemented.

There will be a mock observing run in late summer and then again in Spring 2018. That will be ICS-centric and focused on infrastructure development. We don't need to have quicklook/quicklook framework integrated into the summer run, but it would be nice. The spring run will focus on visualization and will require advanced version of QLF.

Comments

(1) The interface to control pipeline execution should be able to retain prior exposure information. The status bar should trigger an advancement on a milestone or keyword in the QL logs, to ensure clear display of progress.

(2) We need to establish a policy for user control/access to the quicklook controls on the webpage. Can re-use the existing security scheme from the ICS or implement a new lockout scheme that provides access to one person at a time, with potential for override in special cases.

(3) quicklook software developers should re-evaluate content of intermediate files. Should establish new data format independent of offline system to compress resolution matrix

(4) consider benefits of reading from disk for ingestion relative to direct database calls from quicklok software.

(5) consider reading in fibermap files into database, as we may want to have displays of the imaging meta-data and fiber assignment results

(6) Use the collaboration meeting to assess what additional telemetry should be displayed in QLF and what should just be left to ICS. For example, ETC, fiber view camera, focal plane, GFA data.

(7) Embrace week-long visits to DESI collaborator institutions. e.g. a visit to OSU in late summer would be valuable to finalize QLF interface to the Instrument Control System.

(8) establish interface to spectra extracted by quicklook. Could record coordinates and read from flat

files when called, or to load spectra directly into database.

(9) need to converge on configuration manager. Establish roles and ownership amongst ICS, QLF, and QL for defining data locations and processing steps.

(10) consider adding a table to the high-level QA monitor page. Contents of table could include basic diagnostics or failed metrics.

(11) Set policy for external mirrors of database and quicklook flatfiles. Need to be able to deal with computer failure - a rebuild can be expensive

(12) continue to learn from DES experience for lessons on sustainability.

(13) Can push 5000-fiber catalog to DECaLS image viewer. Talk with Dustin Lang about implementation.

(14) consider separate page that allows various plots of y versus x. For example, S/N versus fiber offset from fiberview camera. Assess advantage of flexible format versus pre-defined visualizations. Consider separate expert-level for views corresponding to commissioning/SV tasks. Work with collaboration to determine how much should be displayed through QLF tool and how much should be left to users to run custom software at the command line.

(15) test access to new data as it arrives, as opposed to timed-read of data previously written to disk.

Recommendations

(1) By collaboration meeting in June, establish baseline model for QA metrics. For each metric, need to have high-level scalars that trigger pass/fail warnings according to a specifications table. Need to have outline for drill-down on each QA metric, allowing per-fiber or other deeper assessment for the scalar measurement. Have cartoon level display for each of these. There should be explanatory text at each level of QLF to inform the user of what is being shown and what specifications are being tested against.

(2) By collaboration meeting, move code from current repository to DESI hub

(3) test scalability by collaboration meeting. Requires postgres SQL so that the equivalent of 5yrs data can be loaded. Test performance at late times in survey so as to determine best database infrastructure (for example, host separate DB or separate table for current night's data). Improve efficiency of 30X run.

(4) diagnose data ingest: Determine how much data are loaded per frame. Compare performance of JSON file format relative to YAML for database I/O.

(5) have plenary talk at June collaboration meeting to introduce DESI collaboration to observing system.

(6) present cartoon-level display of metrics versus time by collaboration meeting. No need for data, only display and appearance.

On behalf of the full review panel,

Kyle Dawson
Associate Professor
Department of Physics and Astronomy
University of Utah