Designing the Science Data Quality Analysis (SDQA) Subsystem for the LSST Project

Russ Laher, Deborah Levine, Vince Mannings,
Peregrine McGehee, Jeongee Rho (IPAC),
Dick Shaw (NOAO),
and Jeff Kantor (LSST Corp.)

ADASS XVIII Meeting, Quebec City
November 2-5, 2008

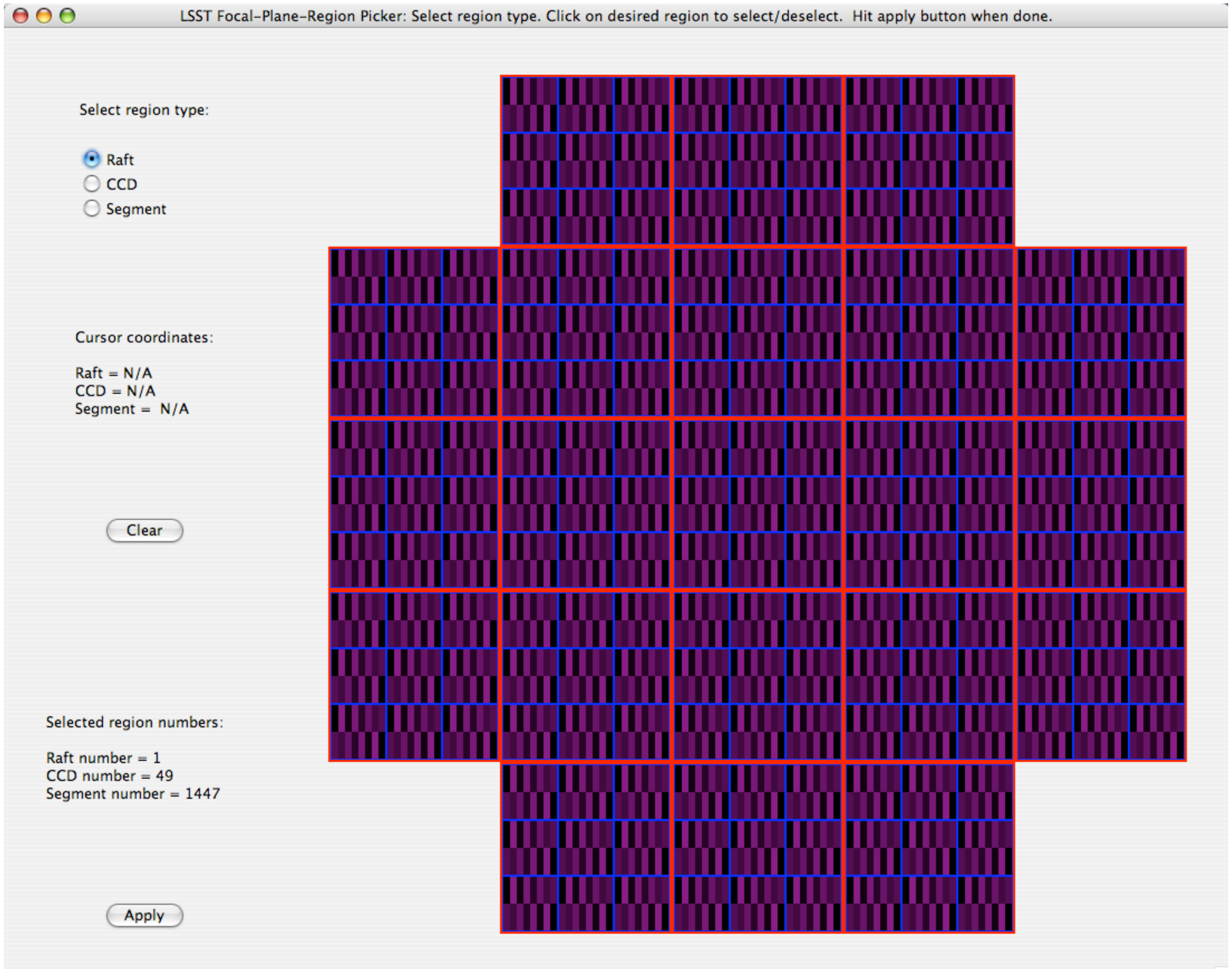
Overview

- LSST project
- SDQA
  - Conceptual design
  - Metrics
  - Data flow
  - Objectives for this year
  - Database schema for images
  - Novel approach
- Conclusions

Large Synoptic Survey Telescope

- Survey the entire visible sky deeply in multiple colors every week
- Telescope and Site
  - 8.4-m primary mirror, 9.6 square degree field-of-view
  - to see first light atop Cerro Pachón, Chile in 2014
- Camera
  - 189 science CCDs, each 4Kx4K pixels (3200 MP)
  - 3024 amplifiers/channels
  - 6 optical filters (ugrizy)

Science CCDs in Focal Plane



LSST Data Products

- Image acquisition
  - Pairs of 15-s exposures on same sky position
  - ~350,000 4Kx4K images per night!
  - 15 TB per night (55 PB in 10 years)
- Pipeline-processed images
- Catalogs built from extracted sources
  - ~50 billion astronomical objects
  - ~16 trillion source detections (high SNR)

Why do SDQA?

- Provide quantitative measure(s) of image and catalog quality to assure requirements are met
- Sanity checking over all temporal, spatial, and color scales
- Integrated summary information gives nearly instantaneous and comprehensive view of system performance
- Allow means of catching bad data and resolving problems quickly
- SDQA reports will feedback directly into the scientific research on the data

SDQA Conceptual Design

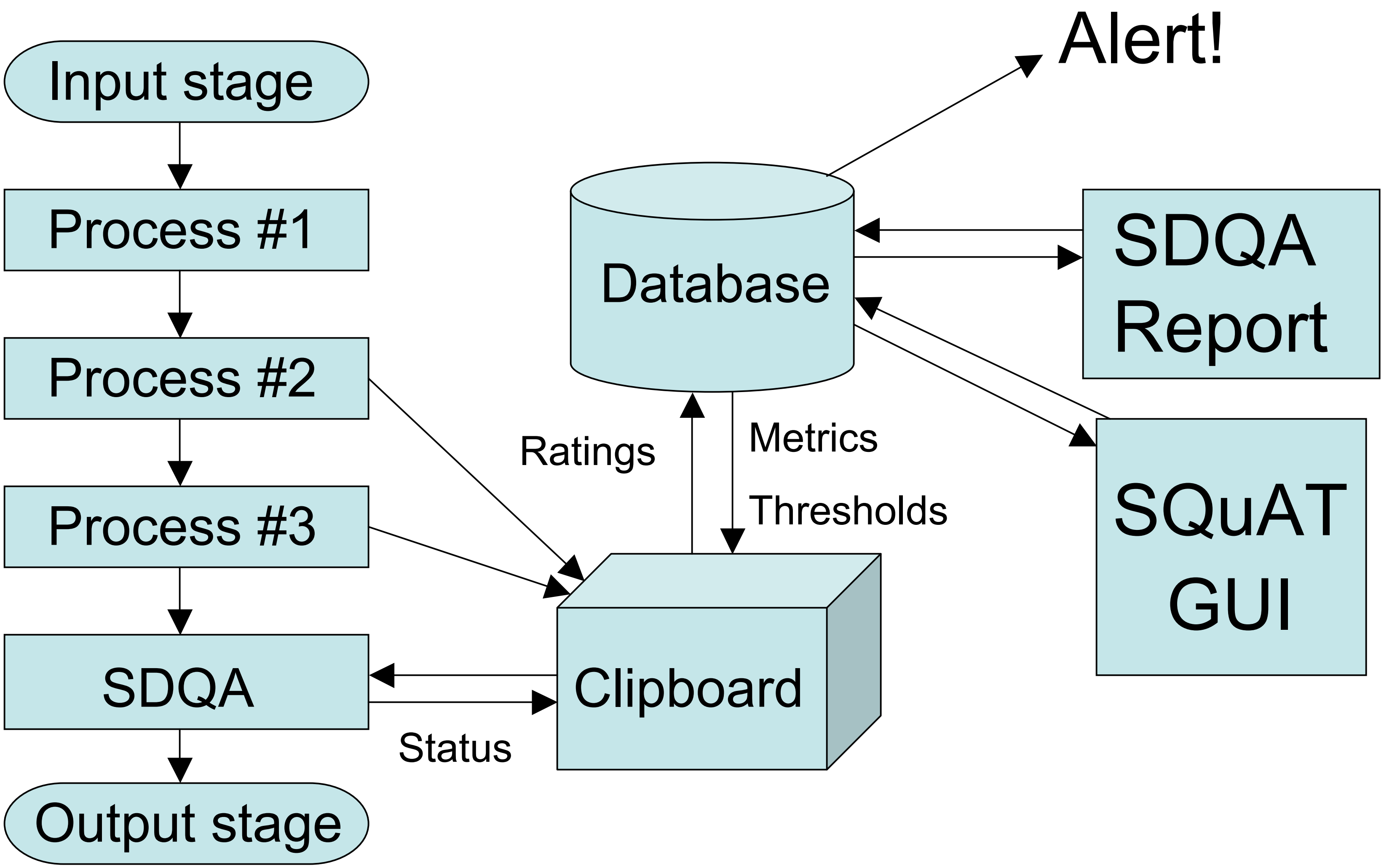
- Automated processes that examine and report on the quality of LSST science data and derived science data products from a primarily scientific perspective.
- Alert responsible personnel to problems found
- Small SDQA team will spot-check data
- Generate nightly and data-release SDQA summaries

Major SDQA Metrics

- Image quality: PSF width & shape
- Photometric accuracy, repeatability, spatial uniformity, band-to-band calibration, etc.
- Astrometric accuracy, CCD-to-CCD transform
- Image and source-detection statistics
- Metrics can depend on wavelength, zenith angle (air mass), seeing, sky brightness, position within FOV, and SNR
- Full description in Science Requirements Document and SDQA Operations Concept Document

Basic SDQA Data Flow

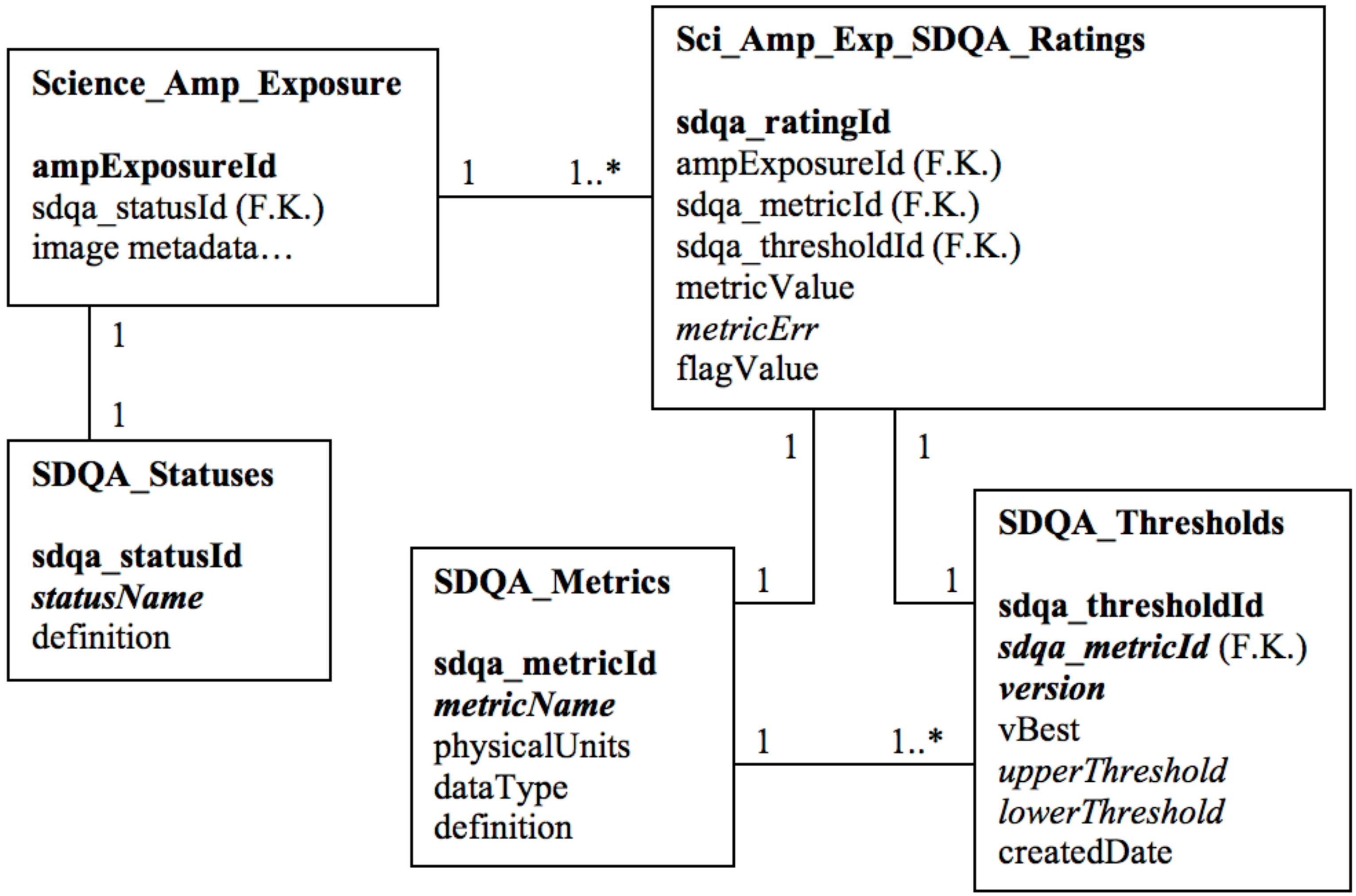
’



DC3 Objectives for SDQA

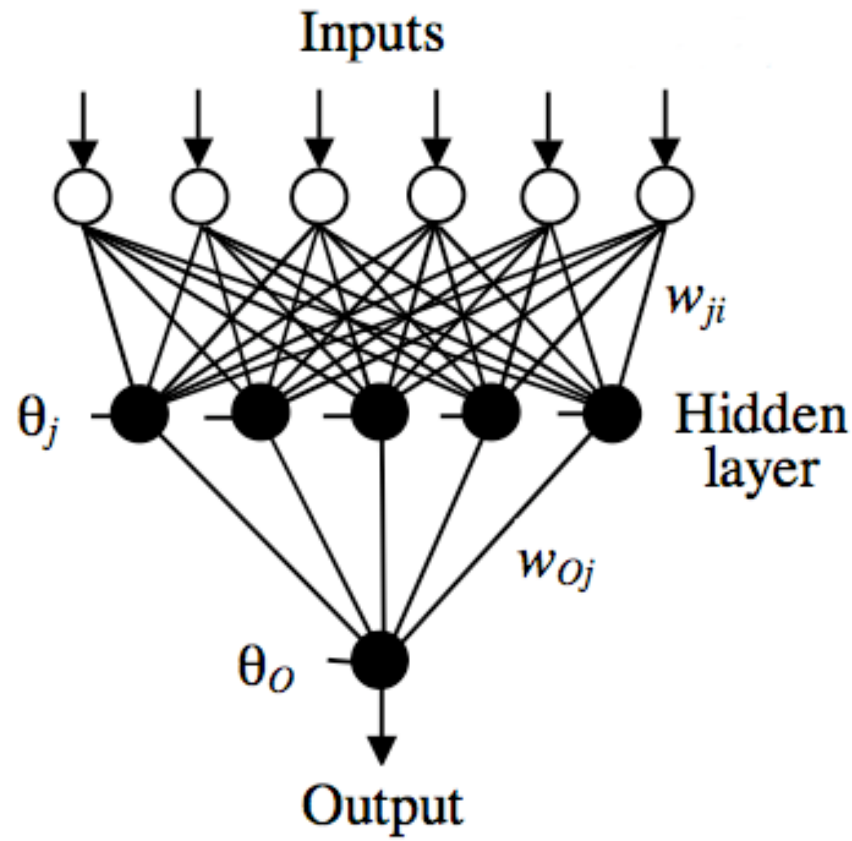
- Single SDQA metric for prototype development
  - E.g., PSF width or convolution-kernel statistics
  - Computed during pipeline processing
- Design database-persistence of SDQA ratings
- Design processes for thresholding SDQA ratings, grading images, and raising alerts
- Automated process for generating nightly SDQA summaries
- Design GUI to query the database for SDQA ratings and display them (e.g., time-series plot, histogram plot, images, movies, etc.)

Basic DB Schema for Image SDQA



Novel Approach for Automated SDQA

- Artificial neural network
- Inputs are selected SDQA ratings
- Output is a value in the [0, 1] range that can be thresholded for decisions
- Completeness & reliability to give performance



Conclusions

- Much more to be done, particularly in the area of SDQA for astronomical sources
- Software development
- Ongoing efforts to precisely define the SDQA metrics (units, range of spatial & temporal applicability, etc.)
- Exploring the questions of how to make the SDQA subsystem flexible so that future unanticipated needs can be met.