LSST Data Management in Final Design Phase

Mario Juric
LSST Data Management Project Scientist

LSST 2012 All-Hands Meeting
August 14th, 2012
LSST Data Management

- Processes the incoming stream of images that the camera system generates to **produce transient alerts** and to **archive the raw images**.

- Approximately once per year **create and archive a Data Release**, a static self-consistent collection of data products generated from all survey data taken from the date of survey initiation to the cutoff date for the Data Release.

- **Make all LSST data available** through an interface that uses community-based standards, and **facilitate user data analysis and production of user-defined data products** at Data Access Centers (DACs) and external sites.
Data Management in Final Design Phase

- Design and prototype key pieces of LSST DM infrastructure and code.

- A 6 month development cycle (a “data challenge”)
  - Software + infrastructure improvements (~4 months)
  - A mini “data release production” run (~2 months)
    • Use simulated and real data

- DC2, DC3, DC3a, DC3b, DC3b PT 1.0, DC3b PT 1.1, DC 2b PT 1.3...

- Going forward: a simpler naming scheme:
  - Every ~February: "Winter 201X" cycle
  - Every ~July: "Summer 201X" cycle
Current Status

- Prototyped(*):
  - Instrumental signature removal
  - Point source photometry
  - Extended source photometry
  - Co-addition of images
  - Moving Object Pipeline (MOPS)
  - Source-Object association
  - ...

- Under development:
  - Deblender
  - Image differencing
  - Object characterization on multi-epoch data (co-adds)
  - Fault-tolerant middleware
  - ...

(*) Does not imply “done”
2012/13 Theme: Preparing for FDR

- FDR: Final Design Review (~October 2013)

- Finalize the design of most critical (highest risk) DM code and infrastructure
  - Prototype middleware (scaling and fault-tolerance)
  - Prototype object characterization algorithms
  - Prototype image differencing
  - Finish prototyping the distributed database
  - Finalize estimates for infrastructure (TFLOPs and PBs) required for data release production

- Demonstrate we’re ready to start construction!
Summer 2012 Data Challenge
1. DM Stack Installability and Usability

2. Extended Source Photometry

3. Co-addition of Images and Detection on Co-adds

4. Forced Photometry
1. DM Stack

- DM Software Stack contains functionality useful throughout the project:
  - The “Application Framework” image manipulation classes
  - Library of measurement algorithms
  - Etc.

- It’s a rapidly maturing astronomical image processing system
  - With modifications, can process data from other cameras
  - **Being used by the Subaru Hyper Suprime-Cam Survey**
    - Talk: Paul Price, Wed/11am (breakout 1.1)

- Traditionally, it was not easy to install.
cd directory/where/lsst/stack/will/be/installed

curl -O http://dev.lsstcorp.org/pkgs/std/w12/newinstall.sh
bash newinstall.sh

source loadLSST.sh
eups distrib install --nolocks -t v5_2 lsstactive

http://dev.lsstcorp.org/trac/wiki/Installing/Summer2012

- Works on Linux (RHEL6, RHEL5, Ubuntu)
- Works on OSX 10.7 Lion
WARNING! ADVERTENCIA! AVERTISSEMENT!

THIS IS STILL NOT A FINISHED, POLISHED, READY-TO-USE END-USER PRODUCT! BEFORE DOWNLOADING, PLEASE MAKE SURE TO READ THE DM STACK FAQ:

http://dev.lsstcorp.org/trac/wiki/DM/Policy/UsingDMCode/FAQ

KEY POINTS:
- YOU’RE DOWNLOADING UNSUPPORTED, PROTOTYPE, CODE
- THIS CODE WILL NOT WORK OUT OF THE BOX FOR CAMERAS OTHER THAN LSST (AND SDSS).
- EXPECT TO HAVE TO WRITE SOME PYTHON
To assess the performance of our code, and inform future work, we used it to reprocess and analyze a subset of SDSS Stripe 82.

Southern stripe, camcol=4, -10 < R.A. < 35
57 SDSS Stripe 82 Runs
~5% of Stripe 82 data

We are releasing this data today. We do this to demonstrate the current capabilities and design of LSST prototype code, and in hope you find the reprocessed data scientifically interesting.

http://dev.lsstcorp.org/trac/wiki/DC/Summer2012
(http://goo.gl/p4Tuq)
2. Extended Source Photometry

Extended Source Photometry

Exponential (Sersic $n=1$), de Vaucouleur (Sersic $n=4$) models, and linear combination thereof (SDSS algorithms).

SDSS-level quality.

Talk: Jim Bosch, Tue/2pm (breakout 2.1)
3. SDSS Stripe 82 Co-adds

**Stripe 82 Co-adds**

Co-add of 57 Stripe 82 runs (compare: Annis et al.; 27 runs). r-band only.

Reaching flux limits 24.4 (defined as 50% completeness vs. DEEP2)

Use it to detect faint objects.

Talk: Andy Becker et al., Tue/2pm (breakout 2.1)
4. Forced Photometry

**Forced Photometry**

For every detection in the deep co-add, perform PSF photometry on individual frames (ugriz). Note that the majority of these will be below the single-frame SNR detection threshold.

Averaging those fluxes allows one to go deeper.

Left: comparison of Ivezic et al. (2004) \(w\) and \(y\) color loci; single frame vs. deep catalog.

Talk: Andy Becker et al., Tue/2pm (breakout 2.1)
Going Forward

− DM Stack:
  − We’re looking forward to discussing with other subsystems and SE how best to use DM’s software stack (where appropriate)
  − We hope that the early adopters in the community will find the software stack of interest.

− Data Release:
  − We hope science collaboration members will find the reprocessed data scientifically interesting
  − While we’ve done some preliminary QA, bugs are likely to be present both in the software and the data

− Please let us know about your experiences!
Summer 2012 DR:
Breakout 2.1, Tue/2pm (Tortolita Salon A)
what is in there? what have we done with it? what is it useful for?

DM Software Stack:
Breakout 1.1, Wed/11am (Tortolita Salon A)
what is in there? who is it for? how to use it?

Questions:
dc-help@lsst.org
To assess the performance of our code, and inform future work, we used it to reprocess and analyze a subset of SDSS Stripe 82.

Southern stripe, camcol=4, -10 < R.A. < 35
57 SDSS Stripe 82 Runs
~5% of Stripe 82 data

We are releasing this data today. We do this to demonstrate the current capabilities and design of LSST prototype code, and in hope you find the reprocessed data scientifically interesting.

(http://goo.gl/H4j6U)