The Large Synoptic Survey Telescope Data Challenges

J. Kantor and the LSST Data Management Team
LSST, 933 N Cherry Tucson, AZ 85721

ABSTRACT

The Data Management system for the LSST will have to perform near-real-time calibration and analysis of acquired images, particularly for transient detection and alert generation; annual processing of the entire dataset for precision calibration, object detection and characterization, and catalog generation; and support of user data access and analysis. Images will be acquired at roughly a 17-second cadence, with alerts generated within one minute. The ten-year survey will result in tens of petabytes of image and catalog data and will require ~250 teraflops of processing to reduce.

The LSST project is carrying out a series of Data Challenges (DC) to refine the design, evaluate the scientific and computational performance of candidate algorithms, and address the challenging scaling issues that the LSST dataset will present. This paper discusses the progress of the DCs to date and plans for future DCs.

Algorithm development must address dual requirements for the efficient use of computational resources and the accurate, reliable processing of the deep and broad survey data. The DCs incorporate both existing astronomical images and image data resulting from detailed photon-level simulations. The data is used to ensure that the system can scale to the LSST field of view and 3.2 gigapixel camera scale and meet the scientific data quality requirements. Future DCs, carried out in conjunction with the LSST Science Collaborations, are planned to deliver data products verified by computer-aided analysis and actual applications as suitable for high-quality science.

Keywords: LSST, Data Management, Data Challenge

1. INTRODUCTION

The LSST Data Management System (DMS) processes the incoming stream of images that the camera system generates to produce transient alerts and to archive the raw images, periodically creates new calibration data products that other processing functions will use, creates and archives an annual 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), and makes all LSST data available through an interface that uses community-based standards and facilitates user data analysis and production of user-defined data products with supercomputing-scale resources.[2]

In this section we give a brief description of the data flow and processing in the DMS as context for the Data Challenges. The flow of data through the LSST Data Management System (DMS) starts with the interface to the camera data acquisition on the summit at Cerro Pachon, continues through Base and Archive Centers for data processing, and on to Data Access Centers for distribution of data products to end user-accessible resources. The first major processing occurs at the Base Center in La Serena, Chile about 90 km from the summit. Here the Alert Production produces transient alerts within 60 seconds of the read-out of the second exposure in a visit (single pointing). Transient alerts are sent immediately to the Archive Center at the Petascale Computing Facility at the University of Illinois in Champaign-Urbana, and to a variety of alerting networks, including GCN, VOEventNet, and eSTAR.

The raw image data is also transferred over national and international fiber optic networks to the Archive Center where the Alert Production is run again for archival purposes. Once a year (twice in year 1), the entire image archive is re-processed in the Data Release Production to generate self-consistent data releases. A Calibration Products Production is run according to the LSST Calibration Plan to produce flats, biases, etc. required for calibration and the other Productions.

Software and Cyberinfrastructure for Astronomy, edited by Nicole M. Radziwill, Alan Bridger,
Proc. of SPIE Vol. 7740, 77401O · © 2010 SPIE · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.857253

The data is distributed to Data Access Centers for redundant storage and to support the distribution to the scientific and public communities alike. Both nightly updates and data releases, including all raw and calibrated images, are available to end users, and all data is available without any proprietary period. There are 2 project-funded Data Access Centers planned, one co-located with the Base Center and one co-located with the Archive Center.[3]

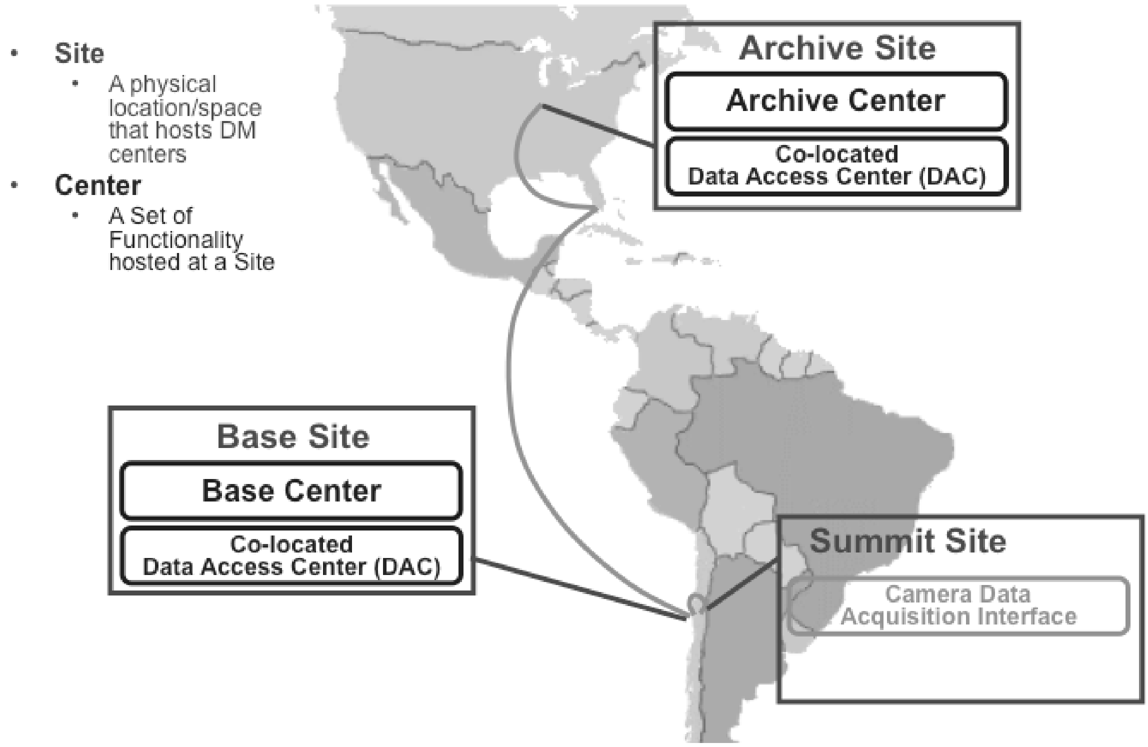

Figure 1: LSST Data Management System Data Flow and Processing Sites

2. DATA CHALLENGE GOALS

During the Research and Development Phase, the DM Team employs a nearly continual process of successive Data Challenges (DCs) in which we prototype key features of the DMS and execute those prototypes at increasing computational scales with increasingly demanding data.[4] The DCs play a crucial role in reducing DM risk during D&D, ensuring:

- Complete coverage of DMS functional scope with functioning prototypes
- Validation of scaling and computing resource estimates
- Verification of the quality of the processing and resultant data products to ensure they will meet science requirements

Data Challenge	Goals
#1 COMPLETE Jan 2006 - Oct 2006	Estimate processing load extrapolated from existing algorithms and data Create resource load simulated applications and data to represent processing workload Validate infrastructure and middleware scalability to 5% of LSST required rates
#2 COMPLETE Oct 2006 - Nov 2007	Validate Alert Production pipeline algorithms on pre-cursor data Create Application Framework and Middleware, validate by creating functioning pipelines Validate infrastructure and middleware scalability to 10% of LSST required rates
#3a COMPLETE Sep 2008 – Aug 2009	Validate Alert Production pipeline algorithms on pre-cursor data and simulated data Extend Application Framework and Middleware to support Alert Production Validate infrastructure and middleware scalability to 15% of LSST required rates
#3b IN PROGRESS May 2009 - Dec 2010	Validate Data Release Production pipeline algorithms on pre-cursor data and simulated data Expand Middleware for full Production Control & Management (orchestration) Assess end-to-end data quality with Science Collaborations Validate infrastructure and middleware reliability Validate infrastructure and middleware scalability to 20% of LSST required rates
#4 PLANNED Jul 2010 - Dec 2011	Validate database query performance Validate Calibration Products Production pipeline algorithms on pre-cursor and simulated data Validate Science User Interface and Data Quality Assessment with Science Collaborations

Table 1: Data Challenge Goals

3. DATA CHALLENGE WORK PRODUCTS

In order to prepare us for the demands of the Construction and Commissioning Phase, every DC is intentionally scoped to challenge the DM team on several fronts simultaneously:

- Computing, storage, and network infrastructure
- Astronomical data handling and algorithms
- Pipeline software framework (Middleware and Applications)
- Software development tools and procedures
- Project management and system engineering processes

Data Challenge	Work Products
#1 COMPLETE Jan 2006 - Oct 2006	3 TeraGrid nodes used to simulate data transfer: Mountain (Purdue), Base (SDSC), Archive Center (NCSA) using Storage Resource Broker (SRB) IA64 itanium 2 clusters at SDSC, NCSA, 32-bit Xeon cluster at Purdue MPI-based Pipeline Harness developed in C and python Simulated nightly processing application pipelines developed (CPU, i/o, RAM) Initial database schema designed and MySQL database configured Data ingest service developed Initial development environment configured, used throughout
#2 COMPLETE Oct 2006 - Nov 2007	10-node, 58-CPU dedicated cluster acquired and configured at NCSA Application Framework and Middleware API developed and tested Image Processing, Detection, Association pipelines developed Moving object pipeline (jointly developed with Pan-STARRS) ported to DM environment, modularized, and re-architected for nightly mode (night MOPS) Major schema upgrade and implementation in MySQL Acquired 2.5 TB pre-cursor data (CFHTLS-deep, TALCS) for testing Complete development environment configured, standardized, used throughout
#3a COMPLETE Sep 2008 – Aug 2009	Long runs on small cluster with CFHTLS CCDs (1 Mpixel) on 10 nodes, 58 cores Shorter, full focal-plane runs on larger cluster (Abe) All applications ported to 64-bit platform Significant redesign of application framework Instrument Signature Removal (ISR) pipeline, Image Characterization pipeline implemented, Visit processing implemented Science Data Quality Assessment (SDQA) framework implemented Added Pipeline “Orchestration” deploy/stage pipelines on multiple platforms
#3b IN PROGRESS May 2009 - Dec 2010	1.5 million CPU-hr, 420 TB storage allocation on TeraGrid iRODs, REDDnet configured for replicating image data 5 TB of CFHTLS data and 5 TB of simulated data prepared Instrument Signature Removal, CCD Assembly, Cosmic Ray Split, Image Characterization, Single Frame Measurement, and Source Association pipelines developed and executed in stand-alone mode Basic user interface configured using Gator, VO, and Thoth tools Non-real-time pipeline framework using Condor DAGman for workflow Data ordering and cooperation with long-term storage (tape)

Table 2: Data Challenge Work Products to date

In the current Data Challenge DC3b, the infrastructure is approaching the distribution and complexity of the operational DMS, with large-scale data transfer, data ingest, and data processing across many sites.

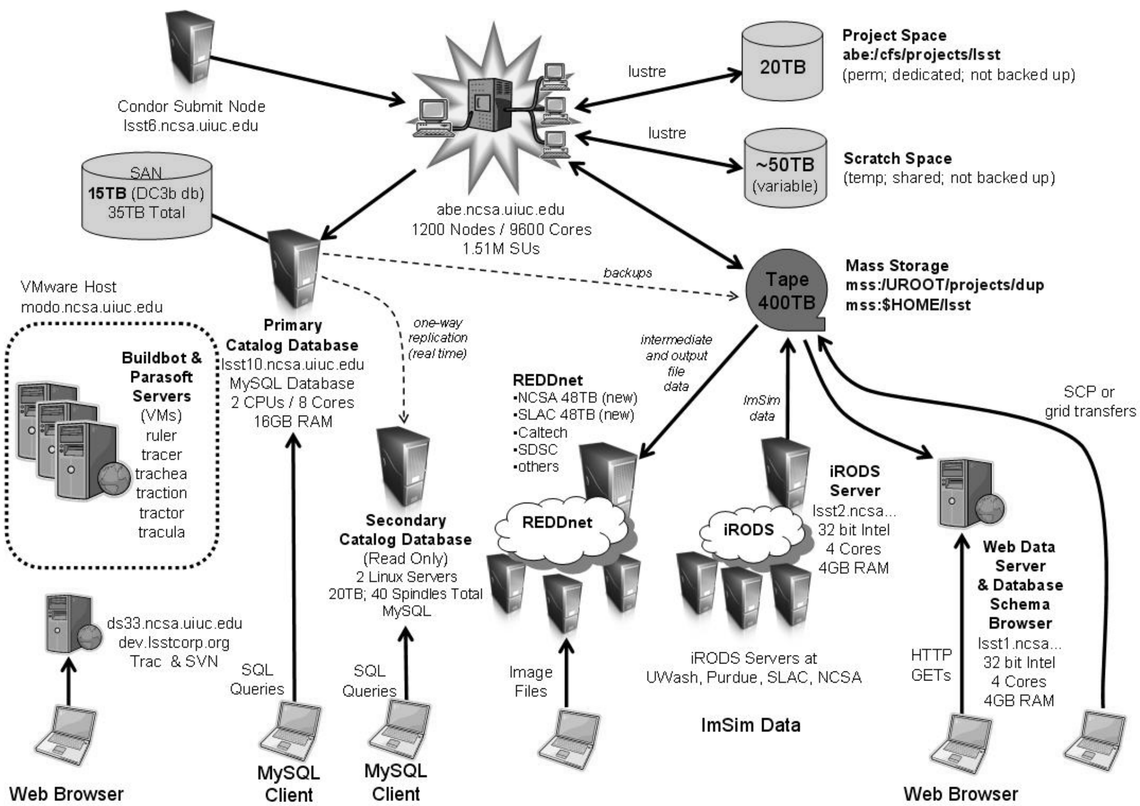

Figure 2: Data Challenge 3b Infrastructure

Due to the software complexity inherent in developing, testing, and running in production mode for18 parallel pipelines and the data volume of 52 TB of input data associated with DC3b, it is being executed in 3 stages, called Performance Tests (PTs). The Performance Tests will execute increasing numbers of pipelines and data volumes until PT3, where the full DRP will be executed on the full data.

Performance Test	Pipelines	CFHTLS Data (TB)	LSST Simulated Data (TB)
#1 May 2010 – Jun 2010	6	5	5
#2 Jul 2010 – Sep 2010	12	5	25
#3 Oct 2010 – Dec 2010	18	5	47

4. DATA CHALLENGE PERFORMANCE RESULTS

Three Data Challenges have been completed to date and one is ongoing. In all Data Challenges, the specified performance tests have been met in data transfer, data ingest, and per node data processing.

Data Challenge	Execution Results
#1 COMPLETE Jan 2006 - Oct 2006	70 megabytes/ second data transfers **(>15% of LSST transfer rate**) 192 CCDs (0.1 - 1.0 gigabytes each) runs processed with simulated pipelines across 16 nodes/32 itanium CPUs with latency and throughput of approximately 141.5 seconds **(>42% of LSST per node image processing rate**) 6.1 megabytes/ second source data ingest **(>100% of LSST required ingest rate at the Base Facility**)
#2 COMPLETE Oct 2006 - Nov 2007	61 visits (0.1 gigabytes each CCD) runs processed through all pipelines (image processing & detection, association, night MOPS) across 58 xeon CPUs with latency and throughput of approximately 257 seconds **(25% of LSST per node processing rate**) Fast nodes only (48 xeon CPUs) run processed in approximately 180 seconds **(30% of LSST per node processing rate)** Data transfer and ingest rates same as DC1
#3a COMPLETE Sep 2008 – Aug 2009	74 visits (0.1 gigabytes each CCD) runs processed through all pipelines (image processing & detection, association, night MOPS) across 58 xeon CPUs with latency and throughput of approximately 300 seconds **(15% of LSST per node processing rate**) Data transfer and ingest rates same as DC1
#3b IN PROGRESS May 2009 - Dec 2010	Performance Test 1 production runs to start in June, 2010 on NCSA Abe cluster

Table 3: Data Challenge Performance Results

In the data transfer and data ingest categories, the demonstrated performance is in excess of the Data Challenge requirements by a large margin and there are no significant risks identified at this time. In the case of the per node processing, while acceptable performance for the Data Challenges has been met for the Alert Production, there is still a 6-fold required improvement in performance that must be achieved before First Light in 2015. A combination of strategies should yield the required improvement:

- Advances in compilers and optimizers targeting multi-core architectures
- Improvements in core processing algorithms
- Porting of selected high-load codes to alternative architectures (e.g. GPU)

The performance for the Data Release Production is largely unknown at this time as it is being prototyped for the first time in the current Data Challenge 3b. Similarly, the performance for the Calibration Products Production is largely unknown at this time as it is being prototyped for the first time in Data Challenge 4.

5. DATA CHALLENGE DEVELOPMENT ENVIRONMENT

Every DC is conducted as a complete project, starting with planning/scoping, requirements definition and design using Unified Modeling Language (UML) and formal design reviews, prototyping in accordance with defined coding standards and coding reviews, and formal unit, integration, and system testing. [1]

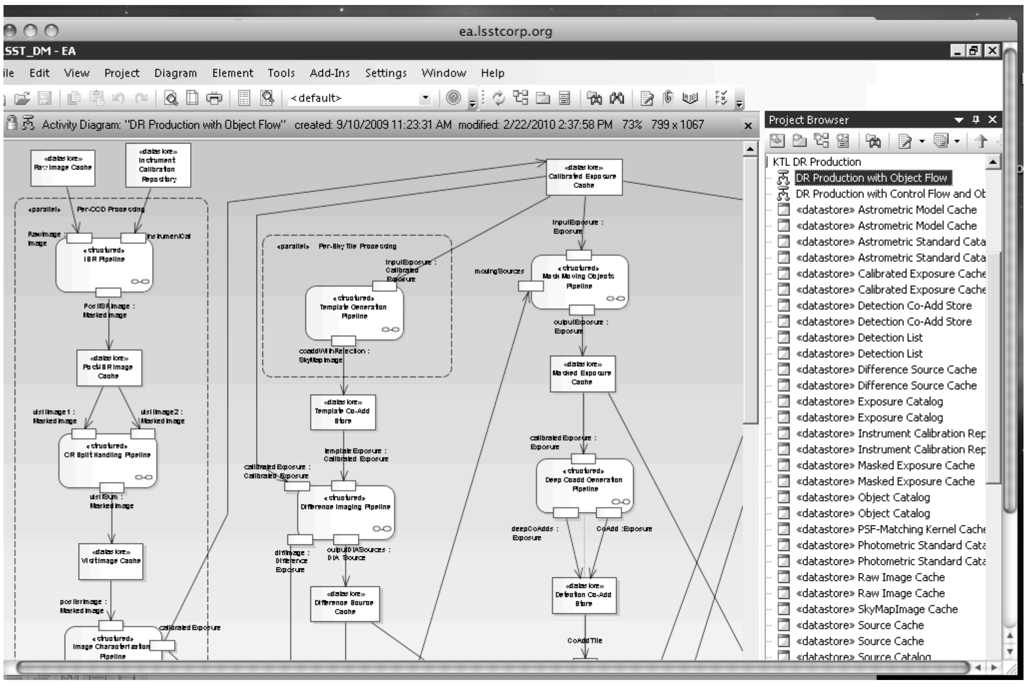

Figure 3: Data Challenge 3b UML model

Post-mortem meetings are conducted after each DC and complete reports on all DCs are available from the LSST project. Data Challenge work products and prototype software are all open source and available at the LSST Data Management trac wiki.

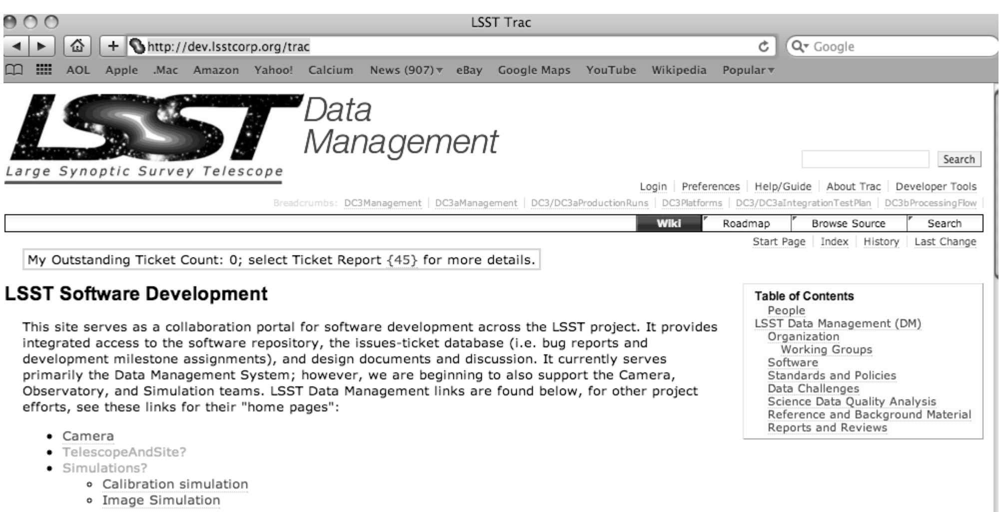

Figure 4: The LSST trac wiki

6. DATA CHALLENGES ROLE IN PREPARING FOR CONSTRUCTION

Due to the complete project nature of the Data Challenges and the process rigor being employed, the represent limited-resource "dry runs" for the incremental, 6-month software releases envisioned during the Construction Phase. All of the engineering, management, and quality assurance processes are being exercised during this phase, in order to ensure that the team can ramp up rapidly to Construction without having to undergo a steep learning curve. Approximately 75% of the eventual construction staff are now participating in the Data Challenges in a part-time mode, and we anticipate converting these staff to full-time during Construction.

7. ACKNOWLEDGEMENTS

The LSST development work is the result of efforts by the LSST collaboration of scientist, engineers, technicians, managers as well as the study work contracted to several outside entities. This team of dedicated and recognized experts in their field is what makes the LSST project a success.

LSST is a public-private partnership. Design and development activity is supported by in part the National Science Foundation under Scientific Program Order No. 9 (AST-0551161) and Scientific Program Order No. 1 (AST-0244680) through Cooperative Agreement AST-0132798. Portions of this work are supported by the Department of Energy under contract DE-AC02-76SF00515 with the Stanford Linear Accelerator Center, contract DE-AC02-98CH10886 with Brookhaven National Laboratory, and contract DE-AC52-07NA27344 with Lawrence Livermore National Laboratory. Additional funding comes from private donations, grants to universities, and in-kind support at Department of Energy laboratories and other LSSTC Institutional Members.

8. REFERENCES

1. C. F. Claver, "LSST systems engineering using SysML," SPIE Proceedings Vol. 7017-02 (2008)
2. J.P. Kantor, et al "An Overview of the LSST Data Management System," SPIE 7740-60, in press (2010).
3. J. P. Kantor et. al., "High performance astronomical data communications in the LSST data management system," SPIE Proceedings Vol. 7019-46 (2008)
4. J. P. Kantor et. al., "Managing the Evolution of the LSST Data Management System," Software and Control for Astronomy. Edited by Lewis, Hilton; Bridger, Alan. Proceedings of the SPIE, Volume 6274, pp. 62740P (2006)