UCRL-CONF-205562

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Simulation of the Future LSST Data Pipelines

Ghaleb M. Abdulla

July 27, 2004

Astronomical Data Analysis Software & Systems XIV
Pasadena, CA, United States
October 24, 2004 through October 27, 2004

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Simulation of the Future LSST Data Pipeline

Ghaleb Abdulla[τ], David Liu[ψ], Jim Garlick[τ], Sergei Nikolaev[τ], Marcus Miller[τ], Kem Cook[τ], Jim Brase[τ], Michael Franklin[ψ]

Abstract

In this paper we describe our approach to build a pipeline simulator for the future Large – scale Synoptic Survey Telescope (LSST). The simulated pipeline will be used to research and evaluate software architectures that are efficient and flexible. It will also be used to define the real-time software and hardware requirements needed to support the anticipated LSST data rates. The LSST data pipeline requirements are still being defined, however, previous surveys can provide a good source for data requirements. Our approach is to use the SuperMacho data pipeline as a prototyping tool to identify a framework for building Modular Data-Centric Pipeline (MDCP) architectures. The prototyping is done in a hierarchical fashion to help capture and define the general data attributes (schema) first. We then model other necessary components based on science and performance requirements. We use identified schemas or data attributes as a way to define a data model for LSST.

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

[τ] Lawrence Livermore National Lab
[ψ] University of California, Berkeley