

Addendum to “FY12 Science Analysis Project for LSST Data Management at Fermilab”

1. Introduction

The LSST Data Management (DM) group has a proposed list of tasks for Fermilab participation in DM for approximately one year beginning in FY11. In order to work on DM tasks at Fermilab, we need DOE approval to establish a task code for computing professionals to work on LSST DM. Kathy Turner agreed (at a Cosmological Computing meeting in Rockville, Maryland on September 14, 2011) to let Fermilab charge the effort of computing professionals to the research portion of the KA-13 B&R category up to a level of 0.5 FTE, with the understanding that Fermilab would use this effort to prepare for developing requirements and software for LSST as long as we tell Kathy what we are working on. Since this agreement did not include any additional funding from DOE, Craig Hogan and Dan Bauer were asked if computing professionals in the Computing Sector or Particle Physics Sector would be allowed to charge effort up to 0.5 FTE to LSST DM. Craig and Dan both agreed that this would be a good use of resources to help Fermilab position itself for future opportunities on LSST.

The purpose of this document is to outline the strategy we envision for Fermilab participation in LSST, and explain why the DM tasks are an important part of the strategy.

2. Strategic Vision for FY11 and FY12

The strategic vision for Fermilab participation in LSST is to identify a few key activities that can grow into significant roles during the next few years. Although Fermilab joined LSST late in the planning stage of the project, there are still a few opportunities that we can take advantage of in order to position ourselves for more significant contributions later on. With a modest level of effort in FY12 we can begin to establish a role for Fermilab in construction of the LSST camera (DOE’s primary deliverable), software development for Data Management (an NSF deliverable), a DOE Dark Energy Science Center, and Dark Energy (DE) science.

3. Goals

Goals for participation in LSST DM are described in the document “FY12 Science Analysis Project for LSST Data Management at Fermilab,” which is also referred to as the “Vision and Goals” document. The goals were defined during a one-day meeting at Fermilab on May 12, 2011. Participants at the meeting included Jeff Kantor (Project Manager for DM), members of LSST institutions working on DM (IPAC, NCSA, and SLAC), two groups from the Computing Division (CET and EAG), and Fermilab personnel involved in Open Science Grid (OSG) user support. A detailed explanation of each goal is presented in the Vision and Goals document, which is being prepared to discuss deliverables and levels of effort with Jeff Kantor and the LSST DM management team. Once we have agreement on deliverables we will prepare an estimate of required resources and the anticipated cost to LSST DM.

Five goals were identified at the meeting. With these goals in mind, Jeff Kantor has adjusted the Work Breakdown Structure (WBS) for DM to include Fermilab as a collaborating institution and to establish institutional responsibilities. Jeff has expressed his interest in having Fermilab develop the Level 3 (L3) Science Analysis Toolkit, an important part of the data analysis environment that LSST scientists will use to do their science. The toolkit will be deployed as part of the software infrastructure that will be used at LSST data access centers and the Dark Energy Science Center. Although work on the toolkit

will not begin until after FY12, it is important for us at Fermilab to use the intervening time to gain experience with LSST software. This will give us time to work with LSST science collaborations, and make it possible for us to take a leading role in developing requirements for the toolkit.

4. Strategy, and Outlook (continue here)

Rewrite this with a focus on simulations (skip the history of how we got here)

Late in 2010, DOE and NSF decided that DOE's sole deliverable for LSST construction is the camera, and that NSF is responsible for all of Data Management. During a Nov. 30, 2010 teleconference that Steve Kahn (Deputy Project Director for the LSST Project, and DOE Project Director for the camera) had with the Office of High Energy Physics in Germantown (specifically Dennis Kovar, Mike Procaro, Kathy Turner and Fred Borcherding) the topic of Data Management and funding was discussed. Dennis Kovar stated that work on DM at both SLAC and Fermilab should be funded by Jeff Kantor using NSF funding. He stated that DOE's contribution to LSST would be the camera and support for scientists working on DE science. During the teleconference they agreed that something like a "Dark Energy Science Center" would have to be created, but that support for the center would be funded out of operations costs rather than MIE costs. The main conclusion from the meeting was that DOE CD-0 (Mission Need) would only contain the camera.

This decision regarding DOE funding for LSST influenced subsequent discussions we had with Jeff Kantor regarding a proposed LSST Data Access Center (DAC) located at Fermilab. Jeff said that due to DOE's funding decision LSST management is now less interested in establishing a DAC at Fermilab. However, Jeff is still interested in having Fermilab work on DM and establishing processing capabilities at Fermilab.

Although most of the challenging DM tasks have been assigned to other institutions, we have an opportunity to contribute to DM based on our experience with SDSS, JDEM and DES as well as previous Fermilab involvement porting LSST image simulations to OSG. The main DM development task envisioned for Fermilab (already mentioned above) is the design and development of the L3 Science Analysis Toolkit. This is an ideal task for us since our experience with HEP frameworks is essential for designing and developing the toolkit, something that Jeff Kantor acknowledges. Naturally, one would think that having a good understanding of the toolkit would be of value to scientists working at Fermilab, since all scientists will need to be familiar with the toolkit to do their science with LSST data. Having local expertise with the toolkit should give scientists at Fermilab an advantage as they prepare to analyze the immense volume of data that LSST will produce.

A second task (highlighted in the Vision and Goals document) where we can contribute to LSST is to enable the use of OSG computing resources for LSST science analyses. The LSST DM group has been working closely with TeraGrid to provide LSST computing resources, and there has been limited interest in the DM group to establish computing capabilities on OSG. An important accomplishment for the OSG has been the ability to demonstrate significant computing resources for LSST image simulations. This seems to have convinced more people, but not everybody, in the DM group that OSG should be considered for the future¹. The OSG success with image simulations explains the emphasis in the Vision and Goals document on porting LSST software to OSG and on establishing a science data processing environment for DE science.

¹ As TeraGrid transitions to the XD/XSEDE project in which OSG is an identified service provider, usage of both should become more transparent.

A potentially very significant opportunity for Fermilab that is not mentioned in the Vision and Goals document is the opportunity for us to take advantage of being involved with both the LSST camera and DM. This is a unique opportunity to improve collaboration between two parts of the LSST Project (funded by DOE and NSF, respectively) that do not appear to be working together in a way that one would expect for a large project. Steve Kahn acknowledges a lack of communication between the camera and DM, and sees Fermilab as being uniquely qualified to provide a bridge between these two parts of the project. In particular, reviewers have stated that LSST should leverage software developed by DM to assist in construction and testing of LSST camera components. We anticipate that the critical software component will be the L3 Science Analysis Toolkit, because it provides the means by which a user interacts with the data processing infrastructure. With this in mind, we have established ourselves in an important role of evaluating and eventually developing key software for the camera data acquisition system (DAQ). Our experience from JDEM and NOvA with the DDS (Data Distribution Service) standard made this possible, and already puts us in a role where communication with both the camera and DM is critical. The reason is that DDS will likely be the software that is used by the DAQ to move image data from the camera to several thousand computing nodes at the upstream end of the DM data processing system. Thinking ahead to the future, our involvement in both the LSST camera and DM could open up other opportunities where experience with camera hardware (DAQ electronics and test systems) and DM software could become critically important as these two parts of the project need to be integrated. Other opportunities in LSST DM may arise in the future, but for now it is experience with LSST software and simulated data that we need in the near term to take advantage of new opportunities when they arise.

The proposal to begin to establish a significant role for Fermilab in LSST DM depends on attracting scientists to the project. Erik Gottschalk has talked to Rick Kessler from the University of Chicago, Scott Dodelson and Albert Stebbins, and all three have expressed interest in learning more about the proposal and getting involved. Rick Kessler is already a member of an LSST science collaboration as co-leader of the supernova science collaboration. However, Rick is also working on DES, and will probably not have much time to work with us on LSST. Scott Dodelson has suggested that University of Chicago postdocs and graduate students may see this as an excellent opportunity to get involved in LSST DE science at an early stage, and Erik is planning to give a presentation on LSST DE science and DM goals in the next few weeks. For members of the LSST group at Fermilab we need involvement from scientists in order to establish a few science analyses as use cases for LSST DM, as explained in the Vision and Goals document.

5. Collaborating with DES

In a follow-up meeting with Jeff Kantor on June 9, 2011 we discussed the possibility of having members of the LSST team at Fermilab collaborate with scientists working on DES in order to develop science use cases for LSST. We asked Jeff (and Tony Tyson) if we could use LSST software for DES science analyses. Both Tony and Jeff encouraged us to use LSST software for DES, and suggested that closer collaboration between the two projects would be beneficial to both.

With this in mind we believe that by working with scientists engaged in DES science analyses we can pursue long-term goals for LSST. We will explore the opportunity to use a few DES science analyses as “science drivers” to gain experience with LSST software. For example, we can look into using software developed for co-adding simulated LSST image data and see if it can be applied to DES. We can also investigate porting the LSST software that is useful for DES analyses to OSG. This would satisfy Jeff Kantor’s interest in getting LSST software ported to OSG, and give Jeff the option to highlight a more significant partnership between LSST and OSG at the upcoming NSF PDR Review, scheduled for August 2011. With regard to NSF funding for Fermilab, Jeff mentioned that he could

provide limited funding, but not before FY12.