

# Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets

## Year 1 Final Report

### 1 Overview

In order to setup the context for this final report, this section covers a brief motivation for our work and summarizes the Year 1 Proposal we originally submitted under grant number NNA06CB89H.

Large datasets are being produced at a very fast pace in the astronomy domain. In principle, these datasets are most valuable if and only if they are made available to the entire community, which may have tens to thousands of members. The astronomy community will generally want to perform various analyses on these datasets to be able to extract new science and knowledge that will both justify the investment in the original acquisition of the datasets as well as provide a building block for other scientists and communities to build upon to further the general quest for knowledge.

Grid Computing has emerged as an important new field focusing on large-scale resource sharing and high-performance orientation. The Globus Toolkit, the “de facto standard” in Grid Computing, offers us much of the needed middleware infrastructure that is required to realize large scale distributed systems. We proposed to develop a collection of Web Services-based systems that use grid computing to federate large computing and storage resources for dynamic analysis of large datasets. We proposed to build a Globus Toolkit 4 based prototype named the “AstroPortal” that would support the “stacking” analysis on the Sloan Digital Sky Survey (SDSS). The stacking analysis is the summing of multiple regions of the sky, a function that can help both identify variable sources and detect faint objects. We proposed to deploy the AstroPortal on the TeraGrid distributed infrastructure and apply the stacking function to the SDSS DR5 dataset, which comprises more than 320 million objects dispersed over 1.5 million files, a total of 9 terabytes of data.

### 2 Completed Milestones

We claimed that our work with the AstroPortal would lead to interesting and innovative research work in three main areas: 1. *resource management* (efficient task dispatch, dynamic resource provisioning); 2. *data management* (data diffusion, data-aware scheduling); and 3. *applications* (performance and scalability). The next few paragraphs describe our achievements in these categories.

**Resource Management:** To enable the rapid execution of many tasks on compute clusters (an essential component for the AstroPortal), we have developed Falcon, a Fast and Light-weight tasK executiON framework. [1] Falcon integrates (1) multi-level scheduling to separate resource acquisition (via, e.g., requests to batch schedulers) from task dispatch, and (2) a streamlined dispatcher. Falcon’s integration of multi-level scheduling and streamlined dispatchers delivers performance not provided by any other system. Microbenchmarks show that Falcon throughput (between 200 to 5000 tasks/sec depending on configuration and optimizations used) and scalability (to 54,000 executors and 1,000,000 tasks processed in just 400 seconds) are several orders of magnitude better than other systems used in production Grids. [1] Large-scale astronomy and medical applications executed under Falcon by the Swift parallel programming system [2, 3] achieve up to 90% reduction in end-to-end run time, relative to versions that execute tasks via separate scheduler submissions. [1, 2] Furthermore, the dynamic resource provisioning proposed in Falcon can allocate resources on the order of 10s of seconds across multiple Grid sites and can reduce average queue wait times by up to 95% (effectively yielding queue wait times within 3% of ideal). [4, 5] Falcon’s latest stable release v0.9 has been tested in various environments, including the TeraGrid [6], TeraPort [7], and Tier3 [8]. Falcon has also been accepted as a Globus incubator project [31].

**Data Diffusion:** We first identified that data locality is important in large scale scientific exploration [13]. We have developed a reference implementation [9, 10] of the data diffusion as part of the AstroPortal [9,

10], where the AstroPortal [11, 12] is an astronomy application that is tightly coupled to Falkon. The reference implementation performed all the basic requirements of the proposed data diffusion (caching of image data on local disk, data-aware scheduler) with some specific assumptions made due to the particular astronomy application and dataset we were using. We investigated the performance of the data-aware scheduler and found it do be sufficiently fast for medium size Grids, but did not scale well to large grids. [10] With the goal to generalize the data management implementation, we ported the reference implementation directly into Falkon for any other application that ran over Falkon to use transparently [14]. We have presented some preliminary results on the data diffusion component of Falkon at the 2007 Microsoft eScience Workshop at RENCi [15, 14].

**Applications:** We have integrated Falkon into the Karajan [3, 2] workflow engine, which in term is used by the Swift parallel programming system [2]. Thus, Karajan and Swift applications can use Falkon without modification, and hence we are able to leverage a wide range of scientific applications that are already implemented and run over Swift. We have tested several Swift-based applications from various domains including astronomy [1, 2, 17], medicine [1, 2, 17], chemistry [18], and economics [19] with varying datasets, workloads, and analysis codes. These applications had end-to-end run time reduction of up to 90%, relative to versions that execute tasks via separate scheduler submissions. Furthermore, we also developed an astronomy application named the AstroPortal [11, 12, 20, 21, 22] (the initial goal of the proposal), which has the analysis codes (i.e. image correction, pixel shifting, stacking) and application logic (i.e. mapping between sky coordinates to files) tightly integrated into Falkon. We have deployed the AstroPortal on the TeraGrid [6] distributed infrastructure and it is now online in beta testing by our collaborator's group Alex Szalay at John Hopkins University, with the goal to have it accessible in the future by the broader astronomy community. We have shown the AstroPortal to scale to 100 processors and 100K+ fine grained analysis tasks on the SDSS DR5 dataset [26] with completion times in 2~104 minutes, depending on the data source location [11, 12, 28, 30]; by contrast, a similar scale study which performed a stacking analysis of 41K+ quasars from the same SDSS dataset took 3 months time to perform. [27]

### 3 Deliverables and Dissemination of the Results

We expect to have the following deliverables upon the completion (09/30/2007) of the NASA GSRP Fellowship under Grant Number NNA06CB89H. As a result of the work funded by the NASA GSRP Fellowship, we have produced a series of software systems, papers, documents, presentations, an online portal and documentation, which can all be accesible online at the main AstroPortal web site at [http://people.cs.uchicago.edu/~iraicu/projects/Falkon/astro\\_portal.htm](http://people.cs.uchicago.edu/~iraicu/projects/Falkon/astro_portal.htm).

- **Online:**
  - Falkon: <http://people.cs.uchicago.edu/~iraicu/projects/Falkon/index.htm>
  - AstroPortal: [http://people.cs.uchicago.edu/~iraicu/projects/Falkon/astro\\_portal.htm](http://people.cs.uchicago.edu/~iraicu/projects/Falkon/astro_portal.htm)
- **Implementations:**
  - Basic Building Blocks: Falkon, a Fast and Light-weight tasK executiON framework; the Falkon project home can be found at <http://dev.globus.org/wiki/Incubator/Falkon>, including related documents, source code, and mailing lists.
  - Application: “*AstroPortal: A Science Gateway for Large-scale Astronomy Data Analysis*”; the web portal to access and use the AstroPortal can be found at <http://s8.uchicago.edu:8080/AstroPortal/index.jsp>.
- **Papers and Reports.**
  - We have produced many papers, technical reports, and proposals based on the work and results of this fellowship. They include the following documents from the reference section [1, 2, 4, 5, 9, 10, 11, 12, 13, 14, 15, 17, 20, 27, 28, 29, 30, 31, 32].
- **Documents for NASA:**

- We have three main documents specifically for NASA, namely the 1<sup>st</sup> year proposal [21], the 1<sup>st</sup> year progress report [22], and the 1<sup>st</sup> year final report [23].
- **Presentations:**
  - "*Harnessing Grid Resources with Grid Resources with Data Data-Centric Task Farms Centric Task Farms*", University of Chicago, Department of Computer Science, Dissertation Proposal, December 12th, 2007.
  - "*Falkon: a Fast and Light-weight tasK executiON framework for Grid Environments*", IEEE/ACM SuperComputing 2007, November 15th, 2007.
  - "*Accelerating Large Scale Scientific Exploration with Falkon*", IEEE/ACM SuperComputing 2007, Argonne National Laboratory Booth, November 14th, 2007.
  - "*A Data Diffusion Approach to Large Scale Scientific Exploration*", University of Chicago, CS Department, DSL Seminar, October 24th, 2007.
  - "*A Data Diffusion Approach to Large Scale Scientific Exploration*", 2007 Microsoft eScience Workshop at RENCi, October 21st, 2007.
  - "*Falkon: a Fast and Light-weight tasK executiON framework for Grid Environments*", CS Seminar, University of Chicago, CS Department, April 30th, 2007.
  - "*AstroPortal: A Science Gateway for Large-scale Astronomy Data Analysis*", IEEE/ACM SuperComputing 2006, November 2006.
  - "*Storage and Compute Resource Management via DYRE, 3DcacheGrid, and CompuStore*", University of Chicago, Department of Computer Science, Distributed Systems Lab Seminar, November 2006.
  - "*Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets*", DSL Workshop 2006, June 2006.
  - "*AstroPortal: A Science Gateway for Large-scale Astronomy Data Analysis*", TeraGrid Conference 2006, June 2006.
  - "*Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets*", University of Chicago, Department of Computer Science, Graduate Seminar, February 2006.
  - "*AstroPortal: A Science Portal to Grid Resources*", University of Chicago, Department of Computer Science, Distributed Systems Lab Seminar, January 2006.

## 4 Contributions & Conclusions

We see the dynamic analysis of large datasets to be a very important due to the ever growing datasets that need to be accessed by larger and larger communities. Attempting to address the storage and computational problems separately essentially forcing much data movement between computational resources will not scale to tomorrow's peta-scale datasets and will likely yield significant underutilization of the raw resources.

There are various fundamental research questions and directions that our work from this proposal has addressed. They center on two main areas, data and compute resource management, and how they relate to particular workloads of data analysis of large datasets.

For the data resource management, we believe that data management architectures is very important to ensure that the data management implementations scale to the required dataset sizes in the number of files, objects, and dataset disk space usage while at the same time, ensuring that data element information can be retrieved fast and efficiently. The data placement and caching strategies is critical to ensure efficient use of raw resources for a variety of workloads, datasets, data locality, and access patterns typically found in the interactive analysis of large datasets. We believe that we addressed these data resource management fundamental research questions with the implementation of Falkon and ultimately the AstroPortal.

In the realm of compute resource management, there are several important issues. Dynamic resource provisioning architectures and implementations must be carefully designed in order to offer the right abstraction while at the same time offer practical and measurable advantages over static resource provisioning. Another important issue is the scheduling of computational tasks close to the data. Again, we believe that we have addressed these compute resource management research questions with the dynamic resource provisioning work within Falcon.

Furthermore, beside the contributions we have made to the Grid Computing domain, we believe the AstroPortal system to offer contributions to the Astronomy domain as well. The AstroPortal prototype supports the “stacking” operation, the summing of image cutouts from different parts of the sky. This function can help to statistically detect objects too faint otherwise. Astronomical image collections usually cover an area of sky several times (in different wavebands, different times, etc). On the other hand, there are large differences in the sensitivities of different observations: objects detected in one band are often too faint to be seen in another survey. In such cases we still would like to see whether these objects can be detected, even in a statistical fashion. There has been a growing interest to re-project each image to a common set of pixel planes, then stacking images. The stacking improves the signal to noise, and after co-adding a large number of images, there will be a detectable signal to measure the average brightness/shape etc of these objects. While this has been done for years manually for a small number of pointing fields, performing this task on wide areas of sky in a systematic way has not yet been done. It is also expected that the detection of much fainter sources (e.g., unusual objects such as transients) can be obtained from stacked images than can be detected in any individual image. The AstroPortal gives the astronomy community a new tool to advance their research and opens doors to new opportunities on a much larger scale than ever was possible before!

## 5 References

- [1] I. Raicu, Y. Zhao, C. Dumitrescu, I. Foster, M. Wilde. “Falcon: a Fast and Light-weight task execution framework”, IEEE/ACM International Conference for High Performance Computing, Networking, Storage, and Analysis (SC07), 2007.
- [2] Y. Zhao, M. Hategan, B. Clifford, I. Foster, G. von Laszewski, I. Raicu, T. Stef-Praun, M. Wilde. “Swift: Fast, Reliable, Loosely Coupled Parallel Computation”, IEEE Workshop on Scientific Workflows 2007.
- [3] Swift Workflow System: [www.ci.uchicago.edu/swift](http://www.ci.uchicago.edu/swift)
- [4] I. Raicu, C. Dumitrescu, I. Foster. “Dynamic Resource Provisioning in Grid Environments”, TeraGrid Conference 2007.
- [5] I. Raicu, Y. Zhao, C. Dumitrescu, I. Foster, M. Wilde. “Dynamic Resource Provisioning in Grid Environments”, Technical Report, University of Chicago, 2007.
- [6] C. Catlett, et al., “TeraGrid: Analysis of Organization, System Architecture, and Middleware Enabling New Types of Applications,” HPC 2006.
- [7] TeraPort Cluster, <http://teraport.uchicago.edu/>, 2007.
- [8] Tier3 Cluster, <http://twiki.mwt2.org/bin/view/UCTier3/WebHome>, 2007.
- [9] I. Raicu, I. Foster, A. Szalay. “3DcacheGrid: Dynamic Distributed Data Cache Grid Engine”, Technical Report, University of Chicago, 2006.
- [10] I. Raicu, I. Foster. “Storage and Compute Resource Management via DYRE, 3DcacheGrid, and CompuStore”, Technical Report, University of Chicago, 2006.
- [11] I. Raicu, I. Foster, A. Szalay. “Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets”, IEEE/ACM International Conference for High Performance Computing, Networking, Storage, and Analysis (SC06), 2006.
- [12] I. Raicu, I. Foster, A. Szalay, G. Turcu. “AstroPortal: A Science Gateway for Large-scale Astronomy Data Analysis”, TeraGrid Conference 2006.

- [13] A. Szalay, J. Bunn, J. Gray, I. Foster, I. Raicu. "The Importance of Data Locality in Distributed Computing Applications", NSF Workflow Workshop 2006.
- [14] Ioan Raicu, Yong Zhao, Ian Foster, Alex Szalay. "Accelerating Large Scale Scientific Exploration through Data Diffusion," under review at IEEE International Workshop on Data-Aware Distributed Computing 2008.
- [15] I. Raicu, Y. Zhao, I. Foster, A. Szalay. "A Data Diffusion Approach to Large Scale Scientific Exploration," Microsoft eScience Workshop at RENCi 2007.
- [16] Y. Zhao, M. Hategan, B. Clifford, I. Foster, G. von Laszewski, I. Raicu, T. Stef-Praun, M. Wilde. "Swift: Fast, Reliable, Loosely Coupled Parallel Computation", IEEE Workshop on Scientific Workflows 2007.
- [17] Y. Zhao, I. Raicu, I. Foster, M. Hategan, V. Nefedova, M. Wilde. "Realizing Fast, Scalable and Reliable Scientific Computations in Grid Environments", to appear as a book chapter in Grid Computing Research Progress, Nova Publisher 2008.
- [18] V. Nefedova, et al., "Molecular Dynamics (LDRD)", 2007, <http://www.ci.uchicago.edu/wiki/bin/view/SWFT/ApplicationStatus>.
- [19] T. Stef, et al., "Moral Hazard", 2007, <http://www.ci.uchicago.edu/wiki/bin/view/SWFT/ApplicationStatus>.
- [20] I. Raicu, I. Foster. "SkyServer Web Service", Technical Report, University of Chicago, 2007.
- [21] I. Raicu, I. Foster. "Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets", NASA GSRP Proposal, Ames Research Center, NASA, February 2006 -- Award funded 10/1/06 - 9/30/07.
- [22] I. Raicu, I. Foster. "Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets: Year 1 Status and Year 2 Proposal", NASA GSRP Year 1 Progress Report and Year 2 Proposal, Ames Research Center, NASA, February 2007.
- [23] I. Raicu, I. Foster. "Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets: Year 1 Final Report", NASA GSRP Year 1 Final Report, Ames Research Center, NASA, February 2008.
- [24] I. Raicu, I. Foster, A. Szalay. "Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets", IEEE/ACM International Conference for High Performance Computing, Networking, Storage, and Analysis (SC06), 2006.
- [25] I. Raicu, I. Foster, A. Szalay, G. Turcu. "AstroPortal: A Science Gateway for Large-scale Astronomy Data Analysis", TeraGrid Conference 2006.
- [26] SDSS: Sloan Digital Sky Survey, <http://www.sdss.org/>.
- [27] R.L. White, D.J. Helfand, R.H. Becker, E. Glikman, and W. Vries. "Signals from the Noise: Image Stacking for Quasars in the FIRST Survey", The Astrophysical Journal, Volume 654, pages 99 – 114, 2007.
- [28] I. Raicu, , I. Foster. "Characterizing Storage Resources Performance in Accessing the SDSS Dataset," Technical Report, University of Chicago, 2005.
- [29] Ioan Raicu. "Harnessing Grid Resources with Data-Centric Task Farms", University of Chicago, Computer Science Department, PhD Proposal, December 2007, Chicago, Illinois.
- [30] I. Raicu, , I. Foster. "Characterizing the SDSS DR4 Dataset and the SkyServer Workloads," Technical Report, University of Chicago, 2006.
- [31] I. Raicu, Y. Zhao, C. Dumitrescu, I. Foster and M. Wilde. "Falkon: A Proposal for Project Globus Incubation", Globus Incubation Management Project, 2007.
- [32] I. Raicu, I. Foster, A. Szalay. "AstroPortal", IEEE/ACM SuperComputing 2006, Argonne National Laboratory Booth, Project Summary Handout, November 2006.