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

Ioan Raicu
Computer Science Department
The University of Chicago
iraicu@cs.uchicago.edu

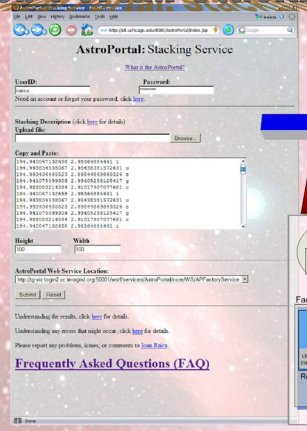
Ian Foster
Math and Computer Science Div., Argonne National Lab.
Computer Science Dept., The University of Chicago
foster@mcs.anl.gov

Alex Szalay
Dept. of Physics and Astronomy
The Johns Hopkins University
szalay@jhu.edu

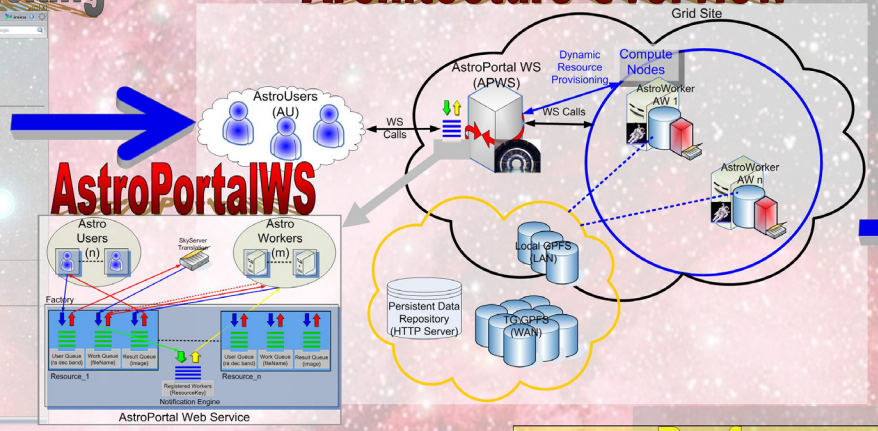
Abstract

Grid computing has emerged as an important new field focusing on large-scale resource sharing and high-performance orientation. The astronomy community has an abundance of imaging datasets at its disposal which are essentially the "crown jewels" for the astronomy community. However, these astronomy datasets are generally terabytes in size and contain hundreds of millions of objects separated into millions of files—factors that make many analyses impractical to perform on small computers. The key question we answer is: "How can we leverage Grid resources to make the analysis of large astronomy datasets a reality for the astronomy community?" Our answer is "AstroPortal," a gateway to grid resources tailored for the astronomy community. To address this question, we have developed a collection of Web Services-based systems that use grid computing to federate large computing and storage resources for dynamic analysis of large datasets. Building on the Globus Toolkit 4, we have built a prototype consisting of various systems (AstroPortal, DYRE – Dynamic Resource pool Engine, 3DcacheGrid – Dynamic Distributed Data cache for Grid applications, and CompuStore – Computational Scheduler) to enable a "stacking" analysis; the analysis sums multiple regions of the sky, a function that can help both identify variable sources and detect faint objects. We have deployed AstroPortal and the related systems on the TeraGrid distributed infrastructure and applied the stacking function to the Sloan Digital Sky Survey (SDSS), DR5, which comprises more than 320 million objects dispersed over 1.5 million files, a total of 3.5 terabytes of compressed data, with promising results. AstroPortal gives the astronomy community a new tool to advance their research and to open new doors to opportunities never before possible on such a large scale.

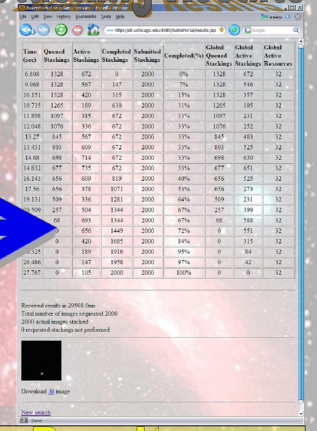
Web Portal: Stacking



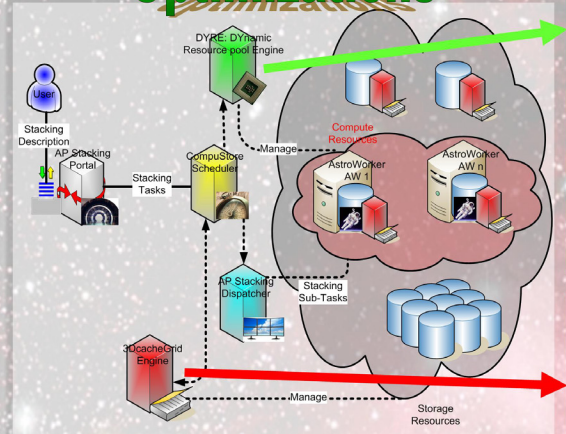
Architecture Overview



Stacking Results

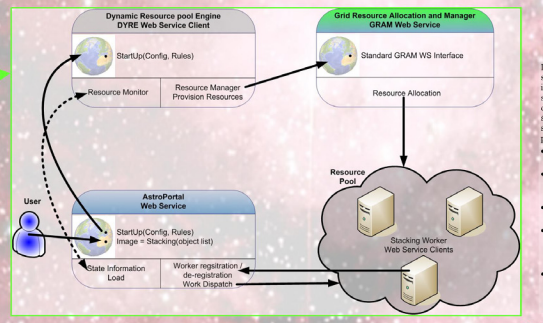


Optimizations



We propose three different systems that can interoperate with each other in order to offer a complete storage and resource management solution to enable the efficient dynamic analysis of large datasets.

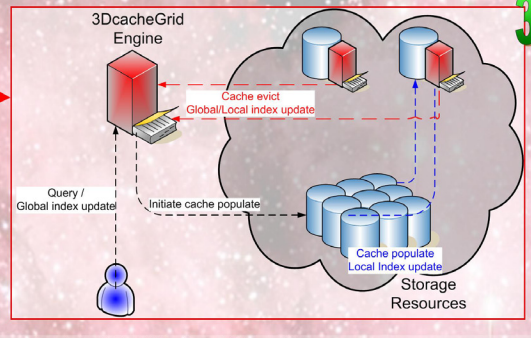
- DYRE, DYnamic Resource pool Engine, is an AstroPortal specific implementation of dynamic resource provisioning. DYRE essentially handles all the necessary tasks associated with state monitoring, resource allocation based on observed state, resource de-allocation based on observed state, and exposing relevant information to other systems
- 3DcacheGrid, Dynamic Distributed Data cache for Grid application, allows applications to achieve a good separation of concerns between their business logic and the complicated data management task of large data sets. We utilize a tiered hierarchy of storage resources to cache data in faster but smaller data storage tiers.
- CompuStore, a work scheduler that uses both the storage and compute resource management systems 3DcacheGrid and DYRE; the goal of CompuStore is to make the best scheduling decisions given some work, the available compute resources (DYRE), and the available data caches (3DcacheGrid) which are stored on the compute resources.



DYRE

DYnamic Resource pool Engine, which is an AstroPortal specific implementation of dynamic resource provisioning. DYRE essentially handles all the necessary tasks associated with state monitoring, resource allocation based on observed state, resource de-allocation based on observed state, and exposing relevant information to other systems. The main motivations behind dynamic resource provisioning are:

- Allows for finer grained resource management, including the control of priorities and usage policies
- Optimize for the grid user's perspective: reduces delays on per job scheduling by utilizing pre-reserved resources
- Give the Resource Provider the perception that resource utilization is higher than it would normally be
- Opens the possibility to customize the resource scheduler per application basis, including the use of both data resource management and compute resource management information for more efficient scheduling
- Reduced complexity to the application developer as the details of the dynamic resource provisioning are abstracted away



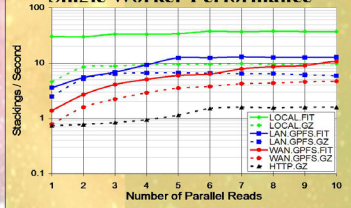
3DcacheGrid

Dynamic Distributed Data cache for Grid application allows applications to achieve a good separation of concerns between their business logic and the complicated data management task of large data sets. We propose to utilize a tiered hierarchy of storage resources to cache data in faster (i.e. higher throughput, lower latency) but smaller data storage tiers. 3DcacheGrid essentially handles the data indexing necessary for efficient data discovery and access, as well as decides the cache content (the cache eviction policy). The 3DcacheGrid engine offers efficient management for large datasets (in both number of files, size of datasets, number of storage resources used, number of replicas, and the size of the query performed), and at the same time, offer performance improvements to those applications which have data locality in their respective workloads and data access patterns through effective caching strategies (i.e. RAND, FIFO, LRU, Perfect LFU, and Hybrid Perfect LFU). The 3DcacheGrid engine gives the following advantages to applications that work with large datasets:

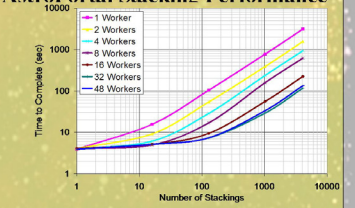
- Improved performance with higher cache hits
- Improved scalability as the data I/O will be distributed over more resources with higher cache hits
- Improved availability as cached data could be accessed without the need for the original data

Performance Results

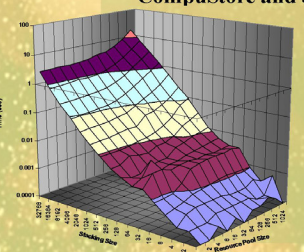
Single Worker Performance

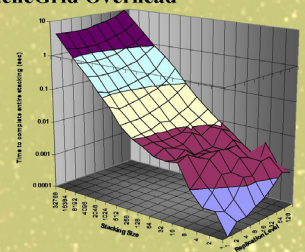


AstroPortal Stacking Performance

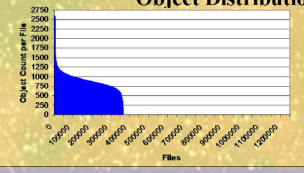


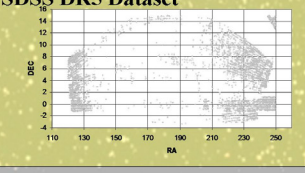
CompuStore and 3DcacheGrid Overhead





Object Distribution in SDSS DR5 Dataset





Related Material, Further Reading, and Acknowledgements

Ioan Raicu, Ian Foster, Alex Szalay, Gabriela Turcu. *"AstroPortal: A Science Gateway for Large-scale Astronomy Data Analysis"*, TeraGrid Conference 2006, June 2006.
Alex Szalay, Julian Bunn, Jim Gray, Ian Foster, Ioan Raicu. *"The Importance of Data Locality in Distributed Computing Applications"*, NSF Workflow Workshop 2006.
Ioan Raicu. "Harnessing Grid Resources to Enable the Dynamic Analysis of Large Astronomy Datasets", NASA GSRP Proposal, Ames Research Center, NASA, February 2006.

For more information, please visit http://people.cs.uchicago.edu/~iraicu/research/ and http://people.cs.uchicago.edu/~iraicu/research/AstroPortal/

This work was funded from the NSF TeraGrid grant from June 2005 - September 2006; the work is currently being funded by the NASA ARC GSRP from October 2006 - September 2007.