Science Gateways on the TeraGrid

A survey of issues for deployment of community gateway interfaces to shared high-end computing resources

Charlie Catlett, Sebastien Goasguen, Jim Marsteller, Stuart Martin, Don Middleton, Kevin J. Price, Anurag Shankar, Von Welch, Nancy Wilkins-Diehr

November 2006

Abstract

Increasingly, the scientific community has been using web portals and desktop applications to organize their work. The TeraGrid team determined that it would be important to create a set of capabilities that would allow TeraGrid services and resources to be integrated, potentially in a transparent way, with these scientific computing environments. This paper outlines the "Science Gateways" program and provides an overview of key lessons learned in developing mechanisms to allow for such integration.

Contents

1 Background

In 2004, the National Science Foundations's TeraGrid facility [i] was made available to the national academic community after a 2-year period of construction and early access. The initial facility consisted of homogeneous environment of four Itanium-1 based clusters deployed at four high-performance computing sites linked by a dedicated 40GB/s network, with an aggregate of roughly 15 Teraflops of computational power. Today the facility includes over 20 computational resources, of a wide variety of architectures, at nine sites. In aggregate, TeraGrid provides over 140 Teraflops of computational power and will grow to over 560 Teraflops in 2007. Beyond computational resources, TeraGrid includes high-performance data archives, a growing number of public data collections, and a suite of remote visualization services.

Initially, TeraGrid use consisted primarily of traditional client-server interactions, with grid functionality such as single sign-on, parallel data transfer, and remote job submission as well as a new allocation scheme that allowed users to obtain allocations "redeemable" on any TeraGrid platform rather than tied to a particular system as in the past. In 2006 the TeraGrid software architecture was augmented with web services, moving to a service oriented architecture to support new usage paradigms such as workflow and access from within applications or web portals.

Concurrent with the TeraGrid effort, many communities have developed customized interfaces to cyberinfrastructure – using emerging technologies such as web portal platforms and web services. Web access to data collections and compute capabilities are increasingly commonplace, though there are a variety of approaches and technologies used within the community, with a resulting variety of architectures and approaches to building such 'science gateways.' For example, the Protein Data Bank [ii] provides an electronic collection of structures of biological macromolecules as well as tools and resources for studying these structures. Similarly, the Network for Earthquake Engineering Simulation Cyberinfrastructure Center [iii] serves those engaged in earthquake engineering research by providing data sharing and simulation capabilities. The Network for Computational Nanotechnology [iv] provides course material, collaboration and simulation capabilities and more to those studying nanotechnology. The National Virtual Observatory [v]provides access to very large digital sky surveys and includes analysis capabilities. There are many such examples.

In 2005 the TeraGrid team initiated a "Science Gateways program" to broaden the availability of TeraGrid resources and services by providing access through community-designed interfaces. Through this program we are effectively

adapting the TeraGrid to the work environment chosen by the user rather than requiring the user to learn and adapt to the TeraGrid environment. Working with developers of community infrastructure, we couple powerful TeraGrid resources to the "back end" of familiar front-end interfaces developed by research communities. By providing a rich set of services to gateway developers, they are able to provide greater capabilities to individual scientists without requiring them to invest in learning how to adapt their work to the TeraGrid environment.

Traditionally, access to high-performance computing resources is granted to individuals, each of whom is provided with their own username and password. These individuals generally log on to a particular computing platform and submit compute jobs to a queuing system. Providing access to high end resources for a community through a shared interface involves a different, indirect, relationship between the resource provider and the end-user, and this impacts many aspects of the system ranging from authorization to usage accounting. Developing mechanisms to support this new model requires adaptation both for the resource providers and for the community interface developers. In this overview we outline some of these issues.

Perhaps most importantly, the Science Gateway model provides an effective and efficient mechanism for providing specialized services (such as high-performance computing) to much larger user communities than was possible in the past. Individual community proxies or organizations can aggregate resources for the larger group while identifying the common community services that will deliver the highest impact. This leads to a collaborative endeavor between those who provide general cyberinfrastructure services (such as TeraGrid), those who provide discipline-specific cyberinfrastructure (such as a science gateway team), and the scientists themselves. Such interactions are critical to developing and evolving a national cyberinfrastructure that is able to improve the productivity of individual scientists.

We will discuss several areas that the TeraGrid has found critical to supporting gateway interfaces to shared resources.

2 Accounting

Some gateways plan to support tens of thousands of users. Creating individual logins for all of these users on all TeraGrid resources presents significant scaling challenges.

One simplification that can streamline access to resources is a shared community account. Shared community accounts are not accounts where many users access resources by entering the same username and password. Rather, the accounts are managed by the gateway developer and used to run codes on behalf ot the scientist through the gateway interface. Individual users may register with a domain-specific front end, and access common community computational services that are provided

via the TeraGrid using a single gateway account/username. Each individual using the gateway need not have an individual account on the TeraGrid; the gateway provides a collective account.

While this level of anonymity can simplify access for end users, there are a number of additional tools required to provide the accountability and security necessary for NSF-funded resources. Developers need additional tools to trace community account resource usage back to individual gateway users and security staff need additional tools to restrict logins that provide web interfaces to supercomputing resources. We will describe these as well as additional capabilities necessary for developers in a complex multi-site environment.

In a shared environment, gateway developers will need to track use of TeraGrid resources and attribute this use to individuals logged on to the gateway. In order to correctly attribute usage to an individual gateway user, a developer must be able to determine how many CPU hours were consumed by a job launched by that user on the TeraGrid. While seemingly straightforward in nature, we found that additional capabilities were necessary to facilitate this tracking.

Many gateway developers use capabilities provided by the Globus Toolkit [vi]to access TeraGrid resources. For example, Grid Resource Allocation and Management (GRAM) [vii] is typically used to support remote job submission and monitoring., However, when a job finishes there is no straightforward way for the gateway to determine how many CPU hours the job consumed. That information is critical to attributing usage to individual users using a Science Gateway account on the TeraGrid.

To enable this functionality, the Globus team defined and created a Web Services (WS) interface and associated mechanisms to provide access to audit and accounting information associated with Grid services. This auditing system was designed to be scalable, secure, and open so that any grid service that TeraGrid deploys can follow this design.

First, the Globus Toolkit's GRAM2 (Pre-WS GRAM) and GRAM4 (WS-GRAM) services were enhanced to create audit records that are written to a database local to the GRAM services. So there will be many audit databases created everywhere TeraGrid's grid services are deployed. These GRAM audit databases and records provide a persistent link between the grid service's job id and the local resource manager's (LRM) job id.

Next, based on use cases and requirements, Open Grid Services Architecture-Data Access and Integration (OGSA-DAI) was selected to provide a service interface for TeraGrid's audit and accounting information. OGSA-DAI is a Globus Toolkit Web Services Resource Framework (WSRF) service that can create a single virtual database from two or more remote databases. A new OGSA-DAI perform document was written which defines the WS operation for returning TeraGrid Service Units (SUs) given a grid job id.

Gateway developers can then use this new interface to manage and account for their allocation on a per job basis. There are some tricky details to accurately retrieving the correct usage for a job from TeraGrid's central database that are best codified in a service/configuration and not exposed to gateway developers. OGSA-DAI can be deployed either centrally providing a TeraGrid-wide usage query service for all jobs, or deployed locally along with each GRAM4 service deployment, providing a local usage query service for local jobs. TeraGrid can decide (and change) which is most strategic.

Gateways will now be able to remotely submit jobs to TeraGrid and account for usage on a per job basis without needing to understand the details of the various local resource managers chosen by TeraGrid resource providers. We feel this accounting capability will be very useful for other projects where per-job usage information is needed. These types of enhancements are essential toward reducing the complexity for gateways to interface with TeraGrid's computational resources, as well as, allowing TeraGrid to simultaneously support an increasing number of gateways.

3 Security

3.1 Risk Mitigation

Additional risks can also arise when providing community account and web interfaces to high performance resources. The TeraGrid security working group has analyzed these risks and is developing approaches to mitigate them. Security officers at each site are alerted when a community accounts is requested and these accounts are uniquely identified in the TeraGrid central database.

TeraGrid resource sites may take independent approaches to account restriction. The current approach being suggested is a command-based restriction approach where a community account may only run certain commands, e.g. only commands in a specific directory. At least one TeraGrid site is using this approach. We believe it provides the necessary security and gateway flexibility when deploying many applications on TeraGrid.

One software package currently being developed at NCSA to suit this purpose is the Community Shell, or Commsh [ix]. Commsh allows for two methods of account restriction: The first method is an implementation of the command-based restriction described in the previous paragraph. Under this method, a configuration file is created that defines which commands (or sets of commands) a given account can execute. These commands can be specified using wildcards and regular expressions to create for a flexible command restriction framework.

The second method is change-root (or chroot) jailing. Change-root jailing effectively creates a filesystem-based "sandbox" for the account, only allowing commands to be executed from within this sandbox. An additional utility,

Chroot_jail, can be used to help construct and manage these change-root jails.

An adapter exists that can allow Commsh to operate in command-based restriction mode for GRAM job submissions. Unfortunately, this adapter does not support change-root jailing at this time.

Gateways may also want to implement shut-off mechanisms so that in the event of a problem, jobs sent to TeraGrid can be restricted from a given gateway user without shutting down the entire community account. As an intermediate step, TeraGrid will be providing a web interface for system administrators to contact developers with information regarding a problem local jobid. Long term, TeraGrid would like to provide a service to enable automatic gateway-level account shut-off.

3.2 Federated Identity Management

In the traditional mode of operation, prior to Science Gateway model, each resource or resource-providing site was responsible for the management of their users identities (a term we use to encompass a user's username, password, attributes and privileges). The science gateway model brought an out-sourcing of identity management from the resource to the gateway, shifting the responsibility for authentication and authorization from the resource to the gateway [xi].

To achieve the maximum scalability, the goal is to shift identity management all the way back to the user's home institution (i.e. their campus or place of employment) and leverage the existing identity management infrastructure. Mechanisms to achieve this based on Shibboleth, GridShib, myVocs are other technologies are currently being evaluated by TeraGrid[x].

4 Metrics and Successful Peer Review

Metrics of success are commonly requested for government funded programs. Successful gateway design will allow principal investigators to highlight gateway usage as well as science accomplishments due to the gateway. In the long term, gateways may set up a mechanism for researchers to cite the use of the gateway in publications. Success both in funding the gateway and in requesting TeraGrid resources can be traced to scientific accomplishments and a history of publications.

For example, the DOE-sponsored Earth System Grid (ESG) project [viii] includes a Metrics Service that tracks logins, file and aggregation downloads, browse and search requests, and the total volume of activity conducted via its portal. Similar to many other funded projects, this information is very useful to principal investigators and sponsors in terms of determining the overall impact of the project. As ESG begins to utilize TeraGrid resources, it will need to track

computational and data services that are delivered to it as a Science Gateway.

At the same time, we need to understand the degree to which the TeraGrid contributes to the success of a given project. In the large, this is a fairly straightforward thing to quantify, but there will need to be an interplay between the Science Gateway projects and the TeraGrid where utilization of resources are associated with impact on the science community.

5 Conclusions and Further Work: Science Gateway Primer

Gateway and portal deployment is an extremely active area, and one objective of the TeraGrid Science Gateway program is to develop standard processes and approaches necessary so that gateways can be enabled in a routine fashion. We have pursued this approach in such a way as to minimize TeraGrid-specific requirements for gateways, thus making it straightforward for a gateway project to use resources from TeraGrid as well as other grid facilities.

To this end, an important aspect of the program has involved capturing and codifying lessons and approaches in the form of a "primer." The initial version of the primer, based on initial experiences integrating TeraGrid resources with gateways, describes available resources and services for gateways as well as requirements necessary for making use of those resources.

To facilitate accurate and timely information in such a dynamic area, the primer has been made available as a Wiki (http://www.teragridforum.org/mediawiki) and will serve as the basis for Science Gateway documentation for TeraGrid resource integration. We expect contributions from a number of active gateway developers to ensure both accurate and timely information on TeraGrid resources and services, and also see the Wiki as an active repository for the many tools used in gateway development. We believe this type of community involvement will provide a rich collection of information for others and will facilitate use of high end resources in an increasing number of gateways.

The primer describes TeraGrid resources and services available to Science Gateways, requirements for using TeraGrid resources, best practices when designing a gateway and a includes a software contribution area. TeraGrid provides a variety of services to gateway developers in addition to hardware and software resources. Developers also have additional responsibilities for securing community accounts and tracking usage.

Accounting services include a variety of accounts made available to gateway developers including single user accounts, community accounts and, in the future, dynamic accounts. The primer includes links to all computing, data and visualization resources

available through TeraGrid. It describes the types of software available – the Common TeraGrid Software Stack, third party packages installed on the TeraGrid and maintained by TeraGrid staff and community software areas available to gateway developers for their own software deployment efforts. In addition, TeraGrid external relations staff are available to assist in publicizing gateway successes.

Requirements outlined in the primer include additional information to be provided when requesting community accounts, recommended audit trails for usage tracking and mechanisms to restrict problem jobs. Best practices described cover gateway planning, design, implementation, operation and metrics collection as well as desirable gateway characteristics.

The goal of the TeraGrid Science Gateway program is to provide streamlined access to developers wishing to integrate high end resources into their portals and desktop applications. The Science Gateway team will continue to develop processes and software functionality necessary to make this possible. Near term future work will address generalized Web Service interfaces to TeraGrid resources and an attribute-based authentication testbed to investigate scaling and accounting issues faced by large communities.

6 References

[i] More information about TeraGrid can be found at http://www.teragrid.org

[ii] Protein Data Bank – http://www.pdb.org

[iii]Network for Earthquake Engineering Simulation – http://it.nees.org

[iv]Nanohub – http://www.nanohub.org

[v] National Virtual Observatory – http:// www.us-nvo.org

[vi]Globus Toolkit – http://www.globus.org

[vii]Grid Resource Allocation and Management - http://www.globus.org/alliance/publications/papers.php#GRAM97

[viii]Earth System Grid – http://www.earthsystemgrid.org

[ix] http://security.ncsa.uiuc.edu/research/commaccts/

[x] Von Welch, Ian Foster, Tom Scavo, Frank Siebenlist, Charlie Catlett. Scaling TeraGrid Access: A Roadmap for Attribute-based Authorization for a Large Cyberinfrastructure (draft) http://gridshib.globus.org/tg-paper.html

[xi] Von Welch, Jim Barlow, James Basney, Doru Marcusiu, and Nancy Wilkins-Diehr. A AAAA model to support science gateways with community accounts. In Concurrency and Computation: Practice and Experience, October 2006.