GridPort 4

TACC’s GridPort software empowers users in the rapid development of highly functional grid portals

Grid computing technologies provide important capabilities such as more efficient utilization of existing resources, aggregation to have more power at once, and coordination of resources to automate workflows. However, grid computing tools can have a steep learning curve. To lower the barrier of entry for grid computing novices, simple interfaces can greatly simplify the use of these tools.

The GridPort Toolkit, a collaborative software project developed under the leadership of the Texas Advanced Computing Center (TACC), presents a consistent, streamlined set of portal interfaces for using grid technologies and services. GridPort augments these grid technologies with rich, customizable web interfaces for displaying resource information, job scheduling, and file/data management—all in a lightweight, modular, easy-to-use package.

“The GridPort team and its collaborators have made it easier for people to use vast amounts of computational power, storage capacity, and visualization and rendering capabilities to be effective in knowledge discovery,” says Dr. Jay Boisseau, director of TACC.

GridPort v4.0, released in November, is a set of portlet interfaces and services that provide access to a wide range of backend grid and information services provided by lower-level grid technologies including the Globus Toolkit, the Grid Portal Information Repository (GPIR), the Condor workload management system, and the Storage Resource Broker (SRB) file collection services.

In essence, web portals provide simple interfaces for people to use underlying resources that could be relatively complex, and grid portals enable the use of combinations of distributed resources behind the scenes—the web portal is the front end, and the software enables the portal to interact with many different resources and for those resources to interact with each other in a standard way. “The marriage is a powerful paradigm for enabling experts and non-experts alike to rapidly harness lots of computational capability,” Boisseau states.

“Historically, GridPort has been a very successful toolkit for building grid portals that help advance scientific research,” says GridPort Project Director Eric Roberts. “The release of GridPort v4.0 is an important milestone because it will be the basis for portals at The University of Texas at Austin, across the state of Texas, on the TeraGrid, and with TACC’s international partners.”

The Inception of GridPort

Necessity is the mother of invention. At the San Diego Supercomputing Center (SDSC) in the late 1990s, Boisseau, then the associate director of scientific computing, had the idea to create a web page, or a portal, that collected information about the current status of each of the systems: Which systems are available? How many jobs are waiting? Can I access high-end resources from my home computer? How busy are the systems? The portal started out of necessity as a way to answer a multitude of user questions from around the country.

The portal was dubbed the NPACI (National Partnership for Advanced Computational Infrastructure) HotPage.” Users could now go to a single location and see information on all of the systems at once. They could access static information such as user guides as well as dynamic information. Boisseau hired Mary Thomas, a graduate student at the time in computer science, to lead the HotPage project.

Boisseau and his team quickly realized that the HotPage could provide interactive capabilities for a relatively new technology called Globus, so they implemented basic interactive capabilities into the portal.
Thomas masterminded taking the underlying grid computing capabilities out of HotPage to create the initial GridPort Toolkit. A number of application portals, including the early BIRN (Biomedical Informatics Research Network) portal, were based on GridPort.

In 2001-2002, the key players in the development of the GridPort Toolkit left SDSC and went to TACC: Jay Boisseau, Mary Thomas and Maytal Dahan. Eric Roberts and Tomislav Urban were hired in 2002, with Akhil Seth completing the team in 2003. They continued working on the project developing GridPort v2.0 and v3.0.

GridPort v4.0 Differentiates Itself

“At TACC, we know that just giving someone a toolkit doesn’t mean they’ll build a portal,” Roberts says. “Along with the software, we provide a general solution that includes support, training, and a close working relationship with the GridPort developers. Our team excels at working with people, supporting them, and ensuring that they get a portal built.”

Mary Thomas, now at San Diego State University (SDSU) and still a key collaborator on GridPort, agrees that the software is unique because of its underlying approach and design philosophy. “We could build a Ferrari, and some projects need that,” she says, “but we’ve always wanted to support smaller teams to bring and keep researchers to the National Science Foundation and Department of Energy research cyberinfrastructure. Cutting edge researchers need to focus on their cutting edge science, not on cutting edge computer science.”

With GridPort v4.0, TACC is ready to make a big splash regionally as it builds user portals for the Southeastern Universities Research Association (SURA) and the Texas Internet Grid for Research and Education (TIGRE); nationally with the TeraGrid; and develops application portals through its International Partners in Advanced Computing (IPAC) and Minority Serving Institution (MSI) collaborations. At SDSU, there are plans for a bio-portal and a Department of Energy fusion grid portal.

Case in Point

In March 2005, SURA collaborated with TACC to develop, deploy and maintain a portal for the SURAgrid project (gridportal.sura.org). The portal went “live” in June 2005 and is providing single sign-on and credential management, status monitoring of grid resources, file transfer between grid nodes, job submission and management, and documentation on how to contribute resources.

More recently, TACC has expanded its role of assisting SURAgrid participants in bringing their resources into the grid. Assistance ranges from reference-level support during grid software installation to the integration of configured resources into the portal and resource monitor.

Mary Fran Yafchak, SURA’s IT program coordinator, says SURAgrid is a unique effort among current grid initiatives due to its diversity, long-term view of grid as generalized infrastructure, and the persistent peripheral objective to discover and understand grid use for and by those outside the scope of expected grid users today.

“It’s important that SURAgrid collaborate with a seasoned developer of grid portal technology to ensure an environment capable of growing from ‘entry level’ to more complex user and resource support,” Yafchak states. “TACC is providing a solid foundation from which to build and is collaborating with other SURAgrid participants, several of which are peers in portal development, to both refine and extend the services the portal will provide for SURAgrid in the future. Such services need to address interoperability and standardization across grid products, meta-scheduling, accounting, and customized access,” she concludes.

GridPort v4.0 and the TeraGrid: User Portal and “Science Gateways”

A science gateway is an interface from a specific community into the TeraGrid. Many new users will join the TeraGrid through a growing set of science gateways that use web services and grid technologies to provide access to TeraGrid resources through familiar web portal and even desktop application interfaces. These gateways will enable large numbers of researchers and educators with common types of scientific problems to use the TeraGrid in ways tailored to the unique requirements of their communities.

One important use of GridPort v4.0 is as a basis for the TeraGrid User Portal, a general purpose science gateway. The Grid Information portlet interface displays information about TeraGrid resources including...
status, load, and job queue listings. This allows users of the TeraGrid to get an “at-a-glance” view of all the resources on TeraGrid to help them make a decision about where to submit their job(s).

“There is a tremendous increase in the number of groups that are building portals, such as in partnership with the TeraGrid science gateways program, tailored to the needs of particular communities,” says Charlie Catlett of Argonne National Laboratory and the University of Chicago, director of the TeraGrid initiative and former chair of the Global Grid Forum. “Gridport represents a significant technology with which these groups can readily build gateways as well as share common components.”

Built on the newly-released GridPort v4.0, the TeraGrid User Portal will greatly simplify the way researchers learn about the resources and capabilities of the TeraGrid and gain access to and use those resources more easily, opening the TeraGrid’s power to many other researchers who may not be technical in nature or even computer literate.

Roberts says it is a crucial step because the TeraGrid User Portal will be a single place where users can see how much allocation they have left, view resource guides, submit questions to the consulting team, get information on upcoming training classes, and submit jobs to and transfer files between resources. Additionally, GridPort and the related services needed to support the TeraGrid User Portal will provide the foundation and framework for TeraGrid Science Gateways and portals.

Grid Computing Takes Flight in Latin America

Scientific progress is built on sharing information and knowledge and building on it, but information and knowledge sharing is greatly limited across geographic regions. The International Partners in Advanced Computing (IPAC) program is a great source of collaborative effort for projects such as GridPort, and TACC is already seeing the fruits of this labor through the Centro de Cálculo Científico de la Universidad de Los Andes (CeCalCULA), one of the IPAC members located in Venezuela.

Relative to the United States, Latin America is resource poor in terms of computing hardware. Grid computing offers the promise of access to increasing capabilities through resource sharing and access to increased knowledge through collaboration on the grid. Many Latin American countries realize that one of the ways to ‘catch up’ is not by trying to build big HPC machines that rival the Department of Energy (DOE), but by adopting grid computing technologies to enhance their ability to access what resources they do have, to aggregate them, and to collaborate in terms of sharing and using them.

“It makes sense for Latin American countries to adopt grid computing technologies,” Boisseau states. “By presenting access to underlying resources through the simplest of interfaces, advanced computing becomes available to the largest number of people.

CeCalCULA is the first IPAC collaborator to use GridPort v4.0. Now that this has been done, we expect to see its use in these countries increase tremendously.” Freddy Rojas, a scientific portals developer at CeCalCULA, knows from hands-on experience that the IPAC program is an ideal way to learn about grid technologies and grid portals. “The knowledge gained from this program is allowing my institute to develop grid portals based on our scientific applications and to share the knowledge with similar institutes in our country.” Rojas is developing a series of quantum chemistry portlets that will be released in 2006.

The Future of GridPort

Future releases of the GridPort Toolkit will focus on integration of new technologies such as Globus 4 (WSRF), metascheduling services, and workflow tools in addition to utilizing cutting edge web technologies such as AJAX and JavaServer Faces (JSF) to provide an even more consistent look and feel across all of the GridPort portlet interfaces.

Also in development are application portlets that provide interfaces to applications in the areas of molecular dynamics, computational chemistry, geosciences, and weather forecasting. The driver behind these interfaces is to simplify the day-to-day use of important scientific applications.

“GridPort will be a flagship software technology for TACC, and will make cyberinfrastructure easier to use and more powerful for all of our user communities and partners,” Boisseau concludes.

For more information, please visit: http://gridport.net