Final Report for DOE Project: Portal Web Services: Support of DOE SciDAC Collaboratories

PI: Mary Thomas (U. Texas), co-PI: Geoffrey Fox (Indiana Univ.)
D. Gannon , M. Pierce (Indiana University)
R. Moore (UCSD/SDSC)
D. Schissel (General Atomics)
J. Boisseau (Univ. of Texas at Austin)

Project Abstract

Grid portals provide the scientific community with familiar and simplified interfaces to the Grid and Grid services, and it is important to deploy grid portals onto the SciDAC grids and collaboratories. The goal of this project is the research, development and deployment of interoperable portal and web services that can be used on SciDAC National Collaboratory grids. This project has four primary task areas: development of portal systems; management of data collections; DOE science application integration; and development of web and grid services in support of the above activities.

1. Year 1 (2002-2003) Annual Report

The Portal Web Services project team is funded to develop and deploy interoperable portal and Grid/Web services that can be used in support of numerous SciDAC Grid projects. Our efforts concentrate on four primary tasks: the development of portal systems, management of data collections, migration of technologies from the DOE sponsored CCA project, and the development of Web services in support of the above activities. Initial deployment will be on the Fusion Grid project, although developed technologies will be of general usefulness to other SciDAC/DOE Grid efforts. As we describe in "Synergistic Activities", we are well positioned to contribute to other SciDAC efforts such as CMCS. The focus in this first year has been on the investigation of emerging grid and portal technologies and starting up the project.

1.1 Project Highlights

- Project website launched: http://www.doeportals.org
- Open Grid Computing Environments Consortium (OGCE) formed
- Portlets developed for Grid information browsing, MyProxy login, GridFTP access, and basic Grid job launching capabilities.
- GPIR Web services up and running on TeraGrid.
- SDSC releases version 2.0 of the Storage Resource Broker
- FusionGrid portal architecture conceptualized
- Fusion codes G2/TRANSP running on FusionGrid.

1.1.1 Deliverables Summary

The primary efforts for this year were focused on researching the best technology choices for the portal framework and web services. The DOE portal framework will be based on the portlets developed by members of Open Grid Computing Environments Consortium (OGCE), which includes members of this project. The initial portal framework will be based on the Jetspeed portlet architecture and the GridPort Toolkit. The Grid Portal Information Repository (GPIR), which is part of GridPort was developed this year, and includes a distributed Web services architecture, is operational on several grids.

1.1.2 Problems/Roadblocks

Funding for this program began in September, 2002. Contract negotiations between U. Texas Office of Sponsored Projects and partner sites have resulted in sites such as General Atomics not getting funded until June 2003. However, the primary thrust of the R&D efforts that this project is based on have been conducted by TACC/U. Texas and Indiana and are on track for a Summer 2003 alpha prototyping phase. Now that all sites are funded, the project management plan will include bi-weekly telcons as well as an onsite visit to General Atomics to facilitate installation and technology transfer of software systems.

1.2 Accomplishments

1.2.1 Task 1: Portal Systems

OGCE Portal

We have investigated the general capabilities for portal systems, available as a Global Grid Forum informational document, http://www.gridforum.org/Meetings/ggf7/drafts/ggf-gce-overview.pdf. From our surveys we have concluded that portal systems are generally built in the same architecture and provide many of the same capabilities, but more work needs to be done in developing interoperable portal systems. Interoperability occurs at two levels. At the service level, component services need to be described in a protocol and programming-language independent fashion. In particular, we need to investigate the role that OGSA and Web services play in developing interoperable services. At the user interface level, there needs to be a well-defined hosting environment that can implement standard portal services such as authentication and access control, and which can be extended in a well defined way to include independently developed user interfaces. We describe this in more detail in the section on Jetspeed.

OGSA consists of a set of core Grid web services defined in terms of the Open Specific activities include:

- Establishment of project website
- Research & Development
- Fusion Grid Portal
- Jetspeed Portal prototype
- SRB/MDS+ integration

Grid Service Infrastructure (OGSI) specification: An OGSI compliant Grid web service defines a subclass of web services whose ports all inherit from a standard Grid Service port. Using this port there are standard ways that a remote portal can interrogate the service to discover such things as the other port types the service implements and what operations can be made on those ports and the public internal state of the service. OGSI services can also implement a simple event subscription and notification mechanism in a standard way. OGSI also provides a mechanism for services to be group together into service collections. The simple and standard nature of OGSI makes it possible for us to build on-the-fly compilers to generate portal portlets interfaces to any OGSI compliant grid service.

The core services defined by OGSA include registries, directories and namespace binding, security, resource descriptions and resource services, reservation and scheduling, messaging and queuing, logging, accounting, data services (caches and replica managers), transaction services, policy management services, workflow management and administration services. Each of these core services is rendered as a Grid web service. (At the time of this writing, this list is probably incomplete and it is certainly not yet official.) Applications that are designed for an OGSA compliant Grid can assume that these services are available and, with the proper authorization, that they can be used.

As part of the standard GGF-GCE portal framework, we can build and distribute portlets for accessing both the client and management ports for these services. OGSA will provide a natural and easy to use building block platform for both GCE portals and applications.

Jetspeed Portal Prototype

As described above in "Portal Research and Development", portals need a consistent and extensible hosting environment that enables sharing user interface components. We have identified the Jakarta project Jetspeed, and more generally the portlet concept, as the solution to this problem. We use Jetspeed as a common baseline technology.

We have been involved in both the development of specific grid portlet instances and basic portlet types. Specific portlet components include Grid information browsing, MyProxy login, GridFTP access, and basic Grid job launching capabilities. We have also developed portlets that are capable of constructing JSP interfaces and invoking remote services from WSDL that is discovered on the fly. Additionally, we have begun prototyping the NPACI HotPage based on portlets, and we have developed portlet interfaces to the GridPort toolkit and underlying user support infrastructure. Further plans include portlet interfaces to SRB and MDS+, and other grid services such as those contained within the NMI releases (e.g. NWS).

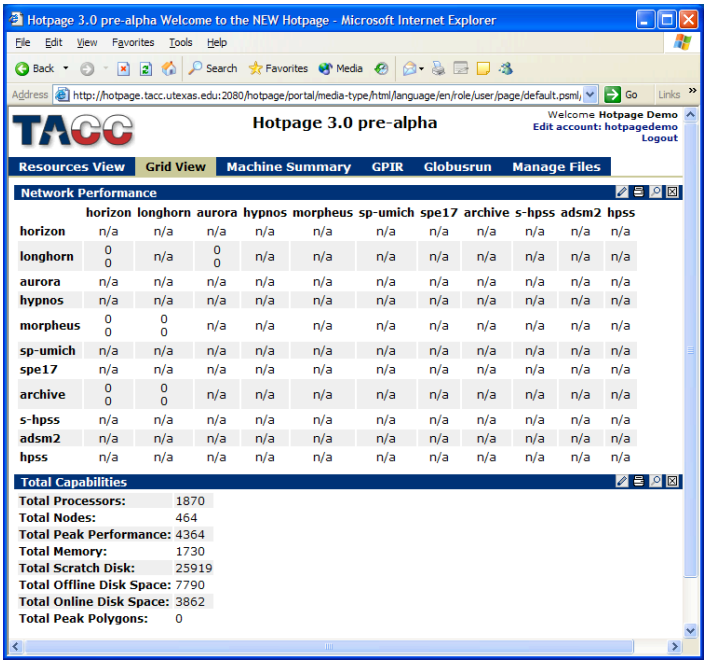


Figure 1. Screenshot of a portlet based HotPage portal interfacing to the GridPort toolkit, displaying information retrieved from the GridPortal Information Service.

In addition to specific portlet capabilities, we have developed portlet component types that can be used as proxies to remote Web servers, enabling secure connections, maintaining session states, transmitting form parameters, and supporting in-portlet navigation of remote web sites. This development is used to support legacy JSP interfaces.

1.2.2 Task 2: Data Collection Management

MDS++/SRB Integration

SDSC has released version 2.0 of the Storage Resource Broker, on Feb 19, 2003. This version will form the basis for the data grid that is implemented for the plasma physics data. SDSC staff are working on an WSDL implementation. This is proceeding in two steps:

- creation of a Java interface for the desired function
- implementation of a WSDL interface on top of the Java code

A complete set of data manipulation interfaces have been created. The next step is to implement collection management support for metadata extraction, publication, and browsing. When the metadata support is complete, we will be able to support all of the Storage Resource Broker functionality through WSDL. The initial implementation has assumed ticket-based access to the data. Once the OGSA authentication environment is defined, we will add the new types of authentication management.

Our immediate plans is to install the SRB data grid at General Atomics, and LBL. As the interface definitions for OGSA are released and stabilized, we will implement a compliant OGSA interface on the SRB. A second challenge is the management of collections that are already organized under existing data management systems. There are multiple mechanisms for providing access to information stored in separate catalogs. The choice of which mechanism to use with MDS+ will depend upon the final set of desired functions that will be used to support researchers.

1.2.3 Task 3: Grid and Web Services

While there are numerous sources of information in a Grid environment, we have found that it is necessary to provide a convenient location to store portal related data. To solve this problem, we have developed the Grid Portal Information Registry (GPIR). The aim is not to replace other information providers such as MDS, but rather to aggregate and cache grid and portal related data in support of rapid and easy portal development. Specifically, GPIR will provide the Fusion Grid Portal with its data persistence needs.

- GPIR is implemented as a set of Java Web Services, one to handle the input of GPIR data (Ingester WS) and another to facilitate the querying of that data (Query WS)
- The Ingester WS accepts or "ingests" several types of XML documents and stores them in a relational database (currently MySQL). These documents are created by a variety of means, including Java Clients that exist on the resources themselves, http "web scraping" of machine-specific flat-file formats, and queries of additional information providers such and MDS, GMS (Grid Monitor Service), and NWS (Network Weather Service).
- Once the data resides inside of GPIR, it can be queried via the Query Web Service which uses the same XML resources used by the Ingester, in addition to some Query specific documents that can return XML such as Machine Summary data

Existing information services already deployed on the Fusion Grid, such as the Globus MDS and the Fusion Grid Monitor can be included into the system.

The team has also developed a number of basic Web services for remote command execution, batch script generation, file uploading/downloading, and cross-loading between two backend hosts. Application metadata web services have been developed that simplify the deployment of existing applications into the portal system, and metadata services have been developed to store general name/value pairs in structured contexts. These are described in http://grids.ucs.indiana.edu/ptliupages/publications/GatewayICCS03v5.01.pdf.

Web service security is another area that is currently open for development. In particular, we have investigated the use of SAML to provide secure SOAP messaging, with implementations built around Kerberos. This system is described at

http://grids.ucs.indiana.edu/ptliupages/publications/GatewaySAM03.pdf.

GCE Shell

The Fusion Grid includes users who desire a more flexible and open-ended working environment than is available from typical Web-based portal systems. These users are comfortable with command line interfaces and UNIX shell-style programming. We are therefore developing a GCE Shell programming environment that is intended to host Grid and Web service client applications. The basic requirements for this system are described in the Global Grid Forum informational document, http://www.gridforum.org/Meetings/ggf7/drafts/GGFWPGCEShell.pdf, by Pierce and Fox. The design of

the prototype system is described in the paper by M. Nacar, Pierce, and Fox, available from http://grids.ucs.indiana.edu/ptliupages/publications/GCEShell_camera.pdf.

In summary, the GCE shell architecture is being implemented to initially host arbitrary Web service commands. We are building an initial command set that mimics basic UNIX shell capabilities (such as copy, move, delete, list, and so on) over a distributed environment. Advanced capabilities will include a) support for complex command line arguments (through "gce-pipes" and "gce-redirects") that will support simplistic workflow; and b) support for shell-like scripting.

1.2.4 Task 4: Applications - FusionGrid Portal

A proposed prototype of a Fusion Grid portal was conceptualized and a diagram is listed below. Deployment of this portal onto the Fusion grid was delayed but will be deployed during Summer of 2003. The prototype will be used to gather user input and requirements and will be used to generate a working portal and demonstration for SC03. Using the Globus Toolkit 2 as a baseline technology, we have been able to independently establish a fusion grid infrastructure and develop general purpose user interface components that can interact with this infrastructure. The general purpose portlets will be used to develop a basic SciDAC Fusion Portal in Summer 2003.

The National Fusion Grid project (Schissel, PI) has installed the Globus Toolkit and has deployed two applications on the fusion grid to interact with data from San Diego Tokomak. The National Fusion Grid (NFG) is a SciDAC Collaboratory Pilot project to create and deploy collaborative software tools throughout the magnetic fusion research community. The goal of the project is to advance scientific understanding and innovation in magnetic fusion research by enabling more efficient use of existing experimental facilities and more effective integration of experiment, theory, and modeling. NFG has 2 applications running on it:

- GS2 is a physics application, developed to study low-frequency turbulence in magnetized plasma. It is presently being tested on FusionGrid as the second code to be made publicly available. GS2 is primarily funded by the USDOE through the SciDAC Project.
- TRANSP, a power balance and simulation code, was the first code released in November 2002. Work on GS2 has been modeled after the TRANSP work and that has greatly reduced the time required to bring GS2 onto FusionGrid. GS2 is running on a Linux cluster at the University of Maryland and is being offered to FusionGrid via the work of Bill Dorland.

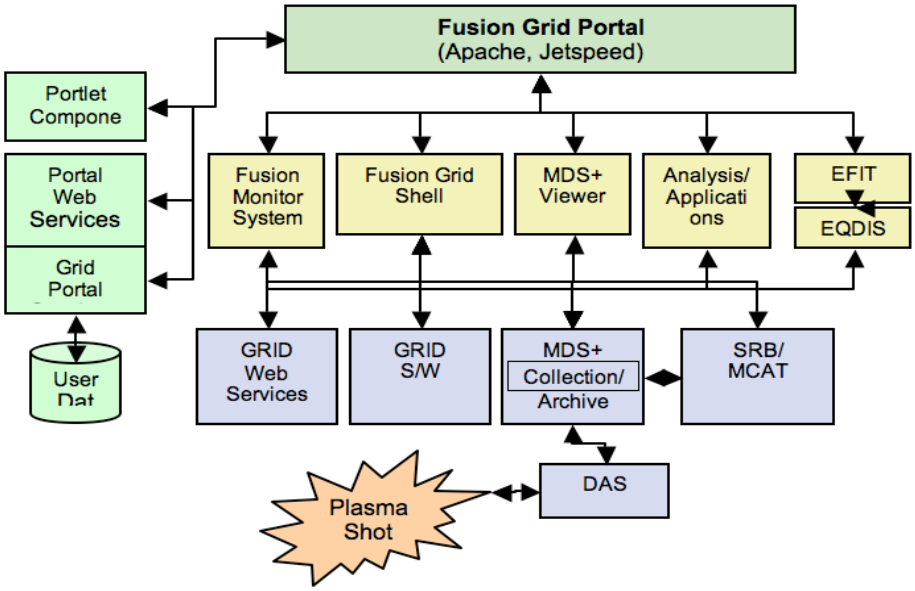


Figure 1: Prototype architecture for the Fusion Grid Portal Infrastructure

Team members involved in portlet development (TACC and IU) have concurrently developed general purpose portlets and services to interface with Globus technologies. These include grid information browsing portlets, MyProxy portlets, gridFTP portlets, and basic job submission capabilities, as well as a portlet interface to GridPort toolkit allowing for secure, credentialed logins. See www.fusiongrid.org for more details.

1.3 Collaboration, Events, Outreach

Several of the team members (Thomas, Fox, Gannon, Pierce) are participants in the Open Grid Computing Environments Consortium. Other team members include representatives from NCSA, Argonne National Labs, and University of Michigan. The aim of this consortium is to provide a high quality, standards-compliant portal framework and portlet components to for building Grid computing environments. This consortium has direct involvement in both the Fusion Grid and CMCS projects, so development work for the DOE is well-placed to be used by other SciDAC projects. Other collaborations include:

- Global Grid Forum: Grid Computing Environments Research Group:
  - Workshop on portal architectures: Tokyo, Feb 2003
- PACI and TeraGrid:
  - NPACI -SDSC, U. Mich (NEESGrid);
  - Alliance: NCSA, Indiana;
- Globus: Argonne, ISI
- Gov: NASA IPG, DoD PET
- Texas Internet Grid for Research and Education (TIGRE): U. Texas, U. Houston, Texas A&M, Texas Tech, Rice
- Commercial: Platform Computing, United Devices, IBM

1.4 Selected Publications

- M. Thomas, J. Boisseau, M. Dahan, C. Mills, S. Mock, K. Mueller. "Development of NPACI Grid Application Portals and Portal Web Services," accepted for publication in Journal of Cluster Computing, 6(3), 2003.
- M. Thomas, J. Boisseau. Building Grid Computing Portals: The NPACI Grid Portal Toolkit. To be published in "Grid Computing: Making the Global Infrastructure a Reality " edited by Fran Berman, Geoffrey Fox and Tony Hey, December 2002 by Wiley and Sons.
- S. Mock, K. Mueller, M. Pierce, C. Youn, G. Fox, M. Thomas. A Batch Script Generator Web Service for Computational Portals. Proceedings of the The 2002 International Multiconference in Computer Science, June 24-27, 2002.
- Choonhan Youn, Marlon Pierce, and Geoffrey Fox Building Problem Solving Environments with Application Web Service Toolkits ICCS03 Australia June 2003.
- Mehmet Nacar, Marlon Pierce and Geoffrey Fox Designing a Grid Computing Environment Shell Engine in Proceedings of the 2003 International Conference on Internet Computing, Las Vegas June 2003
- Geoffrey Fox, Dennis Gannon, Mary Thomas with A Summary of Grid Computing Environments July 2002 as a summary of a special issue of Concurrency and Computation: Practice and Experience
- Marlon Pierce, Choonhan Youn, Ozgur Balsoy, Geoffrey Fox, Steve Mock, and Kurt Mueller Interoperable Web Services for Computational Portals. SC02 November 2002

1.5 Project Management

Team Meetings

The project holds regular telcons on a bi-weekly basis. In addition, the team met at several large meetings held during the year including Supercomputing and the DOE Plasma Physics meeting.

Project Web Site

In order to simplify management of information among a distributed research team and to provide information and downloads to current and future SciDAC collaborators, we have developed an information portal web site, www.doeportals.org , which is scheduled to be on line July 1, 2003. The initial site will contain project information, news, links to publications and presentation material, and links to relevant DOE portal sites. The portal will also serve as a software repository and download site. Future versions may integrate newsgroup systems such as those developed by Indiana University, available from www.xmlnuggets.org.

---

2. Year 2 (2003-2004) Annual Report

2.1 Project Highlights

- DOE Portals All Hands Meeting held Feb 4-5, 2004 at SDSC.
- GridPort Toolkit v3 beta released Feb, 2004.
- GSI integration into JARGON (Java version of SRB) complete
- OGCE NMI Released May, 2004
- OGCE Portal Infrastructure and Software installed and running on Fusion Grid
- Fusion Portal Development Initiated
- DOE Demo Portal project launched
- Initiated DOE Portals Consortium Project
- DOE Portals project contributes to new NSF Middleware Initiative project, the Grid Portal Development program, funded to develop de-facto standard for grid portals (official announcement);
- Successfully demonstrated prototype MDSPlus-SRB system SRB MDSPlus information open and read capabilities;
- Demonstration Fusion portal presented at SC03 based on the alpha NMI/OGCE portals toolkit.

2.1.1 Deliverables Summary

The four partner sites became fully funded and have been focusing on 2 strategic areas: defining architecture and technologies for portals (TACC/IU) and integration of SRB and MDSPlus (SDSC/GA). In 2003, TACC and Indiana won an NSF Middleware Initiative for Grid Portal Development project (led by M. Pierce, IU; Thomas and Gannon co-PI's) and funding started effective 10/1/02. The project is a collaborative effort including Argonne National labs, Indiana Univ., University of Michigan, and NCSA. The proposal was the largest funded proposal awarded by the NSF program and was a direct result of efforts conducted as part of the DOE Portals project for our first efforts conducted under Task 1: Portal Systems (see below for details).

SDSC focused on three major tasks related to data management: Continued development of data handling features required for portals; Integration of the Storage Resource Broker data management system with the MDSPlus system; Integration of the Storage Resource Broker with portal technology. Much of this work was done in collaboration with General Atomics Fusion group.

All teams continued to make progress on their specific task areas with a focus towards implementing the plans formed at the February AHM (see Appendix A) and integration of these tasks. In particular, plans for hardware and software to be deployed on the Fusion grid were formed and implemented. Plans

for the Fusion Grid Portal, with a focus on the PreTRANSP application were solidified, the architecture designed, and initial programming and development were started. In addition, we began development of a DOE Demo Portal, which would provide a test and development framework for generalized DOE Portals requirements. Plans are underway to have beta portals in operation for the August NC2004 meeting at Argonne. Finally, a meeting was held at the HPDC/GGF meeting in early June by several DOE Portal development projects, where it was decided to form the DOE Portals Consortium.

2.1.2 Problems/Roadblocks

Impact of WSRF: The recent announced of the new standard for grid-enabled Web service, the WSRF (and related) standard, which will replace the OGSI standard, will have an impact on short term deliverables for this project. However, the team feels that in the long term this new standard is a good improvement over the previous standard and will not impact our ability to produce useful portal related services. However, as a result, the Fusion Grid project will not use an GT3 type services, but will continue to use GT2.x services. Some adaptation of portal infrastructure will be required and the team is already responding to this.

Fusion Grid Requirements: The Fusion Grid is not a typical Grid: it does not use GridFTP to move files (rather it uses MDS built-in calls), it does not have a large number of distributed HPC resources (only needs a single cluster), and there is currently only a single grid-enable program (preTRANSP). As a result, the Fusion Portal has been limited to a Pre-TRANSP portal (as opposed to a full fledged User Portal/HotPage as originally proposed). The Fusion team at General Atomics has proposed that we move the deliverable of a full portal to year 3 in order to give them time to familiarize the Fusion community with grid computing overall, and to portal usage as well. In addition, we need to solve the Fusion grid account problem. This project, through its collaboration with the DOEPC, will work to identify new portal projects for development and deployment.

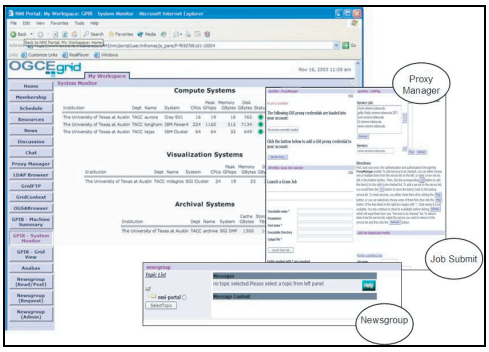
2.2 Accomplishments

2.2.1 Task 1: Portal Systems

Open Grid Computing Environments (OGCE) Project:

In 2003, TACC and Indiana were successful in winning the recently funded NSF Middleware Initiative for Grid Portal Development project (led by M. Pierce, IU). The project is a collaborative effort including Argonne National labs, Indiana Univ., University of Michigan, and NCSA. These institutions form the charter members of the Open Grid Computing Environments (OGCE) consortium (http://www.ogce.org). OGCE provides the grid portal community with sharable components that allow portal developers to quickly create Grid Portals. OGCE libraries will support baseline Grid technologies, freeing developers to concentrate on the specialized needs of a particular scientific community or collaboratory. The project will leverage funding from several national programs including the TeraGrid, the Alliance and NPACI (PACI) programs, the NEESGrid project, as well as several related DOE projects.

Much of the R&D efforts spent by both TACC and Indiana as part of the Task 1 deliverable contributed deeply to the success of this proposal and it is expected that the Fusion portal project will continue to be a primary requirements driver for this project. The Fusion portal has been selected to be one of the principle demonstration of capability portals for this program. A demonstration Fusion portal was presented at SC03 based on the alpha OGCE portals tool.



Significant progress was made on a preliminary OGCE portlet release which resulted in several test portals being demonstrated at SC03, including a Fusion Portal. The current list of features and capabilities include:

1. **Schedule:** This calendar program allows you to schedule appointments and attach associated documents and URLs. Calendars can be used to schedule both user and group activities.
2. **Resources:** Users and groups can upload and share documents and URLs.
3. **News:** Several news feeds are available, including XMLHack and BBC News.
4. **Discussion:** Participate in discussions with other members of your group
5. **Chat:** Engage other members of your group in online discussion.
6. **Proxy Manager and Job Submit:** Obtain and manage one or more Grid credentials that can be used to access Grid resources through your browser. Also shown are GRAM job launchers and a sample "ping" portlet.
7. **LDAP Browser:** Navigate the LDAP server of your choice.
8. **Grid FTP:** Use your Grid credential to browse remote directories and to upload and download files between your desktop and a remote host.
9. **Grid Context:** Store arbitrary web objects (movies, web pages, audio files) in a customizable, annotated directory tree.
10. **GridPort Information Repository Portlets:** Several GPIR portlets are available for browsing *Machine Summaries, System Status,* and *Total Grid Capabilities.*
11. **Anabas Impromptu:** The portal includes a window to the Anabas Impromptu system for real time shared display, audio, and chats.

Future releases will include application monitoring and application factory services, as well as job scheduling and basic workflow.

Indiana developed a portlet version of the Community Grids Lab (CGL) newsgroup system. These include portlets for reading and posting messages, portlets for managing newsgroup user accounts, and portlets for administering news topics. We integrated this into the OGCE portal release, which combines elements of NMI and DOE funded projects. IU also developed a portlet release for the CGL citation browser system, but have not yet integrated this into the OGCE release.

SDSC implemented a portal system based on the prototype OGCE release and hosted on resources at UCSD. In order to test the generality of the portal system, an NSF project, the Southern California Earthquake Center, is using the system to build a digital library portal. The system is supported by funding from NSF, and will integrate the DOE portal technology with dataflow processing systems that manage execution of web services, with access to collection management systems provided by the Storage Resource Broker. This will be an excellent test of the generality of the approach, and the efficacy of web-service based dataflow processing systems.

The OGCE consortium released a production version of OGCE that was synchronized with the NMI R5 release (May 2004). A developers meeting was held, where we began to develop generalized

interfaces for grid functions (FTP, Jobsubmit, etc.) which will be part future releases. This work is being done in close collaboration with Gregor von Laszewski and K. Jackson of the DOE CoG project.

An issue being faced by the portal community is the new JSR168 portlet standard, which is due out in June, 2004. JSR168 is supposed to provide a mechanism that will allow portlets developed under different frameworks (e.g. CHEF, WebSphere, GridSphere) to be interchangeable, which would ideally promote component reuse. The OGCE Framework is currently dependent on the CHEF Framework, which is being migrated to the new Sakai Framework, which is currently slated to be JSR168 compliant. However, the Sakai project is behind schedule which may impact the migration of OGCE portlets to JSR168. As a result, members of this team are investigating alternative portlet frameworks that will allow us to work with JSR168.

GridPort Toolkit

The diagram below shows the revised architecture of a portal based on the GridPort 3.0 Toolkit, which itself has been updated to newer technologies (see Annual report). The portal hosting environment will support OGCE portlets. Recent advances in GridPort include the integration of G. von Laszewski's Java COG (a related DOE funded effort), the integration of Globus 3.0 services, and the inclusion of a new GGF standard, the Community Scheduling Framework (CSF) for job scheduling and reservations.

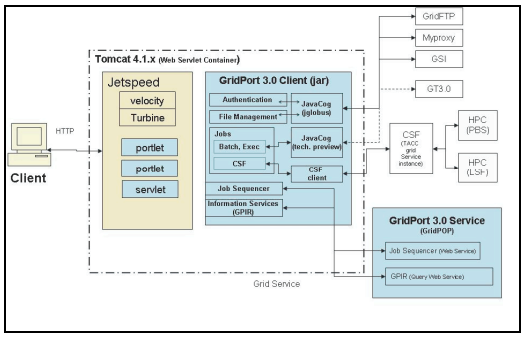


Figure 2. Revised architecture for GridPort Toolkit

TACC Released a production version of GridPort 3.0 and is available for download at, http://www.gridport.net. In addition, TACC held a training class for GP3 and portlet development in March, 2004. The GP3 toolkit was also included in the NMI R5 release package. The GP3 release includes automation for authentication, job submission, file transfer; integration of meta-schedulers; exposing Web and grid services; retention of grid meta-data; historical tracking of grid tasks; support for multiple VOs; support for a broad variety of client types including portals, command line shells, desktop applications. GP3 also provides the following web services: Accounts and Authentication; Batch Job Submission; Command Execution; Job Sequencer; File, Data and Collection Management; Grid Portal Information Repository (GPIR); Demo Portal.

Advanced File Transfer Service: TACC is in the process of designing and developing an advanced file transfer tool that would be incorporated into GridPort as a component. It will utilize GridFTP and the GPIR Web Service, and can be exposed as a web service. The tool will allow for multiple file transfer via GridFTP. The tool will also provide estimated bandwidth and latency rates between source and target hosts for each file before transferring the data.

Collaborative Tools

The Global MMCS group (Wenjun Wu, lead) is planning a May initial release (details below). These will include Access Grid, Polycom, and Real Player clients/portlets. Global MMCS Initial Release:

- Server Package List
  - Web Server

    - AV Session Server
    - Media Server Package
    - Image-Grabber, Audio Mixer, Video Mixer
    - H.323 Server Package
    - H323GW
    - H323Gatekeeper (Open source)
    - Stream Server Package
    - Streaming Gateway
    - Helix Server (Commerial & Open Source)
    - (NaradaBrokering Server
- Supportive Software Package:(packages from other vendors)
    - Tomcat
    - Helix Producer SDK
    - OpenH323
    - Java Media Framework
- Services provided by this package
    - Polycom
    - Access Grid
    - Real Player
    - two modes of the conferencing: broadcasted lecture vs Highly interaction seminars

Global MMCS will allow users to remotely join an Access Grid session from laptops. GA/Fusion personnel were able to connect MMCS clients using "Unicast AG" to the GA Access Grid computer under Windows XP. This worked from both GA and home computers. IU will set up and host the services needed in order to use GlobalMMCS to broadcast the morning meetings. With the current software we can do the following:

- multicast to XP clients (two-way)
- broadcast images to viewers on other clients via an applet (passive)

Indiana has started to work on a design for a Mac OS X client, which is expected to be ready mid-Septemeber.

Collaborative IDL: Work continues on the development of shared IDL sessions for the Fusion user community. Indiana staff obtained an IDL license and have begun developing and testing locally. They plan to have a demo working in xterm that can publish to listening windows managed by subscribing IDL sessions. For this, they will need to modify software to handle arbitrary IDL events

Technology Evaluation

Both TACC and IU continue to evaluate JSR168 containers (Pluto, GridSphere, uPortal, Sakai). In addition, TACC, IU, and SDSC have already begun to work with OGSA/OGSI services, and are now responding to the newly announced WSRF standard proposed by Globus and IBM. This new standard will have a tremendous impact on this project and will end up resulting in serious delays, since much of our work must now be redone.

2.2.2 Task 2: Data Collection Management

Storage Resource Broker (SRB/SDSC)

The activities at the San Diego Supercomputer Center are focused on the integration of the Storage Resource Broker (SRB) with the MDSplus data management system, the integration of WSDL services on top of the SRB, development of Java API extensions as needed, and the installation of the SRB at selected sites. The SDSC staff involved in the collaboration include Wayne Schroeder who supports the

MDSplus integration effort and provides scripts for SRB installation and administration; Lucas Gilbert who supports the Java API extensions; Arcot Rajasekar who supports the metadata catalog upgrades; and Reagan Moore who serves as liaison to DOE workshops. SDSC has volunteered to host the initial metadata catalog service for the SRB/MDSPlus integration on an Oracle database at SDSC. In the interim, a metadata catalog can be installed at GA.

Lucas Gilbert is the developer of a Java interface to the Storage Resource Broker, called JARGON. Additional improvements were made to the Jargon environment in anticipation of the development of a SRB portlet based on Java. The improvements include:

- Data ingestion
  - Added checksum API to JARGON
- Data retrieval
  - Added bulk unload API to JARGON.
  - Developed GUI components to display collection trees for SRB Browsing. This optimized model significantly speeds up file viewing at the GUI level.

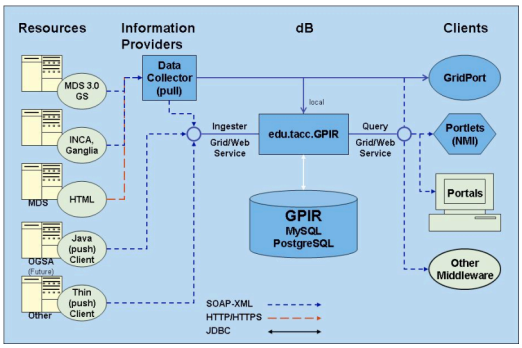
SDSC will work with Tomislav Urban and Maytal Dahan on extensions to the SRB Java API that are needed to support GP3 integration. The goal will be identification of capabilities required for full GP3 functionality.

2.2.3 Task 3: Grid and Web Services

GridPort 3 Web Services

The beta GP3 release also provides the following web services: Accounts and Authentication; Batch Job Submission; Command Execution; Job Sequencer; File, Data and Collection Management; Grid Portal Information Repository (GPIR); Demo Portal.

Grid Portal Information Repository: Improvements to GPIR is ongoing, and include the ability to aggregate and cache grid and portal related data collected by arbitrary information providers. The architecture was revised slightly (see diagram) to allow GPIR services to be hosted within the GridPort toolkit or as a standalone service. Due to technology limitations of SQL, GPIR has been updated to use a PostgreSQL database.



Revised GPIR 1.0 Architecture

GPIR data is now available via a single tool using standard Web service queries (soon Grid services). GPIR provides persistence mechanism for GridPort and includes:

- Archived historical data on Jobs, Load, Machine Status, Node Status;
- Grid configuration info as well Machine info (e.g. IP, Memory, Disk, MOTD, etc.)
- Multiple VO Membership, Admin Contact (if appropriate)
- Integration of MDS 3.0, INCA, and other providers in progress
- Version 1.0 of GPIR (SC2003): PostgreSQL

Portlets were also developed to display all of the GPIR data in both a current/recent snapshot format as well as an historical plot of the data.

Advanced File Transfer Service (TACC): TACC is in the process of designing and developing an advanced file transfer tool that would be incorporated into GridPort as a component. It will utilize GridFTP and the GPIR Web Service, and can be exposed as a web service. The tool will allow for multiple file transfer via GridFTP. The tool will also provide estimated bandwidth and latency rates between source and target hosts for each file before transferring the data. Portlets will be developed that will display this data to the portal user.

Job Sequencer/Community Scheduler Framework: TACC developed 2 new Grid/Web services: the Job Sequencer and an interface to the Community Scheduler Framework (CSF) capabilities as part of our GridPort Toolkit development. Portlets were written to support these services and demonstrated at SC03. These are scheduled to be integrated into future OGCE releases and hence integrated into the DOE project where appropriate

SRB Web Services

SDSC has implemented a complete set of WSDL web services for SRB managed collections. These include services for accessing both data and metadata. The WSDL services are based upon a Java client that implements the SRB protocols. The Java client development was partially funded under the DOE portals project. The Java client has been upgraded to support the peer-to-peer federated SRB system

A grid-compliant set of services is being developed. An OGSA compliant service was developed for an SRB query, using the GT3 toolkit. This required installing a complete GT3 release, and configuring a Tomcat server to run the SRB factory service. This initial version was developed under funding from NSF.

A design effort has been started to integrate dataflow systems with the SRB through use of the JARGON interface. The activities include development of specification for:

- SRB webservice as a dataflow pipeline
- SRB dataflow services as portlets
- Smatrix, command-line matrix tool for describing dataflows

The development of an OGSA compliant set of web services for the SRB depends upon a fully functioning Java API. Under funding from the DOE fusion portal, SDSC is developing a Java API called Jargon. A set of WSDL services has been implemented on top of the Java API. A workflow environment called Matrix is being implemented that builds on the WSDL services.

The initial release of JARGON provided the functionality of pure Java File-I/O for the SRB. The following releases added Metadata querying and modification, as well as the SRB parallel file transfer and bulk file upload. This quarter, using JARGON the SRB Administration Tool was converted from the, often difficult, Java Native Interface (JNI) into a pure Java program. With the appropriate extensions to JARGON in combination with the SRBAdmin tool, SRB Administration can now be done from any java ready computer.

JARGON continues to get good reviews from its users, especially for its effective simplification of what can often otherwise be a complex programming task. This growing user base has also helped build the API into a more robust system, as bugs get found, reported and fixed.

Also of significance was the recent inclusion of the GSI authentication methods. This was a non-trivial effort as the security protocol in GSI is different in the Java implementation than the C implementation. Wayne Schroeder modified the security interface AID library so that it detects whether security headers are not being sent, and if that is the case, switches into a no-header mode for both sends and receives. In testing we found that this works fine (from Java API as well as continuing to work from AID C clients with headers).

In a related security development effort, delegation functionality was developed in collaboration with the UK e-Science Data Grid and added as new API routines in the aid.c library. This makes it possible to base authentication on the user's identify, rather than a SRB server's identity.

JARGON can now verify identity using GSI proxy certificates when connecting to the SRB. This is an important step in creating a grid compatible environment. Further extending the compatible environments of the SRB and improving its acceptance by the greater community, was the creation of SRB URI. The SRB Uniform Resource Identifier (URI) may be used to resolve the address of a network entity to which SRB calls are directed. The SRB URI was submitted as an Internet Draft to the Internet Engineering Task Force (IETF), for registration with the Internet Assigned Numbers Authority (IANA). This process is the first step in the development of an Internet Standard and will ensure that there is no duplication of the URI scheme "srb".

Finally, as the SRB continues to grow, certain protocol changes must occur to scale with the new demands. With the SRB3.1 release, the old metadata query protocol would upload at least 1/4MB per query. The same query with the new protocol can now be less then 100 bytes. Obviously, adding this new protocol to JARGON was an important upgrade to the speed of metadata querying. A listing of all SRB/Jargon commands is given in Appendix B.

2.2.4 Task 4: Applications

FusionGrid Portal

The following tasks were identified in the Feb. AHM (San Diego, CA), and we report on their status below:

- Fusion portal hosting environment and hardware: a set of resources has been configured and set up. The OGCE Portal software has been installed and tested and is operational. A solution needs to be determined in order to move the portal into production (there is a need to have ports open continuously and this has not been approve GA/Fusion security experts).
- Establish final list of desired web services (see Appendix C).
- Investigate the use of messages/events to develop a collaborative IDL capability.: this is described in section above. Collaborative technologies are being used to develop shared IDL sessions and the IU GlobalMMCS softare will be used to allow remote clients to plug into an Access Grid meeting.

We worked on building the required set of web services listed in Appendix D. Our main goal in the next quarter will be to build a functional, production worthy Fusion portal and web services hosting environment that supports the required software while meeting Fusion and DOE security needs (see Appendix E). Progress on these services and the portal are dependent on the production releases of the OGCE software (expected early May).

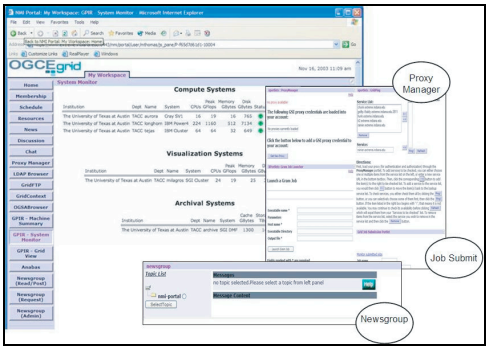


Figure 3. Screenshot of prototype Fusion portal based on NMI Portals software

We continued to develop specific portlets to support pre-TRANSP submission. Because of the nature of the Fusion community, the scope of the initial portal will be limited to Pre-TRANSP capabilities. This is because the user community is not familiar with grids at this time, and there are already accepted web pages for user information that are used by the Fusion community. In Year 3 we will work with Fusion users to migrate those pages into a unified portal.

DOE Demo Portal

The DOE Demo portal project was initiated this quarter. This portal will be the development and capability demonstration venue for current and future Portal and related services and technologies. Initially, the GPIR and SRB portlets will run from this portal. This will be the demo portal used by the DOE Portals Consortium as well.

SCEC Portal:

SDSC has implemented the NMI portal technology under funding from the NSF Southern California Earthquake project. The portal accessed collections registered into the SRB data grid, accessed a 3D visualization system, and supported job execution on the Teragrid. All data files were retrieved through the data grid and all output files were registered back into the data grid. The technology was demonstrated at an NSF review. The system validated the concepts being pursued in the DOE Fusion Portal. The major lesson learned was that the use of a portal to manage the complexities of grid technology was essential for creating a user friendly environment.

In a related effort, SDSC continues the integration of the portal system with a web services based dataflow system developed under NSF Grid Physics Network funding. The resulting system will be applied to the Fusion portal to explore alternate mechanisms for managing dataflows.

An urgent task is to identify Fusion resources on which to host both the portal software systems (GridPort, GPIR, NMI Portals services as needed) and the Storage Resource Broker data grid software. The demonstration for SCEC illustrated the importance of being able to access all data through data grid technology, implying the existence of data grid servers in front of the storage repositories.

2.3 Collaboration, Events, Outreach

Project Meetings

Several of the PI's and team members attended the combined HPDC and GGF meetings in Hawaii. All will present or participate at either GGF or HPDC (see section 4 below for details). In addition, we participated in the OGCE developers meeting held at NCSA in March, 2004.

We are preparing for the August NC2004 meeting at Argonne. For this, we plan to have an alpha version of the Fusion Pre-Transp portal and an alpha version of the DOE Demo portal up and running. Several of the PI's and team members plan to attend the combined HPDC and GGF meetings in Hawaii. All will present or participate at either GGF or HPDC (see section 4 below for details).

Synergistic/Collaborative Activities

- March 7-13 – R. Moore attended the Global Grid Forum meeting in Berlin, and gave presentations on Grid Evolution, Preservation Environments, Grid File Systems, Grid Metadata management, and data grid federation.
- March 10-12 – M. Thomas, M. Pierce, E. Roberts, A. Seth attended OGCE developers meeting at NCSA (Urbana, IL).
- March 16-18 – R. Moore participated in the DOE data management workshop at SLAC
- March 23 - GridPort 3/Portals training class held at TACC.

- June 6-10, 2004 – M. Thomas, M. Pierce, D. Schissel, R. Moore attended the Global Grid Forum meetings.
- June 4-6, 2004 – M. Thomas, and M. Pierce attended the High Performance for Distributed Computing (HPDC) meeting. Thomas participated in HPDC poster session.
- April 1-2, 2004 - M. Thomas, M. Pierce attended the OGCE Portal Developers workshop in Urbana, Il.
- Ongoing – M. Thomas, D. Gannon co-leading TeraGrid Portals Working Group. Led effort to write 3 key documents: the TeraGrid User Portal Plan, TG User Portal Requirements, and the Common TeraGrid Portals Software Stack for Portals (CTSS-Portals). Documents can be found at http://cvs.teragrid.org. The TeraGrid will utilize many of the technologies resulting from work done as part of this DOE Portals Project

DOE Portals Consortium: The PI's from 3 DOE portals projects met at GGF/HPDC to discuss the benefits of forming a portals consortium. They represented the following projects:

- DOE Portals (Thomas, Fox)
- Middleware Technology to Support Science Portals (Gannon, Bramley)
- CoG kits in Support of Applications and Portals (von Laszewski, Jackson)

The goal of the meeting was to explore the potential to have a well defined collaboration where project members can share experiences, software, web services, etc among the R&D teams, and also provide the DOE community with a common gateway (informational portal) to access information, examples, etc. Since we are also all involved in OGCE, this would of course extend that project, but we think that the DOE community needs its own resource center. The benefits include:

- A centralized access point/gateway for potential application projects to learn about DOE specific software and services
- A resource for application developers to connect with our project teams who can help them with their applications
- A mechanism to facilitate inter team collaboration and contributions among each others projects
- Improved efficiency: we can reduce overlap and duplication of effort by coordinating and sharing project resources and tasks

The results of this meeting will be presented at the annual meeting in August.

Traning/Education

The GridPort team presented a portals and GP3 developers class at TACC/Austin in March, 2004. Plans are underway for an August training class.

2.4 Publications

2.4.1 Papers published or in progress:

- M. Thomas, J. Boisseau, M. Dahan, E. Roberts, A. Seth, T. Urban, D. Walling. *GridPort v3: Enabling the Rapid Development of Web Services-Based Grid Portals and Applications*. In progress.
- TACC preparing submission for HPDC-13 on revised GridPort architecture and GT3.0/OGSA experiences.
- Rajasekar, A., Wan, M., Moore, R., Schroeder, W., "Peer-to-peer Federation of Collections", in preparation.
- R. Moore, A. Rajasekar, "Storage Resource Broker", Supercomputing '04 Conference, Tucson, AZ, September, 2003.
- M. Thomas, J. Boisseau, M. Dahan, E. Roberts, A. Seth, T. Urban, D. Walling. *GridPort v3: Enabling the Rapid Development of Web Services-Based Grid Portals and Applications*. In progress.

2.4.2 Presentations Given:

- Jan 20-24 – M. Thomas, M. Pierce, D. Schissel, R. Moore attended GlobusWorld 2004 meeting.
  - Thomas gave presentations on both the OGCE project and the GridPort 3.0 Toolkit
- Feb 4-5, 2004 - SDSC hosted the Portal all-hands meeting.
  - M. Pierce presented OGCE overview
  - M. Thomas presented GridPort 3 overview
  - Wayne Schroeder prepared and gave a short talk on our SRB/MDS+ integration work.
  - D. Schissel presented Fusion Grid/TRANSP update
  - K. S. Schuchardt presented overview of PNNL CMCS portal as potential client of OGCE software
- Feb 18-20 – M. Thomas, M. Pierce participated in the NSF CISE Shared Cyberinfrastructure Division PI Meeting
  - Pierce presented invited talk on OGCE project
- March 22-24 – M. Thomas, D. Schissel, and R. Moore participated in the DOE SciDAC PI meeting in Charleston
  - Schissel presented update on DOE Fusion Grid Project
- May 21, 2004 - M. Thomas gave a presentation at SDSC titled: "GridPort v3: Enabling the Rapid Development of Web Services-Based Grid Portals and Applications."
- June 4-6, 2004 – M. Thomas, participated in HPDC poster session discussing GridPort Toolkit and Fusion Grid Portal.
- June 6-10, 2004 – M. Thomas presented a talk on the GridPort Toolkit at the GGF SAGA
- SC03: NPACI GridPort talk, CSF BOF,

2.5 Project Management

Project Meetings

We are preparing for the August NC2004 meeting at Argonne. For this, we plan to have an alpha version of the Fusion Pre-Transp portal and an alpha version of the DOE Demo portal up and running. Several of the PI's and team members plan to attend the combined HPDC and GGF meetings in Hawaii. All will present or participate at either GGF or HPDC (see section 4 below for details).

We will also begin working with Fusion and DOE personnel to develop a new on line account management system that will be integrated into the portal. Tasks will include:

- Fusion wants a web page for admin
- Akenti has been shut down -- Fusion will rebuild, consider using GPIR admin code

Synergistic/Collaborative Activities:

- March 7-13 – R. Moore attended the Global Grid Forum meeting in Berlin, and gave presentations on Grid Evolution, Preservation Environments, Grid File Systems, Grid Metadata management, and data grid federation.
- March 10-12 – M. Thomas, M. Pierce, E. Roberts, A. Seth attended OGCE developers meeting at NCSA (Urbana, IL).
- March 16-18 – R. Moore participated in the DOE data management workshop at SLAC
- March 23 - GridPort 3/Portals training class held at TACC.

Future Plans

Portal Systems: The OGCE and GridPort teams will continue to work on portals and portlets synchronized with the NMI releases.

Web/Grid Services: We will modify TRANSP and pre-TRANSP portlets to work with GridPort 3 and new OGCE/JSR 168 compliant software. We will continue to develop Web services to manage audio/video collaboration. We will investigate the integration of IDL events with NaradaBrokering event system for shared collaborations.

DOE Portal: We will work with the OGCE team to install a test portal and integrate the SRB portlets into this portal.

Fusion Portal: We will prepare a design and architecture document for the Fusion portal and will begin programming portlets that integrated the OGCE, Fusion Grid, MDS-Plust and related technologies. The portal tasks will include:

- Replace current IDL GUI with a portal no more IDL for clients
- TRANSP only runs at Princeton
- Run management database
- Presentation
- Code run database

We will also begin working with Fusion and DOE personnel to develop a new on line account management system that will be integrated into the portal. Tasks will include:

- Fusion wants a web page for admin
- Akenti has been shut down -- Fusion will rebuild, consider using GPIR admin code

SRB: SDSC has the following SRB dev/deployment plans:

- July 31 - partial-file locking implemented in SRB
- Aug 31 - MDS+ integrated with this version of SRB and the mods (SRB ifdefs) returned to the MDS+ code base.
- Sept 30 – The new version of SRB will be deployed and operating at SDSC and/or GA in support of the DOE Portal. SDSC will collaborate with GA for use cases to test the software. The testing will be a difficult task and will require a good understanding of MDS+ and of the system.

TACC staff will begin design of the SRB Portlets that will run under both the GridPort 3 demo portal framework (Jetspeed 1) and the OGCE framework. These portlets will be installed in the DOE demo portal.

---

3. Year 3 (2004-2005) Annual Report

3.1 Project Highlights

- Prototype Fusion Portal demonstrated to work on the Fusion Grid.
- DOE Portals Consortium Presents Paper at SciDAC 2005 Annual Meeting
- OGCE released JSR 168 Compatible Grid portlets for Proxy Management, Job Submission, GridFTP file management, and GPIR-based information services.
- Delivery of A/V portlets to manage Polycom/H.323 audio and video streams
- OGCE framework used for TeraGrid User Portal
- Introduction to Grid Computing Graduate class taught at SDSU
- No-cost extension granted to project for 2005-2006

3.1.1 Deliverables Summary

This year significant progress has been made on many of the project deliverables, including a prototype Fusion Grid Portal that operates on the production Fusion Grid, the completion of the SRB/MDS+ integration, the completion and testing of innovative collaborative tools such as the GlobalMMCS which facilitates/replaces Access Grid meetings and a new approach to using a common

visualization tool, collaborative IDL. This tool allows a master user (who is running the session) to share the IDL session with other users/clients.

The latest release of the OGCE framework includes portlets for data management (including SRB access and file transfer), identity management using GSI/MyProxy certificates, job submission and control, workflow and application factory portlets, and others. The usefulness of the OGCE framework is evidenced by the fact that it can be used to access a variety of grids including the DOE Fusion Grid, the NSF TeraGrid, and local campus grids.

3.1.2 Problems/Roadblocks

The projects effort to produce a production Fusion Grid portal encountered serious setbacjs associated with FG security. The FG site system administrators put up a firewall that prevented access from outside the GAT firewall. This prevented developers from working easily on the system and slowed down attempts to deploy a production portal that could be tested by a group of friendly users. In addition, the GAT team determined that there was not a significant need for a PRETRANSP capability, so the portlets developed for this application were not needed. Additionally, GAT/FG determined that there was a greater need to have a new approach to security for the Fusion Grid, and that tools and services should be developed using languages such as Python or Perl. This will be address in the next year. Additionally, it was difficult to get the GlobalMMCS services installed and hosted and used by D3D due to hardware and personnel limitations.

In the next year, the software systems will be ready for full technology transfer, as identified in the original proposal. As a result, we expect the Fusion Grid to identify both a security solution that allows users and software to access the portal and services from outside the firewalls. In addition, the FG must identify staff and hardware in order to bring the software into production status. This is also challenged by the fact that SciDAC-1 is ending, and SciDAC-2 is being considered, with the outcome for specific projects unknown.

3.2 Accomplishments

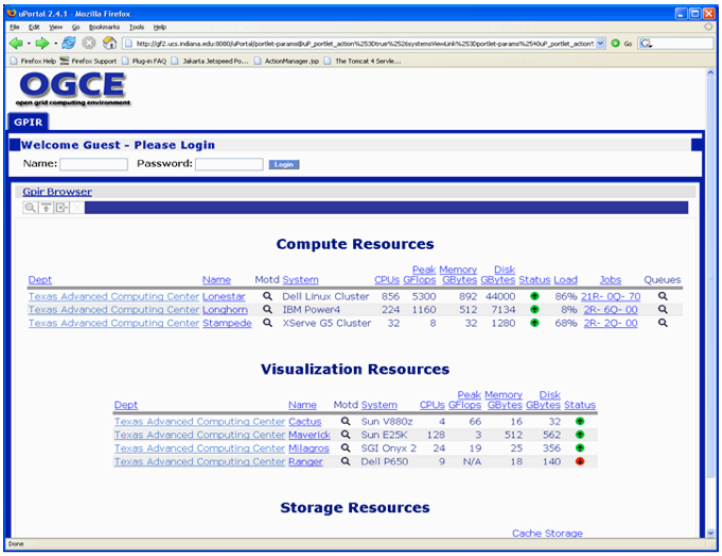


Figure 4. Sample login page and skin for the uPortal-based OGCE.

3.2.1 Task 1: Portal Systems

The OGCE's main focus of effort during this year was its transition to JSR 168 compatible Grid portlets. This was accomplished, and in the process we developed tools for porting older Jetspeed, CHEF, and OGCE1 portlets to the new standard. Below we summarize the core development work:

- Developed JSR 168 Compatible Grid portlets for Proxy Management, Job Submission, GridFTP file management, and GPIR-based information services.
- Developed general purpose Velocity support. All core OGCE2 portlets are written in Velocity. Velocity is the development tool of choice for Jetspeed 1 and CHEF-based development.
- Developed single, high level Grid programming interfaces, co-released in the Java CoG 4 Kit and OGCE2 downloads. These hide the differences between various Globus toolkit versions: GT2, GT3.0, GT3.2, and GT4.0.
- Developed support for multiple Grid clients environments within the same portal.
- Repackaged Open GCE Runtime Engine into component libraries, to ease subsequent reuse.
- Released Trebuchet desktop data management application.
- Developed supplemental data sharing portlet capabilities that allow portals to share proxy credentials.
- Developed an Apache Maven-based build, documentation, and test environment. This allows single portlet builds using remote Java jar repositories and improves over our previous Apache Ant-based build system. Users can deploy into either uPortal or Gridsphere containers.
- Developed an extensible unit testing environment (based on HTTP Unit) that can be used to verify the portal release.
- On-line and downloadable documentation: all system documentation is written in Xdoc format for easy distributions and management within the OGCE2 download.
- Developed new suite of JSR168 GridPort Portlets including GPIR, Job Sequencer, Comprehensive File Transfer, and SRB portlets.

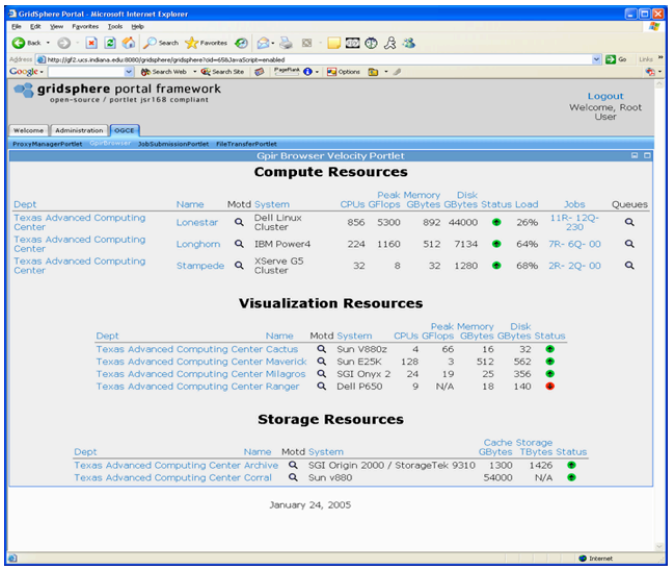


Figure 5. A sample screen of OGCE portlets deployed in GridSphere.

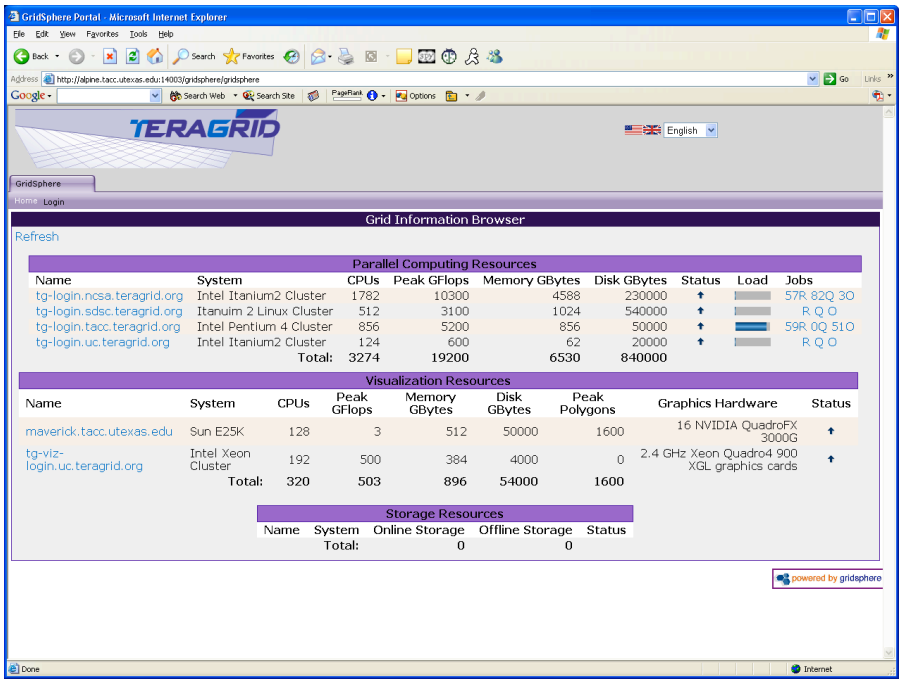


Figure 6. The TeraGrid User Portal is now under development at TACC, using OGCE portlets.

Technology Evaluation

In response to requests by the FusionGrid research scientists, SDSU launched an investigation into the use of portal frameworks such as Plone and TurboGears. Plone is a content management system that was developed in 2001 by Alan Runyan, Alexander Limi, and Vidal Andersen after the dissatisfaction with the state of the CMFDefault product that was shipped with the CMF. TurboGears is a rapid web development python framework developed in 2005 by Kevin Dangoor to create tools for Zesty News. Although both products worked as promised, the fact that they are content management systems (CMS) made it hard to modify layout and content, which is needed to make advanced science portals. This area requires further investigation.

In addition, the SDSU team is evaluating the use of AJAX technologies for the portlet. The use of AJAX will allow for users to have a more interactive experience with the portlets.

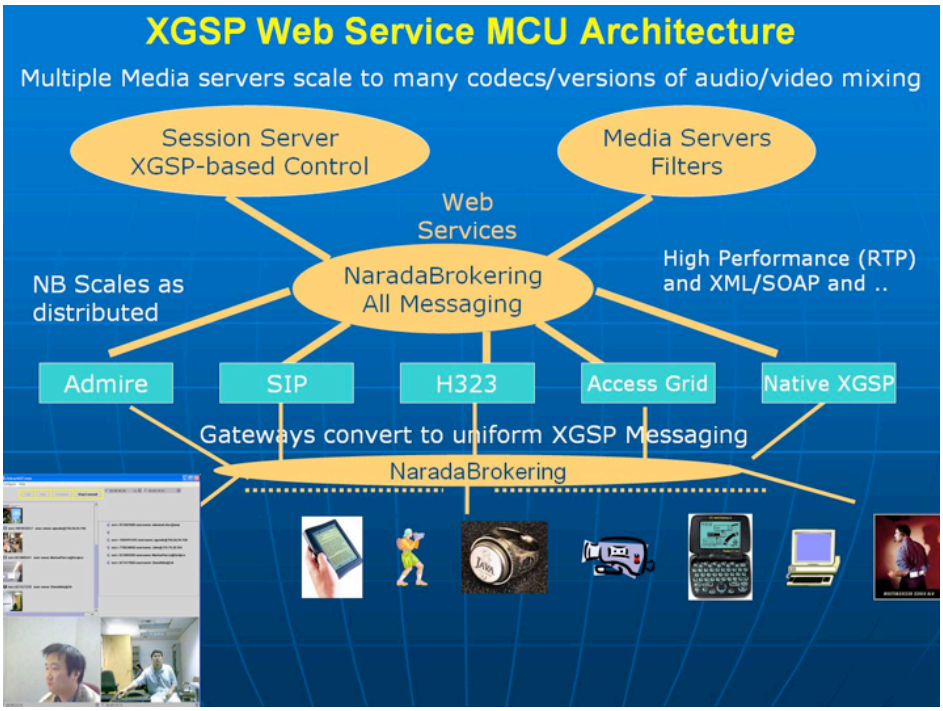


Figure 7. Architecture diagram for the Global MMCS software operating on the Fusion Grid. Inset photos on left show clients at both Indiana and GAT/Fusion Grid sites.

Collaborative Tools

Support for collaborative tools identified as early requirement For the Access Grid Meetings, where remote clients are limited, we determined that deployment of a Global MMCS PC client would allow more interactive participation. Fusion Science visualization software (IDL): need to share sessions during run times: developing event driven shared sessions.

The Global-MMCS Project is a service-oriented collaboration system, which integrates various services including videoconference, instant messaging and streaming. It can be regarded as scalable, robust, virtual MCU (Multipoint Control Unit, H.323 term) supporting multiple videoconferencing technologies (H.323/SIP/AccessGrid) and heterogeneous collaboration environments (Unicast/Multicast, Lossy, Wireless). Figure 4 above shows the architecture diagram for the Global MMCS software operating on the Fusion Grid. Inset photos on left show clients at both Indiana and GAT/Fusion Grid sites.

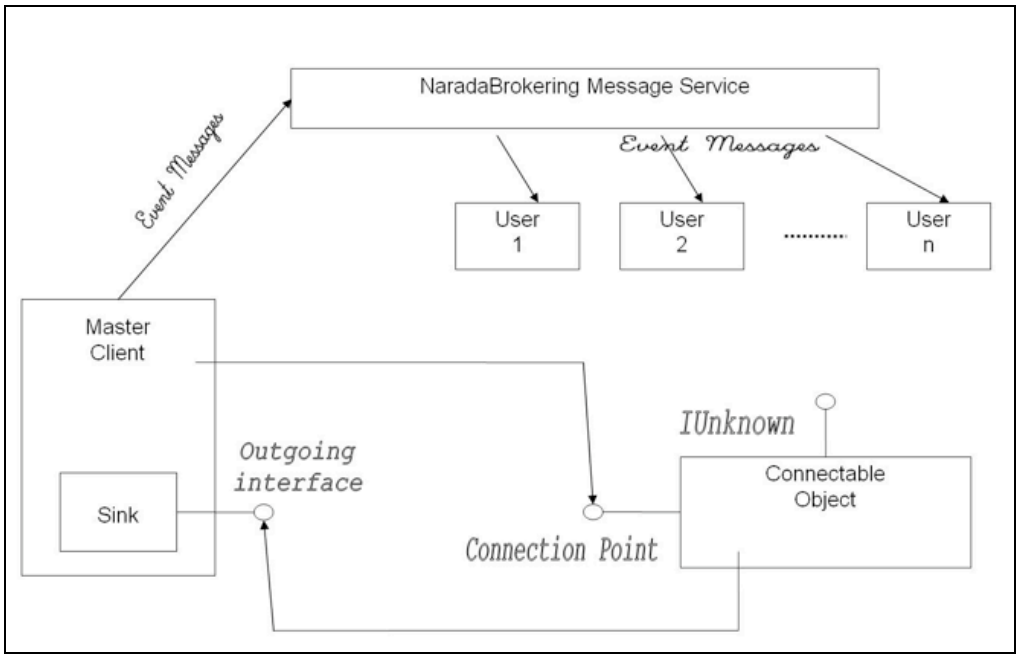


Figure 8. Architecture for the collaborative IDL software system.

The collaborative IDL system was demonstrated to run well between machines at Indiana and the D3D operations room. The D3D 'operator' ran the IDL session while the IU client observered. The next version will allow both users to interact with the IDL session.

This year, both collaborative tools were demonstrated to work on the GAT/Fusion Grid sytems. However, it was difficult to get the GlobalMMCS services installed and hosted and used by D3D due to hardware and personnel limitations. The portlet should help facilitate users having access to meetings.

3.2.2 Task 2: Grid and Web Services

The emergence of services oriented architectures (SOAs) and the Web Services Resource Framework (WSRF) has impacted all levels of work being done on this project. The Globus Toolkit (GT4) includes Web Services. As a result, all software components in the OGCE framework now interface to both the 'new' style GT4 services and the 'old' style (pre-WSRF) GT2 style grid services. The OGCE and DOE Portal teams have converted all services to use Web services. We highlight some of the Grid and Web services developed to date by members of this project:

Generic Application Factory

Indiana released the Generic Application Factory (GFAC) Web service and portlet. The application factory is a tool that wraps a standard application with a web service so that it may be invoked from a workflow or a portal. For any application installed as a command-line tool that takes files as inputs and produces files as outputs, a simple XML document that describes the application (where to find it, the description of the needed input and produced output files) will allow the factory to turn the application into a web service capable of launching versions of your application. The OGCE portlet then allows portal access to the Web service. See http://www.extreme.indiana.edu/portals/tools/factory.shtml

Generic Service Toolkit

The Generic Service Toolkit can "wrap" any command-line application as an application service. An application service is a web service interface to a command-line application. When an application service

is invoked with a given set of input parameters, it runs the application with those input parameters (possibly on a cluster of resources), monitors the application and returns the results to the user. This toolkit can be used by application providers (a.k.a application service providers or service providers to wrap any command-line application as an application service without writing any program code or modifying their applications in any way. The toolkit also provides a generic web service client that allows users to securely access any application service created by the toolkit from the convenience of a Grid Portal. When a user accesses an application service, the user is presented with a graphical user interface (GUI) to that service. The GUI contains a list of operations that the user is allowed to invoke on that service. After choosing an operation, the user is presented with a GUI for that operation, which allows the user to specify all the input parameters to that operation. The user can then invoke the operation on the service and get the output results.

GTLAB

Grid Tag Libraries and Beans (GTLAB) simplifies the process of Grid portlet development by encapsulating common Grid operations as XML tags that can be embedded in portlet pages, enabling rapid development. GTLAB capabilities include credential management, remote file operations, remote job executions, and file transfers. See http://grids.ucs.indiana.edu/users/manacar/GTLAB-website/

Narada Brokering

NaradaBrokering is a content distribution infrastructure, and enables the development of secure, failure-resilient systems. It is used by the collaboration tools developed for the FusionGrid project, as we well as other OGCE portlets and messeging services. The NaradaBrokering substrate itself comprises a distributed network of cooperating broker nodes. Since NaradaBrokering is application-independent, it has been harnessed in a variety of domains. The substrate places no constraints either on the size, rate and scope of the data encapsulated within the streams, or on the number of entities within the system. Some of the salient features in NaradaBrokering include:

- Powerful & efficient long-tail disseminations
- Secure end-to-end delivery of streams
- Robust stream disseminations
- Efficient ordering and synchronization of streams
- Support for rich Quality of Services
- Stream jitter reductions
- Support for multiple transport protocols
- High performance
- Enterprise messaging support: JMS compliant
- Support for Web Services

XBaya

Xbaya is an Java webstart application that can be used for composition, running and monitoring of workflows. It enables the users to run pre-composed workflows or the user can compose and run workflows on the fly.

XRegistry

XRegistry is an application service registry which allows users to register web service descriptions using GFAC portlet interface and access the services from XBaya toolkit.

3.2.3 Task 3: Data Collection Management

The capabilities of the Storage Resource Broker continued to evolve during the project, with 8 releases of the technology. All of the versions support parallel I/O streams and federation of data grids,

but a significant number of additional capabilities were added and integrated into the test platform. used by the Fusion community to be dynamically registered onto files stored in a SRB collection. SRB developed a GSI enabled version of the MCAT server. This allows clients to access SRB services using Grid certificates. The OGCE portal can use its single sign-on service to connect a portal user with the SRB services on the Fusion Grid.

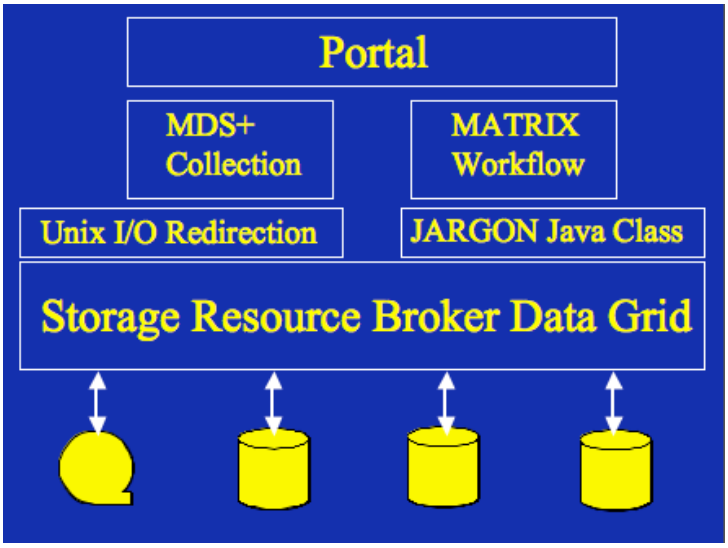


Figure 9. Fusion Data Grid architecture reflecting the SRB components.

As part of the effort to support the Fusion Data Grid portal (see below), the SRB team has begun to work with GAT to set up an MCAT server for the Fusion Grid. The initial set of MCAT keys will be based on the Fusion data sets such as EFIT, EQDISK, etc. SRB will provide the Fusion Data Grid and services will include the following:

- Provide WSDL interface
- Support for Portlet access to data handling system
- Support MDS+ functions
- Partial file locking
- Support large data sets
- Support migration of data to archives
- Provide disaster recovery
- Support replication
- Support grid extensions
- Support dynamic addition of storage resources

3.2.4 Task 4: Applications

FusionGrid Portal

A prototype portal for the Fusion Grid was installed and demonstrated using the OGCE portals, the SRB/MDS+ software and deployed on the Fusion Grid. The portal interfaced to the TRANSP system and pulled status data from Fusion Grid Monitoring service. The portal allowed users to analyze shot data from D3D experiments and to archive results into the SRB system. In addition, users could upload and download files between the portal, the SRB and the users local machine. Using other OGCE portlets,

users could transfer files between any supported host in the system. The portal also authenticated using the Fusion Grid certificate and MyProxy services.

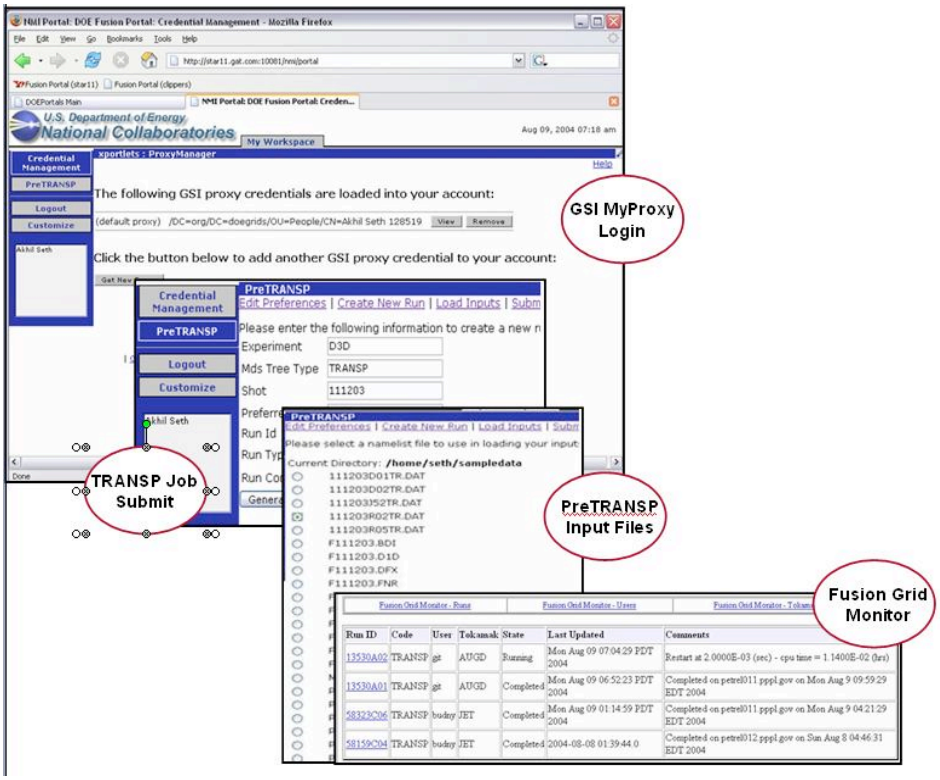


Figure 10. Example of the portal interface for the Fusion Grid Portal

There were, however, serious issues associated with FG security. The FG site system administrators would never allow the project to host a portal accessible from outside the GAT firewall. This prevented developers from working easliy on the system and slowed down attempts to deploy a production portal that could be tested by a group of friendly users. In addition, the GAT team determined that there was not a significant need for a PRETRANSP capability, so the portlets developed for this application were not needed.

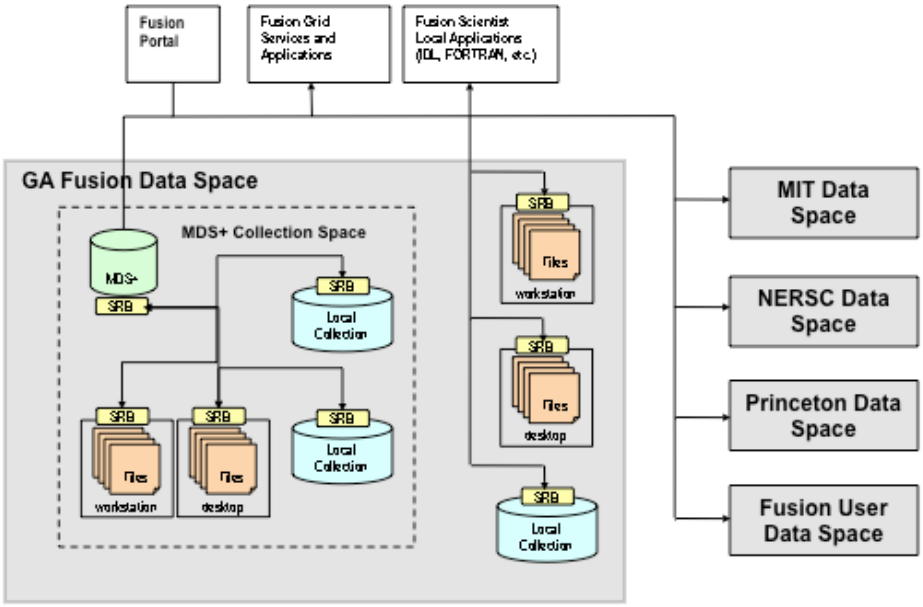


Figure 11. Architecture for a Fusion Grid Data Portal

In addition to the PRETRANSP portal work, the project team met with Fusion researchers at the D3D Tokamak project and determined that there was a deep need for a Fusion Data Grid Portal with capabilities including:

- FDG portal will be gateway to distributed Fusion Data:
- provide access to both MDS+ collections and user collections:
- Use SRB to plug user files into MDS+
- Use SRB to host larger collections that MDS+ cannot handle (has been shown it is limited)
- Fusion Metadata definitions:
- Defining library of metadata for fusion schemas
- To identify/generate metadata, we need to focus on
- Work with real equilibria and simulation data.
- code consumers of EFIT data: GATO, EFIT, NBC, TRANSP:
- CODES that generate EQDISK data: EFIT
- Will develop new approach to viewing metadata (visualization)

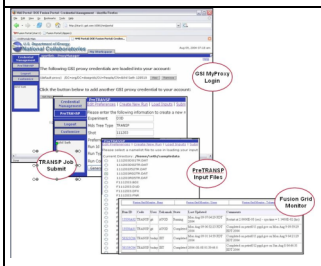
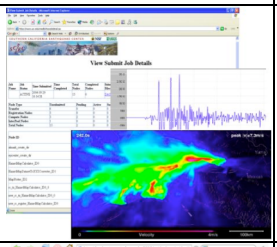
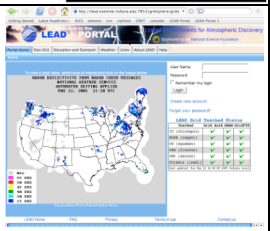
3.3 Collaboration, Events, Outreach

3.3.1 DOE Portals Consortium

The portal development activity within the DOE and the SciDAC program for the work described above has been driven by the need of the HPC community to provide scientific gateways to their communities of users, emphasizing the need for a set of tools and software solutions that can be utilized by this community. With this in mind, the SciDAC National Collaboratories (NC) program [19] and the DOE Office of Science has been funding research projects whose common goal is to produce reusable portal toolkits that can be depended to operate within the DOE environment. These efforts are organized as the *DOE Portals Consortium (DOEPC)* that includes the projects *Portal Web Services: Support of DOE SciDAC Collaboratories* and *Middleware Technology to Support Science Portals,* which are developing the core portal services and portlets described in the previous section, *CoG Kits: Enabling*

Middleware for Designing Science Applications, Web Portals and Problem Solving Environments, which provides the core Java and Python technology to access the foundational Grid Services, and the *Earth System Grid*, which integrates supercomputers, large-scale data and analysis servers and grids for climate research [12]. DOEPC is closely aligned with the *Open Grid Computing Environments Consortium (OGCE):* the NSF Middleware Initiative (NMI) for Grid Portal Development project, which is a collaborative effort including Argonne, Indiana University, NCSA, San Diego State University, and the Universities of Chicago, Michigan, and Texas.

Table 1. Figures 1-4 describe four portals that either use or significantly influence DOEPC directions.

	***Figure 1.*** *Earth System Grid (ESG): Primary Goal is to enable community access to climate simulation results distributed across DOE and NSF supercomputers. Features include user registration, authorization groups, file download, browse collection of usage and discovery-level metadata, free-text search, access to data on archival systems, "virtual data services" that provide flexible spatiotemporal subsetting from aggregated collections of files, user management, and metrics reporting. Communities Served: Worldwide climate research community. Over 1200 users registered. Available Data: IPCC, PCM, and CCSM simulations, over 75TB of data and over 400,000 files available. https://www.earthsystemgrid.org/*
	***Figure 2.*** *The Fusion Grid Portal provides a single view of the distributed data collections (MDS-Plus), local user files, and simulation data sets with computational resources. It will be used to run Fusion community codes such as TRANSP and GATO [5]. Based on GridPort/OGCE portlets, the portal allows users to manage accounts and credentials, view status of resources and Fusion Grid jobs, manage files and collections with the File Transfer and Storage Resource Broker (SRB) portlet [20], and submit jobs to the Fusion Grid. The portal uses Fusion Grid CA and MyProxy server.*
	***Figure 3.*** *Southern California Earthquake Center (SCEC) Portal: The SCEC project [21] uses the OGCE portal environment to build an interface to the SCEC digital library that contains seismic wave simulations, observational data, and visualizations. The SCECLib Portal includes portlets specifically developed to display seismic waves. http://webwork.sdsc.edu:10081/sceclib/*
	***Figure 4.*** *The Linked Environments for Atmospheric Discovery (LEAD) Portal: The LEAD portal effort allows users to manage data-driven, high performance weather simulations. These simulations can also be event-triggered, as data mining analysis on the data stream may indicate severe weather events. Detected events trigger simulation workflows on the Grid that are delivered to end users through the Web portal. See http://lead.ou.edu and [13].*

In many cases, the software developed by the DOEPC utilizes or has been influenced by other NC projects or Collaboratories including projects that are co-funded by other agencies such as the NSF and NASA. Several of these projects are described in Table 1 above.

3.3.2 Other Collaborations

- Demonstrations at Supercomputing 2004: We organized demonstrations at the Alliance, Indiana, and TACC booths.

- TeraGrid tutorial at Supercomputing 2004: TACC/Roberts demonstrated a development version of the TeraGrid User Portal for current and potential TeraGrid users during a comprehensive TeraGrid tutorial.
- Participation at Globus World: All, Von Laszewski gave two presentations.
- Participation in Global Grid Forums: All, Gannon and Fox both serve on GGF steering committees. Gannon and Pierce participated in the GGF 12 Portals tutorial. Gannon also gave a shorter portal tutorial at GGF 13. We are reviving the GCE group for GGF 14.
- Grid Portals Tutorial: ALL. Gannon and Pierce gave a tutorial at GGF 12. Gannon gave a shortened version at GGF 13. Thomas and Pierce gave a tutorial at HPC 2005. See Training and Outreach for more details.
- Supercomputing 2004 Posters: IU and University. of Chicago, Pierce and Laszewski had three posters selected for SC04. Laszewski's poster won the coveted "SC Best Poster" from 109 submissions.
- Distance Education: Indiana,/Pierce delivered a set of lectures on using portlets as part of a distance education course offered to Jackson State University.
- GridPort Training: TACC. Held multiple training classes throughout the year (see training and outreach section below). Hosted Venezuelan visitors for 3 weeks training them in OGCE and GridPort. Built Gaussian application portlet, and conducted GT4 evaluation.
- SDSU Course: Introduction to Grid Computing. Thomas taught a semester-long graduate course on Grid computing using OGCE, Globus, and Grids Center software.

3.3.3 Training and Workshops

- Gregor von Laszewski, Jay Alameda, Greg Daues, and Marlon Pierce, meeting with NMI GridChem Team, NCSA, April 18, 2005. Primary topic of discussion was building workflow clients and tools using the Java CoG kit.
- Eric Roberts, Akhil Seth, Maytal Dahan hosted 4 visitors from Venezuela April 10-30 training them in OGCE and GridPort. Built Gaussian application portlet, and conducted GT4 evaluation collaboratively.
- Mary Thomas and Marlon Pierce, Tutorial on Grid Portals HPC2005 April 4 2005 and NCSA portals meeting, Jan 26 2005.
- Geoffrey Fox and Marlon Pierce, "High Performance Computing and Computational Science." Spring 2005 Graduate Course taught in conjunction with Jackson State University. Course material included OGCE topics on portlet development and simple computational services. Course Web Site: http://grids.ucs.indiana.edu/ptliupages/jsucourse2005/. Course funding was provided by the DOD High Performance Computing Modernization Program.
- Marlon Pierce and Gordon Erlebacher, meeting with Quantum Espresso developers, Cineca, Bologna, Italy. This was part of a VLab effort (an NSF funded ITR project to build a computational chemistry grid). The Espresso team provides a state of the art suite of Molecular Dynamics codes. We will be building a portal and services to support grid applications surrounding the Quantum Espresso suite using the OGCE portal software. We also discussed other OGCE portlet efforts, particularly SRB portlets.
- Eric Roberts, Akhil Seth. "Building Grid Enabled Portals using GridPort 3" TACC training Austin, Texas December 10, 2004
- Maytal Dahan, Building Grid Enabled Portals using GridPort 3 Merida, Venezuela November, 2004
- Eric Roberts, GridPort: A Toolkit for Building Grid-enabled Portals Supercomputing 2004, November 8-12, 2004. Presented at TACC booth throughout the week of Supercomputing 2004.
- Eric Roberts, Maytal Dahan, Akhil Seth. Building Grid Enabled Portals using GridPort 3 TACC training Austin, Texas October 25, 2004.
- Marlon Pierce and Dennis Gannon, Developing Grid Portals Using Portlets Global Grid Forum 12 Portal Tutorial Brussels September 21 2004.
- Maytal Dahan, Eric Roberts, Akhil Seth. Building Grid Enabled Portals using GridPort 3 Monterey Grid Computing Conference at FNMOC, Monterey, California September, 2004.

- von Laszewski, Course project on Grid computing, Illinois Institute of technology, Fall 2004. Von Laszewski supervised 5 students at IIT to conduct projects related to Grid computing.
- Maytal Dahan, Eric Roberts, Akhil Seth. Building Grid Enabled Portals using GridPort 3 TACC training Austin, Texas August 17, 2004.

3.4 Publications

3.4.1 Presentations and Posters

- Eric Roberts, Steve Quinn, Von Welch, "TeraGrid Account Management through the TeraGrid User Portal", TeraGrid All Hands Meeting April 5, 2005.
- Dennis Gannon, Invited Speaker, "Grid Portals Mini Tutorial", Global Grid Forum, Seoul Korea, March 2005.
- Marlon Pierce, "SERVOGrid and GIS Services," SCIGN REASoN Team Meeting, NASA Jet Propulsion Laboratory, Feb 17, 2005. This overviewed portlet-based application portal work with SCIGN Data Portal development team, http://reason.scign.org/scignDataPortal/.
- Eric Roberts, Invited Speaker, GridPort: A Toolkit for Building Grid-Enabled Portals Portals, Portlets and GridSphere Workshop CCT, Louisiana State University, Feb 3-4 2005, Baton Rouge, LA, USA
- Geoffrey Fox iSERVO International Solid Earth Research Virtual Observatory: Grid/Web Services and Portals Supporting Earthquake Science American Geophysical Union Fall Meeting, San Francisco, December 15 2004.
- Geoffrey Fox and Marlon Pierce Data Grids for HPC: Geographical Information System Grids, Internet Seminar to Jackson State University, Florida International University, and University of Hawaii, 7 December 2004.
- Marlon Pierce, Marcus Christie, and Eric Roberts: The Open Grid Computing Environments Collaboratory OGCE portal demo at SC04 Thursday, Nov 11, 12:30 pm to 1:30 pm. with Additional Material
- Marlon Pierce, Greg Daues, Gopi Kandaswamy, and Liang Fang NCSA Alliance Portal Expedition Demo Alliance/OGCE portal demo at SC04 Wed, Nov 10, 1:00 pm - 1:30 pm with Additional Material.
- Eric Roberts, GridPort: A Toolkit for Building Grid-enabled Portals Supercomputing 2004, November 8-12, 2004. Presented at TACC booth throughout the week of Supercomputing 2004.
- Eric Roberts, Invited Speaker, TeraGrid: Learn Once Run Anywhere, TeraGrid Tutorial at Supercomputing 2004 Monday Nov 7, 2004 Gave demonstration of TeraGrid User Portal.
- Marlon Pierce and OGCE Collaboration posters for SC04
    - OGCE Release 1 Combines Grid and Collaboration Portlets in CHEF Framework
    - Standard Portlet Architecture Supports Reusable Components
    - OGCE Portal Applications for Grid Computing I
    - OGCE Portal Applications for Grid Computing II
- Marlon Pierce, Portals and Portlets DOD User Interface Specification group meeting Atlanta September 27 2004.
- Dennis Gannon, Keynote Speaker, "Building Grid Applications and Portals: An Approach Based on Components, Web Services and Workflow Tools," Europar, Pisa, Aug. 2004.
- Geoffrey Fox Web Services and Grid Architecture and their application to Earthquake Science Jade Palace Hotel Beijing (863 Project Grid) August 26 2004. (Earthquake Science in following talk)
- Geoffrey Fox and Marlon Pierce SERVO Grid: Solid Earth Research Virtual Observatory Grid/Web Services and Portals Supporting Earthquake Science Jade Palace Hotel Beijing (863 Project Grid) August 26 2004.
- Dennis Gannon, Invited Speaker, "Component Models and Distributed Services", SC4 DEVO - Workshop on Service Composition for Data Exploration in the Virtual Observatory, Caltech, July 2004.

- Geoffrey Fox, Marlon Pierce SERVO Grid: Solid Earth Research Virtual Observatory Grid/Web Services and Portals Supporting Earthquake Science Fourth ACES (APEC Cooperation for Earthquake Simulation) Workshop, Beijing China, July 13 2004. This was accompanied by an earthquake modeling application portal demonstration using the OGCE-based QuakeSim portal.
- Geoffrey Fox iSERVO and SERVOGrid: (International) Solid Earth Research Virtual Observatory Grid/Web Services and Portals Supporting Earthquake Science US-Australia Workshop on High-Performance Grids and Applications, Swiss Grand Hotel, Bondi Beach, Sydney, June 8 2004
- Dennis Gannon, Keynote Speaker, "Building Applications from a Web Service based Component Architecture". ICS 2004 Workshop on Component Models and Systems for Grid, St. Malo, France, June 2004.
- Dennis Gannon, Invited Speaker, "Experience Building Grid Applications with the CCA Software Component", Second Reality Grid Workshop, Royal Society, London, June 2004.
- Dennis Gannon, Keynote Speaker, "Building High Performance Scientific Applications for Parallel and Distributed Systems", First Workshop on Productivity and Performance in High-End Computing, Feb. 14, 2004, Madrid, Spain.

3.4.2 Book Chapters

- Gregor von Laszewski and Kaizar Amin. Grid Middleware, chapter Middleware for Commnications, pages 109–130. Wiley, 2004. http://www.mcs.anl.gov/~gregor/papers/vonLaszewski--grid-middleware.pdf.
- Geoffrey Fox, Shrideep Pallickara, and Marlon Pierce Building a Grid of Grids: Messaging Substrates and Information Management to appear as chapter in book "Grid Computational Methods" Edited by M.P. Bekakos, G.A. Gravvanis and H.R. Arabnia. Includes sections on OGCE component-based portals and portal services.

3.4.3 Conference Publications

- Sangmi Lee Pallickara, Beth Plale, Scott Jensen, Yiming Sun, "Structure, sharing, and preservation of scientific experiment data", *To appear IEEE 3rd International Workshop on Challenges of Large Applications in Distributed Environments (CLADE)*, July 2005.
- Sangmi Lee Pallickara, Beth Plale, Scott Jensen, Yiming Sun, "Monitoring Access to Stateful Resources in Grid Environments", *To appear IEEE International Conference on Services Computing*, July 2005.
- L. Fang and D. Gannon, "XCAP - An Extensible Capability-based Authorization Infrastructure for Grids", 4th Annual PKI R\&D Workshop: Multiple Paths to Trust, NIST Gaithersburg, MD April 19-21, 2005.
- Shirasuna S., Slominski A., Fang L., and Gannon D., Performance Comparison of Security Mechanisms for Grid Services, the 5th IEEE/ACM International Workshop on Grid Computing, Pittsburgh, Nov. 8, 2004.pp. 360-364.
- D. Gannon, S. Krishnan, L. Fang, G. Kandaswamy, Y. Simmhan, A. Slominski, On Building Parallel and Grid Applications: Component Technology and Distributed Services, Proceedings, Challenges of Large Applications in Distributed Environments (CLADE) In conjunction with the 13th International Symposium on High Performance Distributed Computing (HPDC-13), pp. 44-51, June, 2004. (To appear in a special issue of Cluster Computing.)
- D. Gannon, S. Krishnan, A. Slominski, G. Kandaswamy, L. Fang, "Building Applications from a Web Service based Component Architecture", in In V. Getov and T. Kielmann, editors, Component Models and Systems for Grid Applications. Proc. of the Workshop on Component Models and Systems for Grid Applications, June 26, 2004 held in Saint Malo, France. Springer, 2005, to appear. ISBN: 0-387-23351-2.

3.4.4 Journal Publications and Technical Reports

- Ahmet Sayar, Marlon Pierce, Geoffrey Fox OGC Compatible Geographical Information Services Technical Report (Mar 2005), Indiana Computer Science Report TR610. This includes an early description of our GIS portal client work that is included in NMI Year 2 deliverables.
- Mehmet Aktas, Galip Aydin, Andrea Donnellan, Geoffrey Fox, Robert Granat , Lisa Grant, Greg Lyzenga, Dennis McLeod, Shrideep Pallickara, Jay Parker, Marlon Pierce, John Rundle, Ahmet Sayar, and Terry Tullis iSERVO: Implementing the International Solid Earth Research Virtual Observatory by Integrating Computational Grid and Geographical Information Web Services Technical Report December 2004. This paper includes a description of application portal work based on Year 1 OGCE portal development.
- D. Gannon, J. Alameda, O. Chipara, M. Christie, V. Dukle, L. Fang, M. Farrellee, S. Hampton, G. Kandaswamy, D. Kodeboyina, S. Krishnan, C. Moad, M. Pierce, B. Plale, A. Rossi, Y. Simmhan, A. Sarangi, A. Slominski, S. Shirasuna, T. Thomas, Building Grid Portal Applications from a Web-Service Component Architecture, Proceedings of the IEEE. Special Issue on Grid Technology, vol. 93. No. 3, 2005.
- Eric Roberts, Steve Quinn, Michael Shapiro, Derek Simmel, Von Welch, Nancy Wilkens-Diehr "TeraGrid Account Management through the TeraGrid User Portal", Technical Report for the TeraGrid Requirements and Analysis Team, April 2005.
- Beth Plale, Dennis Gannon, Yi Huang, Gopi Kandaswamy, Sangmi Lee Pallickara, and Aleksander Slominski, "Cooperating Services for Managing Data Driven Computational Experimentation", *submitted manuscript* 2005.
- Beth Plale, Dennis Gannon, Dan Reed, Sara Graves, Kelvin Droegemeier, Bob Wilhelmson, Mohan Ramamurthy, "Towards Dynamically Adaptive Weather Analysis and Forecasting in LEAD", *To appear ICCS workshop on Dynamic Data Driven Applications*, Atlanta, Georgia, May 2005.
- Beth Plale, Dennis Gannon, Jay Alameda, Bob Wilhelmson, Shawn Hampton, Al Rossi, and Kelvin Droegemeier Active Management of Scientific Data *IEEE Internet Computing special issue on Internet Access to Scientific Data*, Vol. 9, No. 1, pp. 27-34, Jan/Feb 2005.
- Beth Plale, "Framework for Bringing Data Streams to the Grid", *Scientific Programming*, IOS Press, Amsterdam, Vol. 12, No. 4, 2004.

3.4.5 Posters

- Marlon Pierce, et al., The Open Grid Computing Environments Project. Supercomputing 2004, Pittsburg, 6-12.

3.5 Project Management

During this year, co-PI Thomas moved from the University of Texas at Austin to San Diego State University. With the approval of DOE Program Manager Dr. Scott, a portion of the grant also transferred to SDSU, in the form of a subcontract to TACC. Both TACC and SDSU now participate in the project.

Project Meetings

The project held regular bi-weekly telcons and met in person when attending related meetings. In addition, as part of the consortium effort, the portals team participated in several OGCE developer meetings where requirements for the OGCE Toolkits were defined and as a result, DOE needs were included in the design decisions and programming tools.

During the year, the team attended many professional meetings including the DOE SciDAC meeting, Supercomputing, High Perfromance Distributed Compting, and others. This is evidenced in the number of presentations and publications listed below.

Future Plans: No-cost extension of the project (approved)

The DOE Portals project was funded with a start and end date of September 1, 2002 through August 31, 2005. This project involves the efforts of team members from San Diego State University, the University of Texas at Austin, Indiana University, the University of California at San Diego, and General Atomics of La Jolla, CA. Due to procedural problems associated with transfer of funds from the DOE to U Texas, and then from U Texas to project subcontractors, many participants did not receive funding until Summer of 2003.

Because of local institutional policies, most participants were not allowed to begin work on the effort until funds actually arrived at their institutions. This caused critical delays in the project and impacted deliverables and the rate at which budgets were spent out. With approval from Program Manager, Dr. Mary Anne Scott, deliverable due dates were adjusted. We note that this project has been producing timely and significant results -- an invited paper and poster were presented at the June, 2005 annual DOE SciDAC meeting in San Francisco, CA. This paper will be published in a Spec. Edtn of Journal of Physics.

Most project participants still have significant funding remaining, and in discussions with Dr. Scott, we have determined that a no cost extension to this grant will allow the project to complete its goals. Therefore, we request an extension to this grant from August 31, 2005 to August 31, 2006.

The main technical challenge that we will address in the next year will be the simplification of portlet development. Most of the portlets developed by the OGCE are based on the Velocity template environment. This provides important backward compatibility to earlier OGCE Release 1, CHEF, Jetspeed, and (in progress) some Sakai portlets. However, Velocity is, for the most part, a scripting and programming environment, and it still requires a great deal of development to build portlets. Just as portals themselves may be built out of reusable portlets, the portlets need to be developed out of reusable parts.After reviewing available technologies, we have decided (at our February 2005 All Hands Meeting) that Java Server Faces provides and excellent baseline technology which may be used to build reusable Grid portlet widgets that can be used to quickly compose portlets. This will also support visual portlet programming in JSF-enabled IDEs (integrated development environments).

---

4. Year 4 (2005-2006) Annual Report

4.1 Project Highlights

- Open Grid Computing Environments (OGCE) Portlet Toolkit v2.0 released
- Global-MMCS: a service-oriented collaboration system completed
- Collaborative IDL prototype completed
- Integration of SRB with the FG MDS-Plus (data base) system completed.
- Pubcookie Web-Based Authentication Framework Demonstrated

4.1.1 Deliverables Summary

A key outcome of this project has been to (1) produce useful portal infrastructure that can be deployed on DOE resources and (2) to test some of these technologies on the SciDAC Fusion Grid. The project has met deliverables set out for part (1) and has produced the following results:

- Open Grid Computing Environments (OGCE) Portlet Toolkit
- Investigation of portal frameworks based on Python (a language in heavy use within the Fusion Community)
- Global-MMCS: a service-oriented collaboration system that integrates services including videoconference, instant messaging and streaming
- Collaborative IDL: IDL is a programming/visualization tool (similar to Matlab) that is heavily in use in the Fusion Community; this effort prototyped a collaborative approach
- Integration of SRB with the FG MDS-Plus (data base) system
- Formation of the DOE Portals Consortium: a collaboration between several SciDAC-1 programs. This has resulted in several significant outcomes, including:

- Integration of the PubCookie authentication scheme into the Fusion Grid security system. This has been done in collaboration with the Fusion Grid project and Lawrence Berkeley Labs (LBL).
- Integration of the PURSE authentication system into the OGCE. PURSE was developed as part of the SciDAC funded Earth Systems Grid (ESG)
- Interactions with the SciDAC funded JavaCoG and PythonCoG projects that has helped drive requirements of those projects.

4.1.2 Problems/Roadblocks

The main roadblocks this year were associated with the maintenance being done on the DIII-D Tokamak, which prevented completion of some software demonstrations (Collaborative IDL and Global MMCS). In addition, the challenge of outreach to DOE projects persisted despite outreach efforts. This is most likely because few projects have budgets sufficient to support a portal team and typically most projects underestimate the level of effort required to develop a capable science portal.

4.2 Accomplishments

4.2.1 Task 1: OGCE2 Portal

The OGCE2 series of software releases is based on the JSR 168 portlet standard and is distinguished from our OGCE1 series in Year 1.

Version 1.0 Release

OGCE released JSR 168 versions of our Grid portal suite, including GPIR, GridFTP, Job Submission, and Proxy Management. These portlets used the portal-compatible Java CoG 4.x series to provide support for Globus Toolkit version 2.4 and 4.0 releases. Versions 1.0.0, 1.0.1, and 1.0.2 were released this year. Additional information is available from http://www.collab-ogce.org/ogce2/ogce2-1.0-download.html.

Version 2.0 Release

Led by the TACC and Indiana teams, we completed a major revision of our installation and packaging system. The current OGCE2 2.0 release includes all software (both OGCE, GridPort, and third party tools such as Sakai and GridSphere) needed to set up a Grid portal and Science Gateway in a single installation package. In addition to the core Globus Toolkit portlets we have developed in OGCE2 1.0, the 2.0 release also contains several new portlets including

- Storage Resource Broker portlet:
- Condor portlets: these work with Condor 6.7's BirdBath SOAP/WSDL enabled services.
- Sakai bridge portlets: these use Web Services to manage connections between the Sakai portlet and remote Sakai services.

PURSe Portlets Release 1.0

The PURSe Portlet system works as a client interface to the NMI PURSe Grid credential account management system. These portlets were designed using Java Server Faces as part of our Grid Faces project. Additional information is available here: http://www.extreme.indiana.edu/portals/purse-portlets/index.html.

OGCE Web Site Revision

TACC and Indiana team members completed a complete revision of the OGCE website, www.collab-ogce.org. The revised site is built with the Maven 1.0.2 XDoc tool. This allows us to easily distribute the task of managing, updating, and controlling revisions to the content of the Web site, which has been checked into the OGCE CVS repository.

Technology Evaluation

Both TACC and IU continued to evaluate JSR168 containers (Pluto, GridSphere, uPortal, Sakai). In addition, SDSU has begun working with Python tools in response to direct requests from the Fusion community, who have begun developing python tools. SDSU continues to work with AJAX technologies.

4.2.2 Task 3: Grid and Web Services

The web services developed for the project during this phase of the work are described above. In general, many of the portlets and gateways that the team worked include accessing web services. This includes the GPIR and SRB portlets, the LEAD and Fusion grid systems and others.

4.2.3 Task 4: Applications - FusionGrid Portal

As part of our work with related technologies, we began investigations into portal systems that use technologies such as Python and new security technologies such as PubCookie. The SDSU team investigated using existing python frameworks such as Zope, Django, and Turbogears for portals and using python for web services. After working with the 3 frameworks above, we have concluded that currently, python frameworks are more oriented towards content management, and this does not meet the needs and flexibility requirements of a science grid portals. SDSU has begun to work on a simple portal framework.

The GAT team, in collaboration with LBL and SDSU began looking at an integrated security solutions for account identity management (see below), which can be used both on the FusionGrid and the portal. This can be a solution for any portal framework, including python and java. As a result, it will be integrated into the OGCE (see Future Plans below).

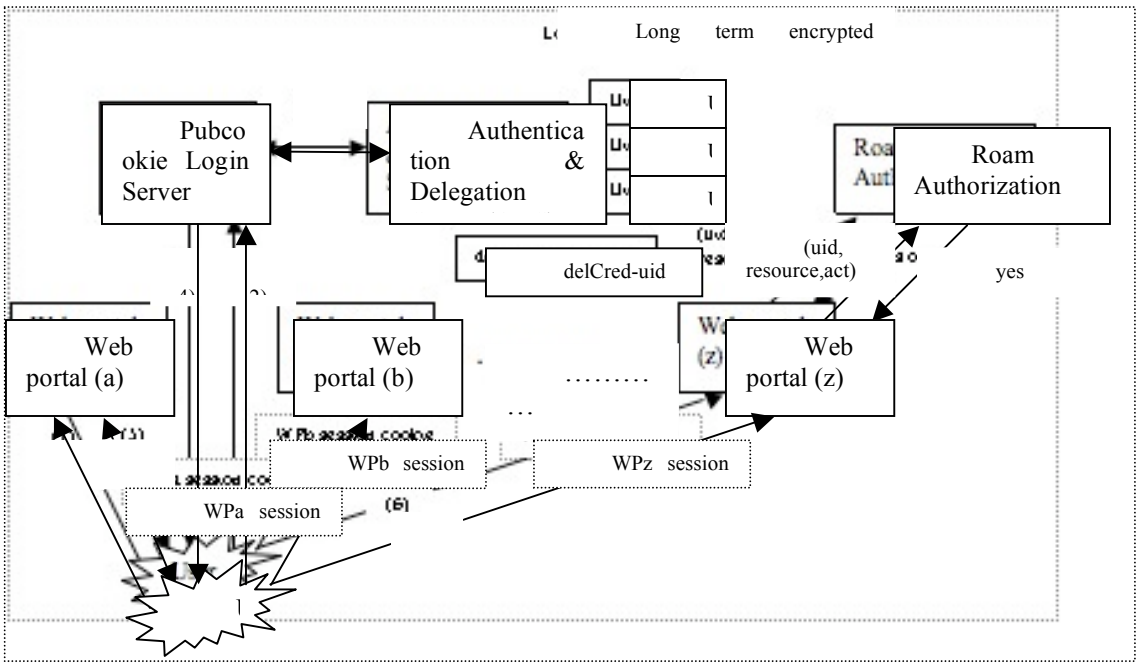


Figure 12. Federated Portal Architecture. (Courtesy: LBNL docs on the federated portal architecture). Note: WP: Web Portal. WPa …WPz illustrate the web portals in the FusionGrid domain.

Web Portals/Application Server Architecture

A web portal may provide both public and restricted services and information. When required, following are the steps (refer to fig.1) taken to authenticate and authorize a user:

First contact (WPa):

(1) When a user requests web portal (a) for a URL tied to an application that uses Pubcookie for authentication, the Pubcookie framework on the app server intercepts the request and generates a redirect response.
(2) The 'redirect' response causes the user's browser to make a request to the Pubcookie login server.
(3) The Pubcookie login server requires the user to sign-on with Fusion-Grid credentials. User is authenticated with the ADS server and a short-term proxy is delegated from stored long-term credentials and placed in the secondary MyProxy server for subsequent portal retrieval.
(4) Upon successful verification , the user is sent a 'redirect' page and a login cookie to be used on any subsequent visits by the user to the login server(single sign-on ability)
(5) The 'redirect' page causes the user's browser to re-request the original resource (URL) on the application server. The app server now successfully authenticates the user sending him the original resource requested. It also generates a WPa session cookie for subsequent requests by the user to the application. Subsequent contact (WPb…WPz)
(6) Receives the cookie and access to web resource is granted. If requested action requires a certificate, the gridId is retrieved from the Apache environment and the app server's host certificate is used to retrieve a proxy. WPb…WPz may call ROAM with gridId, resource, action to authorize an action.

Pubcookie Web-Based Authentication Framework

To demonstrate Pubcookie's web-based authentication framework, a pubcookie protected and an ssl-enabled (server side authentication with X.509 credentials) apache server was set-up on a test system at GA, namely star11. This federated web-portal set-up leveraged off the Credential Manager (residing at the host cert.fusiongrid.org) to demonstrate the use of Pub-cookie's single-sign ability, with FusionGrid id, to a set of portals in the fusiongrid.org domain. The set-up also used Pubcookie authentication for basic web access and a host certificate to retrieve a proxy for job submission on the FusionGrid.

- *TestRun1*- Demonstrated the logon framework for access to a Pubcookie protected resource/web-page :
- When a user, via his browser, initially attempts access to the protected URL https://star11.fusiongrid.org/restricted/ and is not authenticated, they are redirected to the Pubcookie login server running at https://cert.fusiongrid.org/. Upon successful authentication with the user's FusionGrid id, they are redirected back to the originally requested URL.
- TestRun2- Demonstrated use of the single sign-on Pubcookie architecture for FusionGrid job submissions via a web interface.
- When a user accesses https://star11.fusiongrid.org/secured/ to submit a job to FGTestApp (a demo FusionGrid application on star12), the appserver on star11 retrieves a delegated short-term proxy from the MyProxy server using the appserver's host certificate. The proxy thus retrieved is used for the grid job submission to FGTestApp.
- TestRun3- Demonstrated the Pubcookie's single sign-on ability to unify authenticated access across portals in the fusiongrid.org domain. The two sites used in this test were: https://cert.fusiongrid.org/admin and https://star11.fusiongrid.org/secured.

Related Applications Work

In addition to working with the Fusion Grid team, the DOE Portals effort has strong interactions and collaborations with large grid projects and the NSF NMI OGCE project. As a result, the portal software is in use by several large portal projects including:

- **Teragrid**: The OGCE project has provided portal frameworks and tools that are used extensively throughout the Teragrid Gateway project. This includes the LEAD gateway, the Teragrid User Portal, the RENCI Bioportal and the Texas Flood modeling gateway.

- **LEAD**: The OGCE has provided the foundation for the LEAD Portal that is being deployed as part of the "Linked Environments for Atmospheric Discovery – an NSF large ITR. This project is building tools that are designed to enable adaptive, real-time prediction of Tornados and Hurricanes and a platform for experimental research in atmosphic science.

4.3 Collaboration, Events, Outreach

Activity	Description
GCE Working Group Meeting at GGF 14, June 26-30, 2005.	OGCE members reorganized the GCE working group meeting for GGF 14. Our primary activity was planning the upcoming GCE Workshop at Supercomputing 2005. See http://www.gridforum.org/ggf_events_ggf14.htm for additional information.
Science Gateways Meeting at GGF 14 , June 26-30, 2005	Several OGCE members participated in the Science Gateways workshop. These included both presentations by M. Pierce on the OGCE project and presentations by OGCE-supported portals such as LEAD and the North Carolina Bioportal effort. Additional information is available from http://www.gridforum.org/GGF14/ggf_events_next_schedule_Gateways.htm.
Science Portals Workshop at GGF 15, October 3-6, 2005	OGCE members M. Pierce, D. Gannon, C. Severance, and J. Futrelle organized the workshop, "New Technologies for Science Portals." The workshop included 5 presentations and supplemental material. Additional information is available from http://www.collab-ogce.org/GGF15Workshop/.
Richard Tapia Conference, October 19-22 2005	M. Pierce gave a tutorial on portals and portal software in the "Cyberinfrastructure" workshop of the 2005 Tapia Conference. For additional information, see "Presentations" below and http://www.ncsa.uiuc.edu/Conferences/Tapia2005/workshop_abstracts_and_bios.html.
GCE Workshop at Supercomputing November 18, 2005.	OGCE members M. Thomas and M. Pierce helped organized the first GCE Portals Workshop, and several OGCE members assisted as technical committee members. The workshop included five presentations in the plenary session and sixteen poster presentations. The workshop was attended by over 50 participants. Additional information is available from http://pipeline0.acel.sdsu.edu/mtgs/gce05/.
I533 Indiana University Seminar on Chemical Informatics	M. Pierce gave a series of three lectures on Grid computing, science portals, and Web Services for IU's I533 seminar course. Additional information is available from http://www.indiana.edu/~cheminfo/I533/533home.html.
Teragrid 2006	Participation in the Gateways tutorials at the Teragrid 2006 conference, Indianapolis June 2006
SC06 Workshops Chair	Thomas is serving as the Chair for SC06 Workshops. She is in charge of the program committee and organizing the program schedule for SC.

4.4 Publications

4.4.1 Presentations and Posters (Chronological Order)

- M. P. Thomas, J. Burruss, L. Cinquini, G. Fox, D. Gannon, L. Gilbert, G. von Laszewski, K. Jackson, D. Middleton, R. Moore, M. Pierce, B. Plale, A. Rajasekar, R. Regno, E. Roberts, D. Schissel, A. Seth, and W. Schroeder Poster on Grid Portal Architectures for Scientific Applications for SciDAC 2005 meeting June 26-30 2005 San Francisco.
- Marcus Christie, LEAD Portal: a TeraGrid Gateway and Application Service Architecture, GGF 14 Science Gateways Workshop, June 28 2005, Westin Michigan Avenue, Chicago, Illinois.
- Marlon Pierce Grids and Portals for VLAB VLAB Workshop, University of Minnesota July 21-23.
- Marlon Pierce Portals, Portlets, and Clients to Grid Services IRIS Web Services Workshop, Monterey CA, September 21-23.
- Marlon Pierce Using AJAX at GGF 15 Portals Workshop, October 3-6 2005, Boston, MA.
- Marlon Pierce A Portal Architecture Review at GGF 15 Portals Workshop, October 3-6 2005, Boston, MA.
- Marlon Pierce Integrating Geographical Information Systems and Grid Applications at GGF 15 GIS Grid Workshop, October 3-6 2005, Boston, MA.
- Suresh Marru, Grid Computing for Real World Applications, Oklahoma Supercomputing Symposium 2005, October 5th 2005, University of Oklahoma at Norman.
- Marlon Pierce Building Web Portals as Science Gateways Richard Tapia Conference, October 19-22 2005, Albuquerque, NM.

- Dennis Gannon, "Predicting Tornados with Data Driven Workflows: Building a Service Oriented Grid Architecture for Mesoscale Meteorology Research", co-authored with Beth Plale, Microsoft e-Science workshop, Oct. 2005.
- OGCE Collaboration posters for SC05 November 12-18 2005 Seattle
    - OGCE Architecture: Portlets and Services for Science Gateways
    - OGCE Tools Support Collaboration, Data Management, and Application Development
    - OGCE Portal Applications for Grid Computing I
    - OGCE Portal Applications for Grid Computing II
- Dennis Gannon, Keynote Presentation, "Lessons Learned While Building Grids: a View of Some Research Challenges" co-authored with Beth Plale. International Grid Workshop 2005, held in conjunction with SC2005, Seattle, Nov. 2005.
- Mehmet Nacar, Mehmet Aktas, Marlon Pierce, Zhenyu Lu, Gordon Erlebacher, Dan Kigelman, Evan F. Bollig, Cesar De Silva, Benny Sowell, and David A. Yuen VLab: Collaborative Grid Services and Portals to Support Computational Material Science at GCE'05 Workshop on Grid Computing. Seattle, WA. November 18 2005.
- Marlon Pierce Building Web Portals as Science Gateways at I533 Seminar in Chemical Informatics: Molecular Informatics, the Data Grid, and an Introduction to eScience March 2 2006. Pierce also gave lectures on Grid computing and Web Services.
- Mehmet Nacar JSF Custom Grid Tags March 2006 VLAB Meeting, Florida State University.
- Marcus Christie, TeraGrid Science Gateways Tutorial; Using the LEAD Portal, Accelerating Research with Grid Computing, A Workshop for Researchers at Montana State University & Lariat-West Partner Institutions, April 14, 2006, Montana State University, Bozeman, Montana
- Dennis Gannon, Keynote Presentation, "Opportunities and Challenges for Future Generation Grid Research", the Annual Symposium on Advanced Computing Systems and Infrastructures, Osaka, Japan, May 2006.
- 21st IEEE/NPSS Symposium on Fusion Engineering 2005 (Knoxville, TN, September 26-29, 2005) D. P. Schissel, et al., Advances in Remote Participation for Fusion Experimental Facilities
- 10th International Conference on Accelerator and Large Experimental Physics Control Systems (ICALEPCS) (Geneva, Switzerland, 2005), M. Greenwald, et al., Visions for Data Management and Remote Collaboration on ITER [Oral
- 5th Workshop on Advanced Collaborative Environments (WACE) (September 8-9, 2005, Redmond, Washington), G. Abla, et al., Shared Display Wall Based Collaboration Environment in the Control Room of the DIII-D National Fusion Facility [Oral Presentation: PDF]

4.4.2 Books and Book Chapters

- Taylor, E. Deelman, D. Gannon, M. Shields, "Workflows for eScience: Adaptive Workflows for Mesoscale Meteorology", Springer Verlag. To appear Dec. 2006.
- D. Gannon, B. Plale, S. Marru, G. Kandaswamy, Y. Simmhan, and S. Shirasuna, Dynamic, Adaptive Workflows for Mesoscale Meteorology, To Appear in "Workflows for eScience: Scientific Workflows for Grids", Chapter 10,
- OGCE members D. Gannon and M. Pierce are currently editing a book on Grid Portals to be published by Elsevier. This book will include chapters from both OGCE members and collaborators. DOE Portal team contributions include chapters from Thomas (SDSU), Fox, Pierce, Gannon, Plale (Indiana); Boissea, Dahan, Roberts (Texas); and others.

4.4.3 Conference Publications (In chronological order)

- M. P. Thomas, J. Burruss, L. Cinquini, G. Fox, D. Gannon, L. Gilbert, G. von Laszewski, K. Jackson, D. Middleton, R. Moore, M. Pierce, B. Plale, A. Rajasekar, R. Regno, E. Roberts, D. Schissel, A. Seth, and W. Schroeder Grid Portal Architectures for Scientific Applications Proceedings of SciDAC 2005 meeting June 26-30 2005 San Francisco.

- Mehmet Nacar, Mehmet Aktas, Marlon Pierce, Zhenyu Lu, Gordon Erlebacher, Dan Kigelman, Evan F. Bollig, Cesar De Silva, Benny Sowell, and David A. Yuen VLab: Collaborative Grid Services and Portals to Support Computational Material Science GCE'05 Workshop on Grid Portals.at SC05 Seattle, WA. November 18 2005.
- Ahmet Fatih Mustacoglu, Wenjun Wu, and Geoffrey Fox Internet Calendaring and Scheduling Core Object Specification (iCalendar) Compatible Collaborative Calendar-Server (CCS) Web Services IEEE 2006 International Symposium on Collaborative Technologies and Systems CTS 2006 conference Las Vegas May 14-17 2006.
- Beth Plale, Dennis Gannon, Daniel A. Reed, Sara J. Graves, Kelvin Droegemeier, Bob Wilhelmson, Mohan Ramamurthy: Towards Dynamically Adaptive Weather Analysis and Forecasting in LEAD. International Conference on Computational Science (2) 2005: 624-631
- Yogesh Simmhan, Beth Plale, Dennis Gannon: A survey of data provenance in e-science. SIGMOD Record 34(3): 31-36 (2005)
- Sangmi Lee Pallickara, Beth Plale, Liang Fang, and Dennis Gannon, "End-to-End Trustworthy Data Access in Data-Oriented Scientific Computing", CCGrid 2006, Singapore, May, 2006.
- Y. L. Simmhan, B. Plale, D. Gannon, and S. Marru, Performance Evaluation of the Karma Provenance Framework for Scientific Workflows, To Appear in Lecture Notes in Computer Science, 2006 & International Provenance and Annotation Workshop (IPAW), 2006.
- Yogesh Simmhan, Beth Plale, Dennis Gannon: Towards a Quality Model for Effective Data Selection in Collaboratories. ICDE Workshops 2006: 72
- Wei Lu, Kennith Chiu, Aleksander Slominski, and Dennis Gannon, "A Streaming Validation Model for SOAP Digital Signature" The 14th IEEE International Symposium on High Performance Distributed Computing (HPDC-14), 2005.
- Beth Plale, Dennis Gannon, Daniel A. Reed, Sara J. Graves, Kelvin Droegemeier, Bob Wilhelmson, Mohan Ramamurthy, "Towards Dynamically Adaptive Weather Analysis and Forecasting in LEAD". International Conference on Computational Science (2) 2005: 624-631
- L. Fang and D. Gannon, "XPOLA: An Extensible Capability-based Authorization Infrastructure for Grids", 4th Annual PKI R\&D Workshop: Multiple Paths to Trust, NIST Gaithersburg, MD April 19-21, 2005
- Dennis Gannon, Beth Plale, Marcus Christie, Liang Fang, Yi Huang, Scott Jensen, Gopi Kandaswamy, Suresh Marru, Sangmi Lee Pallickara, Satoshi Shirasuna, Yogesh Simmhan, Aleksander Slominski, Yiming Sun, Service Oriented Architectures for Science Gateways on Grid Systems. ICSOC 2005: 21-32
- Sangmi Lee Pallickara, Beth Plale, Scott Jensen, Yiming Sun, Monitoring Access to Stateful Resources in Grid Environments. IEEE SCC 2005: 343-346
- Gopi Kandaswamy, and Dennis Gannon. A Mechanism for Creating Scientific Application Services On-demand from Workflows. Workshop on Web Services-based Grid Applications, to appear Aug 2006
- Yi Huang, Aleksander Slominski, Chathura Herath, and Dennis Gannon WS-Messenger: A Web Services based Messaging System for Service-Oriented Grid Computing. IEEE International Symposium on Cluster Computing and the Grid, Singapore, to appear May 2006
- Performance Evaluation of the Karma Provenance Framework for Scientific Workflows. Y.L. Simmhan, B. Plale, D. Gannon, & S. Marru, International Provenance and Annotation Workshop (IPAW), 2006
- A Framework for Collecting Provenance in Data-Centric Scientific Workflows. Y.L. Simmhan, B. Plale & D. Gannon. Accepted for the International Conference on Web Services (ICWS), to appear 2006.
- Srinath Perera, Dennis Gannon, Enabling Web Service Extensions for Scientific Workflows, The Workshop on Workflows in Support of Large-Scale Science, to appear Paris June, 2006.

- Wei Lu, Kenneth Chiu, and Dennis Gannon. Building a generic soap framework over binary xml. In The 15th IEEE International Symposium on High Performance Distributed Computing (HPDC-15), to appear June 2006.
- Hao Yin, Donald F. McMullen, Mehmet A. Nacar, Marlon Pierce1, Kianosh Huffman, Geoffrey Fox1, Yu Ma, "Providing Portlet-Based Client Access to CIMA-Enabled Crystallographic Instruments, Sensors, and Data." Submitted to GRID 2006.

4.4.4 Journal Publications and Technical Reports

- Dennis Gannon and Geoffrey Fox Workflow in Grid Systems Editorial of special issue of Concurrency&Computation: Practice&Experience based on GGF10 Berlin meeting.
- Beth Plale, Jay Alameda, Bob Wilhelmson, Dennis Gannon, Shawn Hampton, Al Rossi, Kelvin Droegemeier: Active Management of Scientific Data. IEEE Internet Computing 9(1): 27-34 (2005)
- K. Droegemeier, D. Gannon, D. Reed, B. Plale, J. Alameda, T. Baltzer, K. Brewster, R. Clark, B. Domenico, S. Graves, E. Joseph, D. Murray, R. Ramachandran, M. Ramamurthy, L. Ramakkrisshnan, J. Rushing, D. Webeer, R. Wilhelmson, A. Wilson, M. Xue, S. Yalda, "Service-Oriented Environments for Dynamically Interacting with Mesoscale Weather", CiSE, Computing in Science \& Engineering -- November 2005, vol. 7, no. 6, pp. 12-29.
- Beth Plale, Dennis Gannon, Yi Huang, Gopi Kandaswamy, Sangmi Lee Pallickara, and Aleksander Slominski, "Cooperating Services for Data-Driven Computational Experimentation", CiSE, Computing in Science \& Engineering -- September 2005, vol. 7 issue 5, pp. 34-43
- Gopi Kandaswamy, Dennis Gannon, Liang Fang, Yi Huang, Satoshi Shirasuna, Suresh Marru, "Building Web Services for Scientific Applications", IBM Journal of Research and Development, Vol 50, No. 2/3 March/May 2006.
- Jay Alameda, Marcus Christie, Geoffrey Fox, Joe Futrelle, Dennis Gannon, Mihael Hategan, Gregor von Laszewski, Mehmet A. Nacar, Marlon Pierce, Eric Roberts, Charles Severance, and Mary Thomas The Open Grid Computing Environments Collaboration: Portlets and Services for Science Gateways Accepted for publication in Concurrency and Computation: Practice and Experience Special Issue for Science Gateways GGF14 workshop.

4.5 Project Management

In 2006, the Fusion Grid underwent a slow-down because the DIII-D Tokamak, located at General Atomics, has been off line and undergoing significant repairs and upgrades. This caused critical delays in the project and impacted our ability to test and demonstrate software. In addition, personnel changes at GAT have resulted in the need to hire and train a new grid/web developer, which has had the typical impact and resultant slowdown. This has resulted in significant project delays and underfunding of the project budget.

The DOE Portals project was funded with a start and end date of September 1, 2002 and extended through August 31, 2006. This project involves the efforts of team members from San Diego State University, the University of Texas at Austin, Indiana University, the University of California at San Diego, and General Atomics of La Jolla, CA. The PI's have decided to request another no-cost extension and have determined that the following tasks would have the significant impact. This is discussed in the next section.

Extension Plans for 2007

The project PI's met during a workshop (Grid Enabled Problem Solving Environments) this Summer, discussed what tasks remain to be done and what would be of the highest benefit to the DOE Fusion and the SciDAC programs Since the Fusion community and the ITER programs are of high priority for the DOE, and will be the focus of the proposed work:

- Deploy *pubCookie* onto the fusiongrid domain. This will allow FusionGrid to successfully, in a secure way manage FusionGrid credentials, unite the many web servers within fusion into one virtual secure domain. (GAT/LBL)
- Use this implementation of *pubCookie* to make a web browser interface to FG computational services. This should be able to replace the client codes that are presently written in IDL. (GAT)
- Integrate pubCookie into the OGCE Framework which will impact both DOE, NSF, and other programs. (SDSU/GAT/LBL)
- Develop a *Python Grid Portal* framework using TurboGears (GAT/LBL/SDSU). This area of focus is the result of discussions with Fusion scientific community who are currently developing python tools and are interested in the faster integration path between common python tools.
- Demonstrate the *Collaborative IDL* tools during plasma operations. (GAT/IU): it is essential to test and demonstrate this software capability during actual operations. Tests between operations have been successful but we need to run this software during operations to evaluate the impact and usefulness of this approach. This will have impact on future ITER operations.

Note that there will be no further work on the SRB/MDSPlus because SDSC is working on a next generation data management tool, hence any futher work on developing new SRB related capabilities should be put on hold until the new tool is available.

Revised Budget

Most project participants have little funding remaining, with the exception of GAT. In order to facilitate the remaining tasks, a revised budget that will help support the teams involved in the tasks above has been proposed and is listed in the table below:

INSTITUTION	BUDGET	%
GAT	$97	74%
IU	$5	4%
SDSC	$0	0%
SDSU	$24	18%
TACC	$5	4%
Total	***$131***	***100%***

In discussions with the project Program Manager (Dr. Scott), we have determined that a no cost extension to this grant is reasonable and will allow the project to complete its goals and benefit the SciDAC-2 and ITER programs. We have begun the process of requesting an extension to this grant from August 31, 2006 to August 31, 2007.

Project Meetings

The project holds regular bi-weekly telcons and meets in person when attending related meetings. In addition, as part of the consortium effort, the portals team participated in several OGCE developer meetings where requirements for the OGCE Toolkits were defined and as a result, DOE needs were included in the design decisions and programming tools.

During the year, the team attended many professional meetings including the DOE SciDAC meeting, Supercomputing, High Performance Distributed Computing (HPDC), and others. This is evidenced in the number of presentations and publications listed below.

5. Year 5 (2006-2007) Annual Report

5.1 Project Highlights

- Open Grid Computing Environments (OGCE) release (June 2007)
- OGCE-2 project awarded for NSF Science Gateways

5.1.1 Deliverables Summary

The project was granted a final no-cost extension this year in order to spend out the budget and complete some tasks relevant to the Fusion community and the ITER programs, which are of high priority for the DOE. The approved, revised scope of work focused on Tasks 1 and 4.

- Deploy pubCookie onto the fusiongrid domain.
- Use of pubCookie to make a web browser interface to FG.

5.1.2 Problems/Roadblocks

There were several delays getting the final revised budget distributed to the active participants. Thi was due to changes in DOE program managers. As a result, GAT had some difficulty getting the sub award monies to LBL in time for the LBL team to participate. This impacted the IDL testing and the pubcookie work. In addition, the FusionGrid project was not renewed for the SciDAC-2 phase, which resulted in minimal personnel available to migrate and test software. However, the software developed as part of the portal systems and SRB systems is in use by many of the science gateway projects on the TeraGrid. Additionaly, in 2008, the TeraGrid and Open Science Grid (OSG) will evaluate integration and sharing of services. This will result in greater use of the DOE Portals software in the form of the next version of the OGCE portal framework and associated services.

5.2 Accomplishments

5.2.1 Task 1: Portal Systems

OGCE Framework:
During the extension period, there were two OGCE core portal releases.

- Version 2.0.4 (November, 2006) was configured to work using TeraGrid resources (including TeraGrid GridFTP, GRAM, and GPIR services). The PURSe portlets were also added to the main build.
- Version 2.1 (June 2007): This version included major revisions to the Condor portlet and added support for Condor-G.

Condor and Condor-G Portlets

Updates to the Condor Web Service interface ("Birdbath") in Condor's 6.8 production series were incompatible clients based on the 6.7 research and development series. IU rewrote the OGCE Condor portlet from scratch using Java Server Faces and extended this to provide support for Condor-G. This acts as a universal Grid client.

Technology Evaluations

SDSU completed its investigations into the use of Python CMS frameworks as candidates for grid portals. Frameworks tested included Zope, Django, and others. Conclusion: content management systems are complex, with data base systems not easily modified/altered and thus not useful for programming environments needed to develop useful HPC grid portals.

Python Tools: Investigated the use of Python based WSRF services and integration with OGCE portlets. The use of the NMI PyWSRF and pyGlobus toolkits proved to be flexible and easy to use, with easy integration behind both python and java based portlets. In addition, SDSU worked with the LBL development team to provide valuable feedback and debugging of the software.

Pubcookie

The implementation of PubCookie was used to create a demonstration web browser interface to the FusionGrid computational service TRANSP. Additionally, this work was used to create a demonstration interface for a computational service as part of the SWIM (Center of Simulation of Wave Interactions with MHD) SciDAC-2 Project.

5.2.2 Task 2: Grid and Web Services

See Section 5.3 below.

5.2.3 Task 3: Data Collection Management

No significant activity on this task during this year.

5.2.4 Task 4: Applications

5.3 FusionGrid

A pubCookie single-signon server was integrated with the a MyProxy server. The resulting Credential Manager allows a user to login once in a 24-hr period at the Credential Manager portal and to establish a short-term (24 hr) proxy that can be used for authentication and authorization to run the TRANSP code at PPPL. If desired, the user can also specify that a longer term (2-week) proxy be created that can be used by the TRANSP server to renew the shorter term proxy until the TRANSP job is completed. This service was being used for some PPPL TRANSP production runs by 2QT 2006. The Credential Manager service was transferred from LBNL to PPPL by May 2007.

5.4 Collaboration, Events, Outreach

OGCE team members gave frequent presentations and lectures, including Supercomputing, the TeraGrid annual meeting, and many other symposia and workshops. These are described below in the section on Presentations and posters, which includes tutorials.

SDSU (Thomas) taught a graduate level course in advanced grid computing infrastructure. Course was based on developing pyWSRF-based mathematical services that run on the TeraGrid, behind the OGCE portal, and used TG security services as well.

In addition, the OGCE team organized a second and third Grid Computing Environments workshops. Grid Computing Environments 2005 workshop with the even larger GCE 2006, held in Tampa in conjunction with Supercomputing 2006. The workshop consisted of peer-reviewed papers from an international group of authors over two days. More information is available from the GCE 06 Web Site: http://wiki.cogkit.org/index.php/GCE06.

GCE 07 was also planned during the extension period. We followed the same format at GCE 06, with two days of peer-reviewed papers. More information on the workshop is available from http://www.collab-ogce.org/gce07. Finally, the proceedings for GCE05 were published in *Concurrency and Computation: Practice and Experience*, Volume 19, Issue 12. For the table of contents, see http://www3.interscience.wiley.com/journal/114298946/issue.

5.5 Publications

5.5.1 Presentations and Posters (Chronological Order)

- Mehmet A. Nacar, Jong Youl Choi, Marlon Pierce, Geoffrey Fox, "Building Grid Portals with OGCE: Big Red Portal and GTLAB" Presentation at TeraGrid 2007 Conference Madison Wisconsin June 5 2007.
- Marlon Pierce, "OGCE Short Summary," Software Provider Forum at TeraGrid 2007 Conference Madison Wisconsin June 5 2007.
- Marlon Pierce and Jim Myers Gateways and Cyberenvironments and Web 2.0 Birds of a Feather at TeraGrid 2007 Conference Madison Wisconsin June 5 2007.

- Geoffrey Fox and Marlon Pierce, "Web 2.0 in a Web Services and Grid Context Part I: CTS2007 Web 2.0 Tutorial," Part I of Tutorial at 2007 International Symposium on Collaborative Technologies and Systems (CTS 2007) May 21 2007.
- Marlon Pierce and Geoffrey Fox, "Web 2.0 Tutorial: Part 2," Part II of Tutorial at 2007 International Symposium on Collaborative Technologies and Systems (CTS 2007) May 21 2007.
- Marlon Pierce, "Web Service Foundations: WSDL and SOAP," I590 Class IUPUI April 5 2007.
- Marlon Pierce, "Open Grid Computing Environments" at OGF19 Software Developers Track, Friday Center Chapel Hill NC January 29 2007.
- Marlon Pierce, Jong Youl Choi OGCE Portal Software for Big Red and the TeraGrid Booth Presentation at SC06 Tampa Florida, November 12-17 2006.
- Mehmet Nacar, Marlon Pierce, Gordon Erlebacher, Geoffrey Fox, "Designing Grid Tag Libraries and Grid Beans." Second International Workshop on Grid Computing Environments GCE06 at SC06, Tampa, FL. Nov. 12-13 2006.
- Hao Yin, Sofia Brenes-Barahona, Donald F. McMullen, Marlon Pierce, Kianosh Huffman, Geoffrey Fox, "A PERMIS-based Authorization Solution between Portlets and Back-end Web Services." Second International Workshop on Grid Computing Environments GCE06 at SC06, Tampa, FL. Nov. 12-13 2006.
- Marlon Pierce, Dennis Gannon Charles Severance, Gregor von Laszewski, Mary Thomas, and Eric Roberts, "OGCE Briefing to NSF OCI", August 22 2006
- Mary Thomas, José Castillo, Germán Larrazábal, Carlos Torres The General Curvilinear Ocean Model (GCOM) Cyberinfrastructure Environment, Presented at the annual SDSU CSRC HPC Symposium, April, 2007.
- Mary Thomas. SDSU Computational Sciences and Cyberinfrastructure Research Grid (CSCRGrid). Presented to SDSU Management (Deans, Provost, VP of Research, etc.), and at the kickoff meeting for the SDSU Research Grid Project, 2007.

5.5.2 Books and Book Chapters

- Geoffrey C. Fox, Rajarshi Guha, Donald F. McMullen, Ahmet Fatih Mustacoglu, Marlon E. Pierce, Ahmet E. Topcu, and David J. Wild, "Web 2.0 for Grids and e-Science," INGRID 2007 - Instrumenting the Grid 2nd International Workshop on Distributed Cooperative Laboratories - S.Margherita Ligure Portofino, ITALY, April 18 2007.
- Marlon E. Pierce, Geoffrey Fox, Huapeng Yuan, and Yu Deng, "Cyberinfrastructure and Web 2.0", Proceedings of HPC2006 July 4 2006 Cetraro Italy.

5.5.3 Conference Publications

- Security on the US Fusion Grid, J. R. Burruss, T.W. Fredian, M.R. Thompson, Fifth International Atomic Energy Agency (IAEA) Technical Meeting onControl, Data Acquisition, and Remote Participation for Fusion Research, July 2005 Budapest, Hungary; Fusion Engineering and Design, Elesevier, to appear Fall 2006, LBNL-58693
- Mehmet A. Nacar, Marlon E. Pierce, Geoffrey C. Fox GTLAB: Grid Tag Libraries Supporting Workflows within Science Gateways 3rd International Conference on Semantics, Knowledge and Grid SKG2007 Xian China October 28-30 2007.
- Mehmet A. Nacar, Jong Y. Choi, Marlon E. Pierce, and Geoffrey C. Fox, "Building a Grid Portal for Teragrid's Big Red," Proceedings of TeraGrid 2007 Conference Madison Wisconsin June 4-8 2007.
- Mehmet Nacar, Marlon Pierce, Gordon Erlebacher, Geoffrey Fox. Designing Grid Tag Libraries and Grid Beans Second International Workshop on Grid Computing Environments GCE06 at SC06, Tampa, FL. Nov. 12-13 2006.
- Hao Yin, Sofia Brenes-Barahona, Donald F. McMullen, Marlon Pierce, Kianosh Huffman, Geoffrey Fox, A PERMIS-based Authorization Solution between Portlets and Back-end Web

Services. Second International Workshop on Grid Computing Environments GCE06 at SC06, Tampa, FL. Nov. 12-13 2006.

5.5.4 Journal Publications and Technical Reports

- Mehmet A. Nacar, Mehmet S. Aktas, Marlon E. Pierce, Zhenyu Lu, Gordon Erlebacher, Dan Kigelman, Evan F. Bollig, Cesar R. S. da Silva, Benny Sowell, David A. Yuen: VLab: collaborative Grid services and portals to support computational material science. Concurrency and Computation: Practice and Experience 19(12): 1717-1728 (2007).
- Jay Alameda, Marcus Christie, Geoffrey Fox, Joe Futrelle, Dennis Gannon, Mihael Hategan, Gopi Kandaswamy, Gregor von Laszewski, Mehmet A. Nacar, Marlon E. Pierce, Eric Roberts, Charles Severance, Mary Thomas: The Open Grid Computing Environments collaboration: portlets and services for science gateways. Concurrency and Computation: Practice and Experience 19(6): 921-942 (2007).
- Hao Yin, Donald F. McMullen, Mehmet A. Nacar, Marlon E. Pierce, Kianosh Huffman, Geoffrey Fox, Yu Ma: Providing Portlet-Based Client Access to CIMA-Enabled Crystallographic Instruments, Sensors, and Data. GRID 2006: 317-318.
- M. Thomas. GCE 2005: Workshop on Grid Computing Portals. Editorial, Concurrency and Computation: Practice and Experience, Volume 19 Issue 12 , Pages 1563 - 1748 (25 August 2007.
- M. Thomas. Investigations into the Parallelization of the General Curvilinear Ocean Model (GCOM) and its Operation Across Cyberinfrastructure Environments. Internal report, SDSU Computational Sciences Research Center. January, 2007.

5.6 Project Management

During this final year, the primary participants were SDSU, General Atomics, and Lawrence Berkely Laboratories (as subcontract to GAT).

For further information on this project, contact:

Mary P. Thomas
San Diego State University
Department of Computer Science
San Diego, CA 92181
Phone: 619-594-1694
E-mail: mthomas@sciences.sdsu.edu

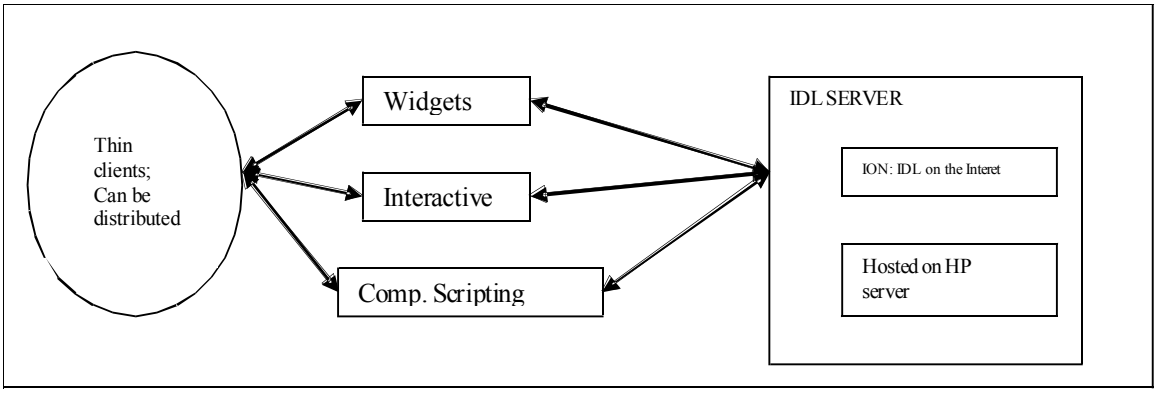
Appendix A. DOE Portals All Hands Meeting Notes

February 4-5, 2004
San Diego Supercomputer Center/General Atomics

- Note: the full report of the meeting can be found on the project website:
  - www.doeportals.org
- Agenda: Weds:
  - 9:00 am: Opening Remarks (Mary)
  - 9:15 -10:15: Collaborator Updates: TACC, IU, SDSC, GA
  - 10:15 - 10:30: Break
  - 10:30 - 11:30: NMI Portals/OGCE
  - 11:30 - 11:40: Data grid federation technology (Reagan)
  - 11:40: SDSC Computer Room Tour
  - 12:00 - 1:00: lunch
  - 1:00 - 2:00: Open discussion
  - 2:00 – 4:00: Fusion Portal
  - 4:00 pm: drive to GA for tour
  - 4:30 - 5:30 pm: Fusion facility tour
- Agenda: Thurs: dev discussions/open

Fusion discussion Notes: Thursday

- Talking about between shots
  - Things that MDS+ does not do right…
  - Moving data between shots more efficiently
- Computer System Requirements
  - Meet with Kristie
- Geoffrey wants to lean about review+
  - GA Uses IDL and Review+, JET Uses SCOPE for graphical package but not as sophisticated
  - Need for a central meta data catalogue or name translation system / concept space
  - ED: Electronic Display:
  - Store session, context data in a relational database
  - Grabs data and events from MDS+
- Security issues:
  - Need to meet with GA and explain the GP3 account model
- Question of visualization services needed:
  - Some portals out there: SCEC will have, S. Cutcheons has GridPort based portal;
  - How to use IDL?



PreTRANSP

- Replicating PreTRANSP as a portlet
  - Widgets
  - Job submission
  - MDS+ access (via the MDS+ java API)
  - Run database interaction
  - User configuration
  - FGM interaction
- build portlet on top of one or more services)
  - Between-shot info
  - Sharing review plus (using video codec or shared event model)
- PreTRANSP code details
  - Ufile extension: name of trigraph node to put in MDS+
  - Prefix may have 2 letters
  - Mds put (path, TDI expression)
  - Build_signal(build_with_units($ “”), *, build_with_units(... (see code)

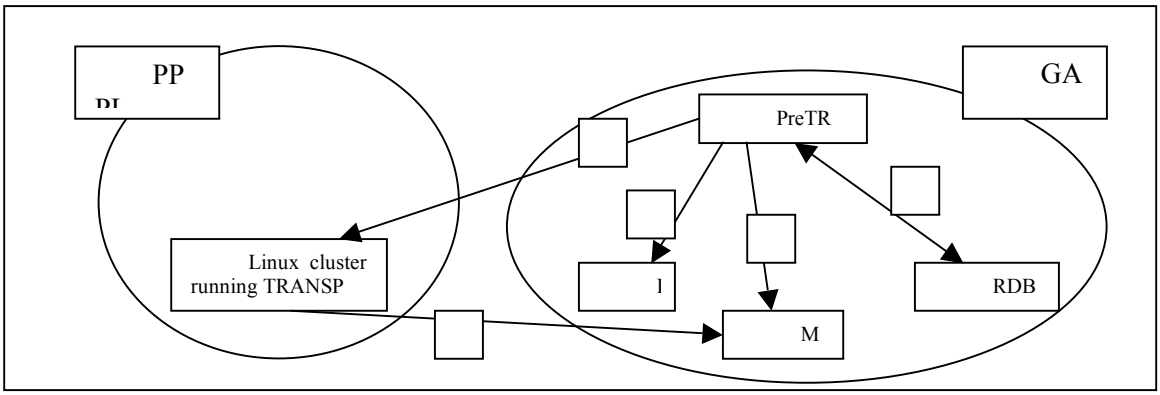


Figure 13. Diagram of where things are running (see related ppt files)

- PreTRANSP Steps:

- Zero: on connection, determine site, tree for subtype of coderun
1. PreTRANSP interacts with the run management database to load a user's run
2. PreTRANSP loads the user's input files into MDS+
3. PreTRANSP submits the job on behalf of the user to the PPPL cluster using the GT2 GSI security model
4. PreTRANSP alerts the FGM that the job has been submitted
5. TRANSP (not PreTRANSP) reads the user's input files from MDS+

- Notes:
  - Connect to right MDS+ configuration for each site/and data tree
  - Communication across the circles is through firewalls…only allowed from known IP addresses
  - *Sendpost* (to FGM): using old version of java
  - Need a mapping on MDS+ server
  - There was a firewall issue present for job submission?
  - Unique data in each RDBMS (there's an rdbms at MIT as well)
  - D3D – GA; CMOD – MIT
    - Each site reads config data from MDS+, and each sites has unique set of steps, and each step will have its own series of IDL widgets
    - Portal will pull steps from MDS+ and emulate widgets where possible.
    - Name of machine, type of coderun (lightweight D3D), tree type, connection type, etc.
  - Logbook, code run db schemas the same
  - The following is the functionality needed by a preTRANSP portlet:
    - User configuration
    - Proxy management
    - Uploading
    - Accessing
    - Run loading (manipulation of the run metadata db)
    - Input files
    - Parsing
    - Loading into MDS+
    - Job submission (GT2)
    - Interaction with FGM
- Services-based:
  - Most important to get the services that can be re-used by other software (i.e. other portlets, applications, etc.)
  - Some of these can be abstracted as services.
  - Most probably should if for nothing else to make the portlet itself (and thus portal) as lightweight as possible.
  - Having this portlet access services, if they are remote, also allows the work to be distributed among compute resources.
- Security Challenges
  - Need secure web services solution: SOAP, Kerberos, other solutions (IU/Gannon?)
  - Mary Thompson is DOE credential manager
  - What about use of MyProxy?
  - 7 day short term Kerberos tickets (?)
  - Form a team

ACTION ITEMS:

1. Hardware/resource identification (GA/TACC)
2. Collaborative Sharing (IU)
    a. RT display of between shot information
    b. Develop new event model
3. Security Solution
    a. Form a team:
        i. SDSC: Wayne Schroeder
        ii. IU: Marlon Pierce
        iii. TACC: Eric or Akhil
        iv. GA: Justin
    b. Need secure web services solution
4. Fusion Portal (All)
    a. Review/Define use Requirements, features
    b. Revise/Update list of Technologies
    c. Revise/Update architecture
    d. Schedule for deliverables
    e. team assignments
    f. PRETRANSP Web Portal (TACC/GA)
        i. Keep as lightweight as possible
        ii. job submission
        iii. replicate 'PT' functions in java
        iv. Accounts (TACC/GA)
        v. Need to create a plan and arch doc.
        vi. A PreTRANS service using a CODERUN service

Appendix B. SRB/Jargon Command Listing

The following SRB Scommand (shell command) equivalents are currently available in the Java API:

Sannotate - SRBFile.modifyMetaData
Sappend - SRBRandomAccessFile seek and write
Sattrs - javadoc
Sbload - LocalFile.copyTo(SRBFile) and SRBFile.copyFrom(LocalFile)
Scat - SRBRandomAccessFile.read
Scd - SRBFile constructor
Schmod - SRBFile.changePermissions
Scp - SRBFile.copyTo
Senv - SRBAccount setters and getters
Serror - SRBException
Sexit - SRBFileSystem finalize
Sget - SRBFile.copyTo
SgetColl - SRBFile.query
SgetD - SRBFile.query
SgetR - SRBFileSystem.query
SgetT - SRBFileSystem.query
SgetU - SRBFileSystem.query
Shelp - javadoc
Sinit - SRBFileSystem constructor
Sls - SRBFile.list
Slscont - SRBFileSystem.query
Smeta - SRBFileSystem.query, SRBFile.query and SRBFile.modifyMetaData
Smkcont - SRBContainer.createNewFile
Smkdir - SRBFile.mkdir
SmodColl - SRBFile.modifyMetaData
SmodD - SRBFile.modifyMetaData
Smv - SRBFile.renameTo
Spasswd - SRBFileSystem.srbModifyUser
Spcommand - SRBFileSystem.executeProxyCommand
Sphymove - SRBFile.setResource
Sput - SRBFile.copyFro
Spwd - SRBFile.getAbsolutePath
Sreplcont - SRBContainer.replicate
Sreplicate - SRBFile.replicate
Srm - SRBFile.delete
Srmcont - SRBContainer.delete
Srmdir - SRBFile.delete
Stls - SRBFileSystem.query
Stoken - SRBFileSystem.query

For comparison, the SRB Scommands C library client does the following:

Sannotate - Annotate SRB Objects and display annotations
Sappend - Appends a local file, standard in, or a SRB object to a target SRB object
Sattrs - Display meta attributes used in conditions
Sbload - imports in bulk one or more local files and/or directories into SRB space.

Scat - Streams SRB Objects to standard output
Scd - Changes working SRB collection
Schmod - Modify access permissions for SRB objects and collections
Scp - Copies an object to a new collection in SRB space
Senv - Displays environmental file content
Serror - Display Error Information from Error Number
Sexit - Clears environmental files created during SRB operation
Sget - Exports SRB objects into local file system
SgetColl - Displays information about SRB collections
SgetD - Displays information about SRB files (objects)
SgetR - Displays information about SRB resources
SgetT - Displays information about SRB tickets
SgetU - Displays information about SRB users
Shelp - Displays one-line help messages for all Scommands
Sinit - Initializes environmental files for SRB operation
Sls - Lists SRB objects and collections
Slscont - Lists SRB containers
Smeta - Displays and ingests meta data values
Smkcont - Creates a new SRB container
Smkdir - Creates a new SRB collection
SmodColl - Modifies some system metadata information on SRB collections
SmodD - Modifies some system metadata information on SRB objects
Smv - Changes the collection for objects in SRB space
Spasswd - Change SRB (Text) Password
Spcommand - Execute proxy command in SRB
Sphymove - Moves a SRB object to a new SRB resource
Sput - Imports local files into SRB space
Spwd - Displays current working SRB collection
Sregister - Registers an object as a SRB object in MCAT
Sreplcont - Replicates an existing SRB Container
Sreplicate - Replicates an existing SRB object in a (possibly) new resource
Srm - Removes SRB objects
Srmcont - Removes SRB container
Srmdir - Removes an existing SRB collection
Srmticket - Remove a ticket
Ssyncd - Synchronizes copies of an SRB object
Ssyncont - Synchronize SRB containers
Stcat - Streams ticketed SRB Objects to standard output
Sticket - Issue tickets for SRB objects and collections
Stls - Lists ticketed SRB objects and collections
Stoken - Displays information about SRB-aware types.

More information and an expanded list of functionality is provided in the JARGON API javadocs, website and example programs. http://www.npaci.edu/DICE/SRB/jargon/

Appendix C. Fusion Portal Features & services:

The following list describes the set of features and services that will be used for the Fusion Portal.

a. Managing certs/proxies → online account management
    i. Certificates are a PIA
b. PreTRANSP
    i. Replace current IDL GUI with a portal no more IDL for clients
    ii. TRANSP only runs at Princeton
    iii. Run management database
    iv. Presentation
    v. Code run database
c. AG Interface
    i. Attend morning 8:05 meeting through a web portal
    ii. Hardware needs
d. Fusion Grid Authorization
    i. Akenti has been shut down, so may want an admin portal in front of this.
    ii. Fusion will rebuild, could use GPIR admin code possibly
    iii. Each service will have its own db for access
    iv. Codes have access lists, and services have access lists
    v. Fusion wants a web page for admin

e. File and Data management
    i. MDS+ Node Management
        1. listing/traversal of nodes
        2. viewing plots of data
        3. SRB
            a. can use SRB as the MDS+ web service
            b. what is win with SRB? Adding layer of complexity and what do we gain?
            c. Dictionary based on SRB metadata capability.
            d. Is it really needed for pretransp?
    ii. Grid FTP
        1. not really needed by Fusion because they don't really move files directly
        2. need to think about what benefits GridFTP can bring
f. Fusion grid monitor – not really needed at this time as a web service
g. http://physics.gat.com
    i. Can we give them internet/extranet capability
h. OGCE
i. Running pretransp

Appendix D. Fusion Web Services Plan (March, 2004)

Proposed Web Services Suite (from DOE AHM):

1) Events and Messages
   a) MDS+ has events and batch queuing
   b) DAQ (Data AcQuisition) events
   c) Analysis events
   d) IU look at Events/Narada: WS-Notification → WS-Event → NARADA
2) Data:
   a) SRB Web Services
   b) MDS+
      i) Need to define what this means and how to do this
3) Informational
   a) GPIR
   b) Portal monitoring
   c) Fusion Grid Monitor (FGM)
      i) what features and capabilities belong in the portal
   d) Data Analysis Monitoring System (DAM)
   e) MDS (MDS+?) status
   f) Other existing FG tools
   g) PreTRANSP web service
   h) DAM/FGM data into GPIR, convert DAM/FGM
      i) Data currently is saved in RDB
      ii) Used to go back and get timing
4) Web service to broker for IDL events for ReviewPlus in order to share ReviewPlus sessions

Appendix E. Fusion Portal Hardware Plan (March, 2004)

Fusion Portal Hardware Plan

The goal is to move development of portal from external test sites to internal Fusion machines to work with real infrastructure and to host a production portal by mid-Summer so it can be in use before the end of the D3D Fusion experimental year.

ACTION	COMPLETION
• Identify hardware on which to install portal software and services	April 30, 2004
° Install core software for all services	April 30, 2004
° Install OGCE, GP3, SRB on portal machine (dependent on release dates)	May 31, 2004
° Obtain security approval and test	May 31, 2004
• Define web services	April 30, 2004
° Install core software needed	May 15, 2004
° Install GPIR web service (test version)	May 31, 2004
° Install SRB web services (test version)	May 31, 2004
° Install other services TBD	
• Fusion TRANSP Portal (demo)	June 30, 2004
° portlet interface for TRANSP	June 30, 2004
° portlet interface for SRB	June 30, 2004
° online portal account creation	June 15, 2004
° other portlets/interfaces TBD	
• Production Fusion Portal in operation	July 15, 2004

General Notes:

- Need to host OGCE and GP3. both require authentication/credential mgmt
- Hosting web services (new SRB web services)
  - CVS
  - SRB services
  - GPIR
  - Others as needed
- GA currently hosts web services on a couple of hosts, but they are production
  - Apache/linux; Has 25 GB, 13 available
  - MSFT IIS web server/XP
  - Need development system asap:
    - Will have a couple of linux nodes, dual processor machines; 450 MHz Pentium 2or3, dual processor, 1GB RAM. This seems likely and doable to team.
    - Only has 2 GBytes; need more like 15GB, but can NFS mount more.
- Set up web portal with CVS repository on GA host:
  - Hydra is set up for CVS, should be able to add a repository to that
- Data Analysis Monitoring System (DAM) – add this as web service?
- May be another place to host services – but need to identify
- GP3 installation and demo portal requirements:
  - OS: Linux (tested on linux 7.3, 9)
  - Java 1.4.1 or higher
  - Ant 1.5 or higher
  - RDBMS: PostgreSQL 7.3 or higher – requires ~ 15GB disk space
  - GridPort services server: JBoss 3.2.x with jetty 4.2.x – uses port 8080, 8090-8095
  - Demo Portal web container: Tomcat 4.1.x – uses port 9080, easily configurable
  - GPIR needs information providers running on remote resources ingesting to GPIR
  - GT3 (3.0.2) on remote resources – highly memory intensive

- Community Schedule Framework uses GT3, is a grid service, optional use for general job submission
- MyProxy server needs to run on Fusion Grid

- OGCE portal
  - Current installation size is 95MB but expected to be much smaller after next release
  - includes all libraries and src code for the portlets, gridport3 client code(after April 30), and the tomcat web server.