

SciDAC DataGrid Middleware
A High-Performance Data Grid Toolkit:
Enabling Technology for Wide Area Data-Intensive Applications
Quarterly Report January 2002 thru March 2002

Accomplishments this Quarter:

- ***Conducted expanded, in-depth Developers Tutorials in Edinburgh and Chicago***
The tutorials were a tremendous success with over 150 attendees in Edinburgh, Scotland and over 200 at Argonne National Laboratory in Chicago, IL. The tutorials covered all aspects of software development with the Globus Toolkit including presentation, code walk-throughs, and demonstrations. The Argonne tutorials were recorded and are being indexed for streaming playback over the web.
- ***Tested a prototype of an Open Grid Services Architecture based Reliable File Transfer Service***
We have successfully modified the RFT service to meet the developing OGSA specifications as well as incorporating input from DOE science Grid and the Earth Sciences Grid. This will be an ongoing process as the OGSA specifications are still changing, but we are gaining valuable experience and providing feedback to the OGSA developers.
- ***Conducted a 54 hour transfer of 263 GB of data using the Reliable File Transfer Service***
Multiple tests were conducted between ANL and the NERSC HPSS system, the longest of which lasted 54 hours. The tests recovered from multiple failures including a shutdown of the HPSS system, NFS failures, a network outage, and us intentionally killing the service, the transfer client, and the GridFTP servers.
- ***Completed Beta Testing on GridFTP***
A number of bugs were found and fixed in the server. This has resulted in much improved stability in the server. Interestingly, a number of the bugs were not found in our testing because we did not have a network SLOW enough to find them.
- ***Designed and developed an alpha prototype of the Giggle Replica Location Service***
We designed a distributed service for maintaining mappings between logical file names and physical file locations. We delivered an alpha prototype of the Giggle Replica Location Service on February 28th to Work Package 2 of the European DataGrid project. During March, the WP2 group and the Globus group performed functional testing on the alpha prototype, resulting in bug fixes and two additional code releases.
-
- ***Design proposed for Replica Management Service***
We have designed a higher-level service that will replace the current replica management API of Globus 2.0 Beta. This service will combine data transfer operations controlled by the Reliable File Transfer Service with calls to the Replica Location Service to register and discover replicas.

Plans for next quarter

- ***Execute a feasibility study for integration of GridFTP into HPSS***
We will be working with Los Alamos National Laboratory and the HPSS development team to assess the feasibility of replacing the HPSS data mover protocol with GridFTP. The proposal is that this will make it into production HPSS, making all HPSS systems natively accessible by GridFTP clients.
- ***Release an alpha of the GridFTP V1.1 that includes usability enhancements***
We are currently working on a number of enhancements to increase usability of the server and make support easier. These include runtime version dumps to verify versions of all components including ones that are dynamically linked, options to control the frequency of restart and performance markers, and support for connection decisions to be based on filenames to support proxy applications.
- ***Continue development work on RFT***
Plans for RFT next quarter include further refinement of the interface in cooperation with DOE Science Grid, Earth Science Grid, and Particle Physics Data Grid representatives, continued work on integrating into the Open Grid Services Architecture Framework, and continued stability testing.
- ***Design for Server Development***
We are going to be doing a light weight, from scratch server development. This primarily to support the ability to start up a striped server from within a parallel application running on a scheduled resource where you can not pre-install servers. This will also allow us to bring the striped and non-striped code bases back together.
- ***Continued testing and development of alpha prototype of Replica Location Service***
We will continue to work with EDG WP2 on functional, integration and stress testing of the RLS prototype. We plan to extend the set of alpha testers to include users from PPDG, ESG and GriPhyN communities. We plan to add functionality to the alpha prototype, including support for master copies, integration with the Community Authorization Service, support for partitioning and redundancy of the replica location indices, and development of a SOAP/OGSA interface for the RLS.
- ***Design and alpha prototype development of Replica Management Service***
We will reach consensus on our design proposal for the Replica Management Service and implement and test an alpha prototype of the service.
- ***Design and alpha prototype development of Metadata Service***
We will implement a prototype of a simple, centralized metadata service. This may be based on the Spitfire database service from the European DataGrid project or may be based on our own server in front of a MySQL database.
- ***Globus Toolkit Final Release***
The Globus Toolkit V2.0 should be in final release by May 1st. This will be the first official release of GridFTP in final form.

Papers Published or in Progress

Giggle: A Framework for Constructing Scalable Replica Location Services. Ian Foster, Adriana Iamnitchi, Matei Ripeanu, Ann Chervenak, Ewa Deelman, Carl Kesselman, Wolfgang Hoschek, Peter Kunszt, Heinz Stockinger, Kurt Stockinger, Brian Tierney. Draft Document of the Data Replication Research Group of Global Grid Forum 4, February 2002.

Giggle: A Framework for Constructing Scalable Replica Location Services. Ann Chervenak, Ewa Deelman, Ian Foster, Adriana Iamnitchi, Carl Kesselman, Peter Kunsz, Matei Ripeanu, Heinz Stockinger, Kurt Stockinger. Submitted to HPDC11.

Presentations Given

1. 28-31 Jan 2002: Globus Developers Tutorial, Argonne National Laboratory, Chicago, Illinois, USA
2. 21-24 Jan 2002: Globus Developers Tutorial, Edinburgh Scotland
3. 10 Jan 2002: Presentation of Latest Replica Catalog Design at PPDG Reliable Replication Meeting, Jefferson National Accelerator Laboratory, Newport News, Virginia, USA
4. 18-20 Feb 2002: As Working Group Chair of GGF GridFTP working group, led working group discussion of current protocol draft, interactions between GridFTP and Storage Resource Managers, and desirability of organizing an expanded Transport Research Group within GGF.
5. 18-20 Feb 2002: As Working Group Chair of GGF Data Replication research group, led working group discussion of the Giggle Replica Location Service. Participated in discussion of the replication functionality of Giggle compared with Storage Resource Broker and JLAB schemes.
6. 20-21 Feb 2002: Attended PPDG meeting and made presentation on current status of Reliable File Transfer service.
7. 6-8 March 2002: European Data Grid Meeting in Paris. Participated in discussions with various Work Packages regarding GridFTP, Reliable File Transport and Globus in general. Presented the Globus Roadmap for their planning purposes.