



RELIABLE FILE TRANSFER

Contact: Madduri@mcs.anl.gov

The Globus Toolkit's Reliable File Transfer Service is a Web Services Resource Framework (WSRF)-compliant Web service that provides job scheduler-like functionality for data movement.

While globus-url-copy and GridFTP in general constitute a powerful set of tools, certain features may not always be optimal. First, the GridFTP protocol is not a Web service protocol: it does not employ SOAP, WSDL, and the like. Second, GridFTP requires that the client maintain an open socket connection to the server throughout the transfer. For long transfers this approach may not be convenient, such as if a job is run from the user's laptop. While globus-url-copy uses the robustness features of GridFTP to recover from remote failures (network outages, server failures, etc.), a failure of the client or the client's host means that recovery is not possible because the information needed for recovery is held in the client's memory. What is needed to address these issues is a service interface,

Web: http://dev.globus.org/wiki/Reliable_File_Transfer based on Web services protocols, that maintains the transfer state in reliable storage. We provide such a service, called the Reliable File Transfer (RFT) service.

With RFT, the user simply provides a list of source and destination URLs (including directories or file globs). The service writes the job description into a database and then moves the files on the user's behalf. Once the service has taken a job request, interactions with it are similar to any job scheduler. Service methods are provided for querying the transfer status, or the user may use standard WSRF tools (also provided in the Globus Toolkit) to subscribe for notifications of state change events.

We provide the service implementation, which is installed in a Web services container (like all Web services) and a simple client. Java classes are available for custom development; but because of lack of time and resources, work is still needed to make this easier.

New Features Coming in RFT in GT 4.2

- **Command line client in C:** A new command line client in C will enable users to get around the JVM startup time of the current Java command line client.
- **Transfer time prediction:** New resource properties will provide transfer time predictions based on performance markers from the GridFTP servers and based on historic data of the transfers.
- **Transfer priorities:** Priorities will be implemented in transfers.
- **Data scheduler functionality in RFT:** Users will be able to schedule and manage data transfers.
- **Multiple transfer protocols:** Currently, RFT supports third-party GridFTP transfers only. In GT 4.2, RFT will support multiple transfer protocols.
- **Changes in transfer requests:** Users will be able to change a transfer request after the request is submitted. Also provided will be the ability to add new transfers to an existing RFT resource.