

SSN COLLEGE OF ENGINEERING
Department of Computer Science and Engineering
CS6712 Grid and Cloud Computing Laboratory

Assignment -13 : Study on GridFTP

Assigned Date: 26.09.2017.

Due Date: 26. 09.2017 & 26.09.2017

Grid FTP Components

GridFTP is a high-performance, secure, reliable data transfer protocol optimized for high-bandwidth wide-area networks. The GridFTP protocol is based on FTP, the highly-popular Internet file transfer protocol.

One of the foundational issues in HPC computing is the ability to move large (multi Gigabyte, and even Terabyte), file-based data sets between sites. Simple file transfer mechanisms such as FTP and SCP are not sufficient either from a reliability or performance perspective.

GridFTP extends the standard FTP protocol to provide a high-performance, secure, reliable protocol for bulk data transfer.

GridFTP Protocol

GridFTP is a protocol defined by Global Grid Forum Recommendation GFD.020, RFC 959, RFC 2228, RFC 2389, and a draft before the IETF FTP working group. Key features include:

Performance - GridFTP protocol supports using parallel TCP streams and multi-node transfers to achieve high performance.

Checkpointing - GridFTP protocol requires that the server send restart markers (checkpoint) to the client.

Third-party transfers - The FTP protocol on which GridFTP is based separates control and data channels, enabling third-party transfers, that is, the transfer of data between two end hosts, mediated by a third host.

Security - Provides strong security on both control and data channels. Control channel is encrypted by default. Data channel is authenticated by default with optional integrity protection and encryption.

Globus Implementation of GridFTP

The GridFTP protocol provides for the secure, robust, fast and efficient transfer of (especially bulk) data. The Globus Toolkit provides the most commonly used implementation of that protocol, though others do exist (primarily tied to proprietary internal systems).

The Globus Toolkit provides:

- i. a server implementation called globus-gridftp-server,
- ii. a scriptable command line client called globus-url-copy, and
- iii. a set of development libraries for custom clients.

While the Globus Toolkit does not provide a client with Graphical User Interface (GUI), Globus Online provides a web GUI for GridFTP data movement.

Globus GridFTP framework implements all the key features of GridFTP protocol mentioned above. It supports both Grid Security Infrastructure (GSI) and SSH for securing the data transfer. Unlike sftp, SSH based GridFTP supports multiple security options on the data channel - authentication only, authentication and integrity protection, fully encrypted. Globus implementation of GridFTP is modular and extensible. XIO based Globus GridFTP framework makes it easy to plugin alternate transport protocols. The Data Storage Interface (DSI) allows for easier integration with various storage systems.

Globus GridFTP has been deployed at thousands of sites with more than 10 million data transfers per day.

GridFTP Clients

Globus Online is the recommended interface to move data to and from GridFTP servers. Globus Online provides a web GUI, command line interface and a REST API for GridFTP data movement. It provides automatic fault recovery and automatic tuning of optimization parameters to achieve high performance.

The Globus Toolkit provides a GridFTP client called globus-url-copy, a command line interface, suitable for scripting. For example, the following command:

```
globus-url-copy gsiftp://remote.host.edu/path/to/file file:///path/on/local/host
```

would transfer a file from a remote host to the locally accessible path specified in the second URL. Finally, if you wish to add access to files stored behind GridFTP servers, or you need custom client functionality, you can use our very powerful client library to develop custom client functionality.

References.

<http://www.gridftp.org>

<http://www.globus.org/toolkit>

gridftp-user@globus.org

<http://toolkit.globus.org/toolkit/docs/latest-stable/gridftp/key/>