
GT 5.0.0 Release Notes: GridFTP

Table of Contents

1. Component Overview	1
2. Feature Summary	1
3. Summary of Changes in GridFTP	3
4. Bug Fixes	3
5. Known Problems	3
6. Technology dependencies	3
7. Tested platforms	4
8. Backward compatibility summary	4
9. Associated Standards	4
10. For More Information	5
Glossary	5

<titleabbrev>Release Notes</titleabbrev>

1. Component Overview

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. We have selected a set of protocol features and extensions defined already in IETF RFCs and added a few additional features to meet requirements from current data grid projects.

2. Feature Summary

Features new in GT 5.0.0:

- Improved failure restart capability in globus-url-copy:

A new option to store untransferred urls for later restarting is available. In case of any failures, this option allows users to restart transfers from a checkpoint rather than restarting from scratch. This option can be used for directory transfers, single file transfer or a list of files transfer. See ??? for more details.

- Stall detection:

It is possible that the transfer hangs due to filesystem errors, network errors or GridFTP server errors. A new option is available in globus-url-copy to specify how long before canceling/restarting a transfer with no data movement. See the section called “Reliability Options” for more details.

- Client-side host aliasing:

This allows concurrent transfers to be extended to multiple different hosts rather than multiple connections to the same host, without relying on DNS.

Features that continue to be supported from previous versions

- GridFTP over UDT
- SSH security for GridFTP control channel

- Running the GridFTP server with GFork GridFTP
- Multicasting / Network overlays
- Netlogger's bottleneck detection for GridFTP transfers
- GSI security: This is the PKI based, de facto standard security system used in Grid applications. Kerberos is also possible but is not supported and can be difficult to use due to divergence in the capabilities of GSI and Kerberos.
- Third-party transfers: Very common in Grid applications, this is where a *client* mediates a transfer between two servers (both likely at remote sites) rather than between the server and itself (called a *client/server transfer*).
- Cluster-to-cluster data movement or Striping: GridFTP can do coordinated data transfer by using multiple computer nodes at the source and destination.
- Partial file access: Regions of a file may be accessed by specifying an offset into the file and the length of the block desired.
- Reliability/restart: The receiving server periodically (the default is 5 seconds, but this can be changed) sends “restart markers” to the client. This marker is a messages specifying what bytes have been successfully written to the disk. If the transfer fails, the client may restart the transfer and provide these markers (or an aggregated equivalent marker), and the transfer will pick up where it left off. This can include “holes” in the file.
- Large file support: All file sizes, lengths, and offsets are 64 bits in length.
- Data channel reuse: Data channel can be held open and reused if the next transfer has the same source, destination, and credentials. This saves the time of connection establishment, authentication, and delegation. This can be a huge performance difference when moving lots of small files.
- Integrated instrumentation (Performance Markers).
- Logging/audit trail (Extensive Logging in the server).
- Parallel transfers (Multiple TCP streams between a pair of hosts).
- TCP Buffer size control (Protocol supports Manual and Automatic; Only Manual Implemented).
- Server-side computation (Extended Retrieve (ERET) / Extended Store (ESTO) commands).
- Based on Standards: RFC 959, RFC 2228, RFC 2389, IETF Draft MLST-16 , GGF GFD.020.

Other Supported Features

- On the client side we provide a scriptable tool called `globus-url-copy`. This tool can take advantage of all the GridFTP protocol features and can also do protocol translation between FTP, HTTP, HTTPS, and POSIX file IO on the client machine.
- We also provide a set of development libraries and APIs for developers wishing to add GridFTP functionality to their application.

Deprecated Features

- None

3. Summary of Changes in GridFTP

- The default flavor of the GridFTP server has been changed to non-threaded.

4. Bug Fixes

- [Bug 6783](https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6783)¹ - Gridftp intermittently gives users access to groups that root is in but the user is not
- [Bug 6601](https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6601)² - globus-gridftp-server -ssh hangs when using OpenSSH 5.1p1
- [Bug 6851](https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6851)³ - Memory violation in globus_ftp_control_auth_info_compare()
- [Bug 6867](https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6867)⁴ - GridFTP FTP_INFO logs sometimes have tiny or negative durations
- [Bug 6848](https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6848)⁵ - Change in behaviour with -f flag between versions 3.23 and 4.14

5. Known Problems

The following problems and limitations are known to exist for GridFTP at the time of the 5.0.0 release:

5.1. Known Bugs

The following problems are known to exist for GridFTP at the time of the 5.0.0 release:

- GridFTP Server
 - There are some small memory leaks, though they should not grow much.
 - Some error responses are unclear.

5.1.1. Limitations

In order to use custom XIO drivers with GridFTP and globus-url-copy, the flavor of the driver must match the flavor of globus-gridftp-server and globus-url-copy (for 2-party transfers).

6. Technology dependencies

GridFTP depends on the following GT components:

- Non-WS (General) Authentication & Authorization
- C Common Libraries
- XIO

GridFTP depends on the following 3rd party software:

¹ https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6783

² https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6601

³ https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6851

⁴ https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6867

⁵ https://bugzilla.mcs.anl.gov/globus/show_bug.cgi?id=6848

- OpenSSL (version is included in release)

7. Tested platforms

Tested platforms for GridFTP

- i386 Linux
- ia64 Linux (TeraGrid)
- AIX 5.2
- Solaris 9
- PA-RISC HP/UX 11.11
- ia64 HP/UX 11.22
- Tru64 Unix
- Mac OS X

While the above list includes platforms on which we have tested GridFTP, it does not imply support for a specific platform. However, we are interested in hearing reports of success or bug reports on any platform.

8. Backward compatibility summary

Protocol changes since GT 4.2.x

- None

API changes since GT 4.2.x

- None

Exception changes since GT 4.2.x

- Not Applicable (GridFTP is not Java-based)

Schema changes since GT 4.2.x

- Not Applicable (GridFTP is not SOAP-based)

9. Associated Standards

Associated standards for GridFTP:

- [RFC 959 Base FTP protocol](http://rfc.net/rfc959.html)⁶
- [RFC 2228 gssapi security extensions](http://rfc.net/rfc2228.html)⁷ for FTP RFC 2389 FEAT, OPTS, etc.

⁶ <http://rfc.net/rfc959.html>

⁷ <http://rfc.net/rfc2228.html>

- [extensions to FTP \(IETF FTP Working group draft\)](#)⁸ for structured directory listings, SIZE, MDTM commands.
- [GFD.020 GridFTP extensions](#)⁹

10. For More Information

See [GridFTP](#) for more information about this component.

Glossary

C

client A process that sends commands and receives responses. Note that in GridFTP, the client may or may not take part in the actual movement of data.

⁸ <http://www.ietf.org/internet-drafts/draft-ietf-ftpext-mlst-16.txt>

⁹ <http://www.ggf.org/documents/GWD-R/GFD-R.020.pdf>