NAME

gtransfer − The GridFTP data transfer script

SYNOPSIS

{gtransfer|gt} [−−source|−s *sourceUrl***] [−−destination|−d** *destinationUrl***] [−−transfer−list|−f** *transferList***] [−−auto−optimize|−o** *transferMode***] [−−recursive|−r][−−checksum−data−channel|−c] [−−encrypt−data−channel|−e][−−guc−max−retries** *gucMaxRetries***] [−−gt−max−retries** *gtMaxRetries***] [−−gt−progress−indicator** *indicatorCharacter***] [−−verbose|−v] [−−metric|−m** *dataPathMetric***] [−−logfile|−l** *logfile***] [−−auto−clean|−a] [−−configfile** *configurationFile***] [−−** *gucParameters***]**

DESCRIPTION

gtransfer or short **gt** is a wrapper script for the **tgftp(1)** tool and provides an advanced command line interface for performing GridFTP data transfers.

gt has the following features:

Multi−step data transfers
gtransfer can transfer files along predefined paths by using transit sites and can therefore bridge different network domains. See **dpath(5)** for more details.

Data transfer using multipathing
gtransfer can distribute a data transfer over multiple paths. This way users can benefit from the combined bandwidth of multiple paths. See option **−m** for usage details.

Optimized data transfer performance
gtransfer supports usage of pre−optimized data transfer parameters for specific connections. See **dparam(5)** for more details. In addition **gt** can also automatically optimize a data transfer depending on the size of the files.

Data transfer interruption and continuation
gtransfer supports interruption and continuation of transfers. You can interrupt a transfer by hitting `CTRL+C`. To continue an interrupted transfer simply issue the very same command, **gt** will then continue the transfer where it was interrupted. The same procedure works for a failed transfer.

Data transfer reliability
gtransfer supports automatic retries of failed transfer steps. The number of retries is configurable.

Bash completion
gtransfer makes use of bash completion to ease usage. This supports completion of options and URLs. URL completion also expands (remote) paths. Just hit the `TAB` key to see what's possible.

Host aliases
gtransfer can use host aliases as alternatives to host addresses. E.g. a user can use "myGridFTP:" and "gsiftp://host1.domain.tld:2811" synonymically.

Persistent identifiers (PIDs)
gtransfer can use persistent identifiers (PIDs) as used by EUDAT and provided by EPIC as source of a data transfer.

OPTIONS

The options are as follows:

[−s, −−source *sourceUrl***]**
Set the source URL for the transfer.

Possible URL examples:

- {[gsi]ftp|http[s]}://FQDN[:PORT]/path/to/file
- [file://]/path/to/file

"FQDN" is the fully qualified domain name.

[−d, −−destination *destinationUrl***]**
Set the destination URL for the transfer.

Possible URL examples:

- [gsi]ftp://FQDN[:PORT]/path/to/file
- [file://]/path/to/file

"FQDN" is the fully qualified domain name.

[−f, −−transfer−list *transferList***]**
As alternative to providing source and destination URLs on the command line, one can also provide a list of source and destination URLs in a transfer list; **gtransfer** will then perform a *list transfer* instead of an *URL transfer* when using source and destination URLs.

The format of each line of the transfer list file is as follows (including the double quotes!):

"<PROTOCOL>://<FQDN1>:<PORT>/path/to/file" "<PROTOCOL>://<FQDN2>:<PORT>/path/to/file[s/]"

Throughout all lines the source URL host part (e.g. "<PROTOCOL>://<FQDN1>:<PORT>") has to be identical. This is also required for the destination URL host part.

[−o, −−auto−optimize *transferMode***]**
This option activates an automatic optimization of transfers depending on the size of files to be transferred. If less than 100 files are going to be transferred, gtransfer will fall back to list transfer. The *transferMode* controls how files of different size classes are transferred. Currently "seq[uential]" (different size classes are transferred sequentially) is possible. To define different file size classes use the file *[...]/chunkConfig*. See **FILES** section below for more details.

[−r, −−recursive]
Transfer files recursively.

NOTICE: globus−url−copy(1) (even with option **−cd**) and therefore also **gt** will not create directories on the destination side that are empty on the source side.

[−c, −−checksum−data−channel]
Enable checksumming on the data channel. Cannot be used in conjunction with **−e**!

[−e, −−encrypt−data−channel]
Enable encryption on the data channel. Cannot be used in conjunction with **−c**!

[−−guc−max−retries *gucMaxRetries***]**
This option sets the maximum number of retries **globus−url−copy(1)** will do for a transfer of a single file. By default this is set to 1, which means that **globus−url−copy(1)** will tolerate at max. one transfer error per file and retry the transfer once. Alternatively this option can also be set with the environment variable **GUC_MAX_RETRIES**.

[−−gt−max−retries *gtMaxRetries***]**
This option sets the maximum number of retries **gt** will do for a single transfer step. By default this is set to 3, which means that **gt** will try to finish a single transfer step three times or fail. Alternatively this option can also be set with the environment variable **GT_MAX_RETRIES**.

[−v, −−verbose]
Be verbose.

[−m, −−metric *dataPathMetric***]**
Set the metric to select the corresponding path of a data path. To enable multipathing, use either the keyword "all" to transfer data using all available paths or use a comma separated list with the metric values of the paths that should be used (e.g. "0,1,2"). You can also use metric values multiple times (e.g. "0,0").

[−l, −−logfile *logfile***]**
Set the name for the logfile, **tgftp(1)** will generate for each transfer. If specified with ".log" as extension, **gt** will insert a "__step_#" string to the name of the logfile ("#" is the number of the transfer step performed).

If omitted **gt** will automatically generate a name for the logfile(s).

[−a, −−auto−clean]
Remove logfiles automatically after the transfer completed.

[−−configfile *configurationFile***]**
Set the name of the configuration file for **gt**. If not set, this defaults to:

1. "/etc/gtransfer/gtransfer.conf" or
2. "/etc/gtransfer.conf" or
3. "/etc/opt/gtransfer/gtransfer.conf" or
4. "$HOME/.gtransfer/gtransfer.conf" or
5. "$(dirname $BASH_SOURCE)/../etc/gtransfer/gtransfer.conf" in this order.

[−− *gucParameters***]**
Set the **globus−url−copy(1)** parameters that should be used for all transfer steps. Notice the space between "−−" and the actual parameters. This overwrites any available dparams and is not recommended for regular usage. There exists one exception for the **−len|−partial−length X** option. If this is provided, it will only be added to the transfer parameters from a dparam for a connection or − if no dparam is available − to the builtin default transfer parameters.

NOTICE: If specified, this option must be the last one in a **gt** command line.

General options:

[−−help]
Prints out a help message.

[−V, −−version]
Prints out version information.

ENVIRONMENT VARIABLES

GUC_MAX_RETRIES
See option **−−guc−max−retries** for details.

GT_MAX_RETRIES
See option **−−gt−max−retries** for details.

GT_KEEP_TMP_DIR
If set to 1, **gt** will keep its used temporary directory below ~/.gtransfer/tmp for inspection when exiting.

GT_NO_RELIABILITY
If set to 1, **gt** will not make use of the reliabilty functionality of **globus−url−copy(1)**. This means that transfers always start from the beginning. I.e. transfers cannot be interrupted and later continued from where they were interrupted and transfers that failed temporarily will also start from the beginning, when retried.

FILES

[...]/gtransfer.conf
The **gt** configuration file.

[...]/chunkConfig
The chunk configuration file. In this file you can define the different file size classes for the auto−optimization. Practically the file is a table with three columns: **MIN_SIZE_IN_MB**, **MAX_SIZE_IN_MB** and **GUC_PARAMETERS** separated by a semicolon.

Each line defines a size class. The value for **MIN_SIZE_IN_MB** is not included in the class. The value for **MAX_SIZE_IN_MB** is included in the class. Use the keyword "min" in the column **MIN_SIZE_IN_MB** to default to the size of the smallest file available in a transfer list. Files of this size will be included in this class then. Use the keyword "max" in the column **MAX_SIZE_IN_MB** to default to the size of the biggest file available in a transfer list. The third column **GUC_PARAMETERS** defines the transfer parameters to use for the specific file size class.

Example:

```
#MIN_SIZE_IN_MB;MAX_SIZE_IN_MB;GUC_PARAMETERS
min;50;-cc 16 -tcp-bs 4M -stripe -sbs 4M -cd
50;250;-cc 8 -tcp-bs 8M -stripe -sbs 4M -cd
250;max;-cc 6 -p 4 -tcp-bs 8M -stripe -sbs 8M -g2 -cd
```

[...]/dpaths/
This directory contains the system dpaths usable by **gt** and is configurable.

[...]/dparams/
This directory contains the system dparams usable by **gt** and is configurable.

$HOME/.gtransfer/dpaths/
This directory contains the user dpaths usable by **gt**. Can be created with **dpath(1)**. If existing, dpaths in this directory have precedence.

$HOME/.gtransfer/dparams/
This directory contains the user dparams usable by **gt**. Can be created with **dparam(1)**. If existing, dparams in this directory have precedence.

SEE ALSO

dparam(1), **dparam(5)**, **dpath(1)**, **dpath(5)**, **globus-url-copy(1)**, **tgftp(1)**, **uberftp(1C)**

AUTHORS

Frank Scheiner.