Globus Toolkit 5.0.3 Commandline Tools
Globus Toolkit 5.0.3 Commandline Tools

Abstract

You can also download a PDF version here¹.

¹ gtCommands.pdf
## Table of Contents

I. GSI Commands .......................................................................................................................... 6  
  globus-update-certificate-dir .................................................................................................. 7  
  grid-cert-diagnostics ........................................................................................................... 8  
  grid-cert-info ..................................................................................................................... 10  
  grid-cert-request ................................................................................................................ 12  
  grid-default-ca .................................................................................................................. 16  
  grid-change-pass-phrase .................................................................................................... 18  
  grid-proxy-init .................................................................................................................... 19  
  grid-proxy-destroy ............................................................................................................... 22  
  grid-proxy-info ................................................................................................................... 23  
  grid-mapfile-add-entry ....................................................................................................... 25  
  grid-mapfile-check-consistency .......................................................................................... 27  
  grid-mapfile-delete-entry ................................................................................................... 29  

II. GridFTP Commands ................................................................................................................. 31  
    globus-url-copy .................................................................................................................. 32  
    globus-url-sync ................................................................................................................ 45  
    globus-gridftp-server ........................................................................................................ 48  

III. Replica Location Service (RLS) Commands ........................................................................... 61  
     globus-rls-admin ............................................................................................................. 62  
     globus-rls-cli .................................................................................................................. 64  
     globus-rls-server ............................................................................................................. 68  

IV. GRAM5 Commands .................................................................................................................. 73  
    globusrun ......................................................................................................................... 74  
    globus-job-cancel ............................................................................................................. 78  
    globus-job-clean ............................................................................................................. 79  
    globus-job-get-output ..................................................................................................... 80  
    globus-job-run ................................................................................................................ 82  
    globus-job-status .......................................................................................................... 85  
    globus-job-submit .......................................................................................................... 87  
    globus-personal-gatekeeper ......................................................................................... 90  
    globus-gram-audit ......................................................................................................... 92  
    globus-job-manager ....................................................................................................... 93  
    globus-job-manager-event-generator ............................................................................... 98  

V. GSI-OpenSSH Commands ......................................................................................................... 100  
    gssh ............................................................................................................................... 100  
    gscp ............................................................................................................................... 101  
    gsiftp .............................................................................................................................. 102  

Glossary ...................................................................................................................................... 103
List of Figures

1. Effect of Parallel Streams in GridFTP ................................................................. 43
List of Tables

1. URL formats ................................................................................................................................. 34
2. URL formats ................................................................................................................................. 46
3. Options for globus-rls-admin ........................................................................................................... 63
4. Options for globus-rls-cli ............................................................................................................... 64
5. Commands for globus-rls-cli ......................................................................................................... 66
6. Options for globus-rls-server ........................................................................................................... 71
GSI Commands
Name
globus-update-certificate-dir -- Update symlinks in the trusted CA directory
globus-update-certificate-dir [-help] [-d DIRECTORY]

Description
The **globus-update-certificate-dir** program creates symlinks between files (CA certificates, certificate revocation lists, signing policy, and certificate request configuration files) using the certificate hash the installed version of OpenSSL uses. OpenSSL 1.0.0 uses a different name hashing algorithm than previous versions, so CA distributions created with older versions of OpenSSL might not be able to locate trusted CAs and related files. Running **globus-update-certificate-dir** against a trusted CA directory will add symlinks to the files to the hash if needed.

_NOTE_
To run globus-update-certificate-dir on Linux, modify that script so that the first line is
```bash
#!/usr/bin/env perl
```
instead of
```bash
#!/usr/bin/env perl -w
```

The full set of command-line options to **globus-update-certificate-dir** consists of:
- `-help` Display a help message to standard output and exit
- `-d DIRECTORY` Create links in the trusted CA directory `DIRECTORY` instead of using the default search path.

Environment
If the following variables affect the execution of **globus-update-certificate-dir**

- `X509_CERT_DIR` Default trusted certificate directory.
- `HOME` Path to the current user's home directory.
- `GLOBUS_LOCATION` Path to the Globus installation.
Name
grid-cert-diagnostics -- Print diagnostic information about certificates and keys
grid-cert-diagnostics [-h] [-help] [-p]

Description
The grid-cert-diagnostics program displays information about the current user's security environment, including information about security-related environment variables, security directory search path, personal key and certificates, and trusted certificates. It is intended to provide information to help diagnose problems using GSIC.

By default, grid-cert-diagnostics prints out information regarding the environment and trusted certificate directory. If the -p command-line option is used, then additional information about the current user's default certificate and key will be printed.

The full set of command-line options to grid-cert-diagnostics consists of:
- h,   Display a help message and exit.
- help
- p    Display information about the personal certificate and key that is the current user's default credential.

Examples
In this example, we see the default mode of checking the default security environment for the system, without processing the user's key and certificate. Note the user receives a warning about a cog.properties and about an expired CA certificate.

% grid-cert-diagnostics
Checking Environment Variables
=================================
Checking if X509_CERT_DIR is set... no
Checking if X509_USER_CERT is set... no
Checking if X509_USER_KEY is set... no
Checking if X509_USER_PROXY is set... no

Checking Security Directories
=================================
Determining trusted cert path... /etc/grid-security/certificates
Checking for cog.properties... found
   WARNING: If the cog.properties file contains security properties, Java apps will ignore the security paths described in the GSI documentation

Checking trusted certificates...
=================================
Getting trusted certificate list...
Checking CA file /etc/grid-security/certificates/1c4f4c48.0... ok
Verifying certificate chain for "/etc/grid-security/certificates/1c3f2ca8.0"... ok
Checking CA file /etc/grid-security/certificates/9d8788eb.0... ok
Verifying certificate chain for "/etc/grid-security/certificates/9d8753eb.0"... failed
globus_certificate: Error verifying credential: Failed to verify credential
globus_gsi_callback_module: Could not verify credential
    The certificate has expired:
    Credential with subject: /DC=org/DC=example/OU=grid/CN=CA has expired.

In this example, we show a user with a mismatched private key and certificate:

% grid-cert-diagnostics -p

Checking Environment Variables
================================
Checking if X509_CERT_DIR is set... no
Checking if X509_USER_CERT is set... no
Checking if X509_USER_KEY is set... no
Checking if X509_USER_PROXY is set... no

Checking Security Directories
============================
Determining trusted cert path... /etc/grid-security/certificates
Checking for cog.properties... not found

Checking Default Credentials
============================
Determining certificate and key file names... ok
Certificate Path: "/home/juser/.globus/usercert.pem"
Key Path: "/home/juser/.globus/userkey.pem"
Reading certificate... ok
Reading private key... ok
Checking Certificate Subject... "/O=Grid/OU=Example/OU=User/CN=Joe User"
Checking cert... ok
Checking key... ok
Checking that certificate contains an RSA key... ok
Checking that private key is an RSA key... ok
Checking that public and private keys have the same modulus... failed
Private key modulus: D294849E37F048C3B5ACEEF2CCDF97D88B679C361E29D5CB5
219C3E948F3E530CFC609489759E1D751F0ACF0F0515A614276A0F4C11A57D92D7165B8
FA64E3140155DE44B8D45C182F4657DA13EDA2B8422F5B9D169DFF3822EFD81EB2E6403
CE3CB4CCF96B65284D92592BB1673A18354DA241B9AFD7F494E54F63A93E15DCAE2
Checking that public and private keys have the same modulus... failed
Public key modulus : C002C7B329B133BAF87BAF214EACE3DC3490165ACEB791790
600708C544175D9193C9BAC5AE03B7CB49BBB6A6D29B7E635FAC751E9A6D1CEA98022
6F1B6300292D6623A319E4682E7BBF0968DCE962CF218A9D95FAAD6A0BA5C42AA9A9EF
7FD3D2B37C6E2B2F0E311310A55F8B9EAFDF5B995C7D9EEAD8D5D81F3531E0AE5
Certificate and and private key don't match
Name

grid-cert-info -- Display information about a certificate

grid-cert-info [-file CERTIFICATE-FILE] [-rfc2253] [-all]
  [-subject] [-s]
  [-issuer] [-i]
  [-issuerhash] [-ih]
  [-startdate] [-sd]
  [-enddate] [-ed]

Description

The grid-cert-info program displays information contained within a certificate file. By default it shows a text representation of the entire certificate. Specific facts about the certificate can be shown instead by using command-line options. If any of those options are used, then the default display is suppressed. This can be added to the output by using the -all command-line option.

If multiple display options are included on the command-line, the facts related to those will be displayed on separate lines in the order that they occur. If an option is specified multiple time, that fact will be displayed multiple times.

The full set of command-line options to grid-cert-info are:

- help, -usage
  Display the command-line options to grid-cert-info and exit.

- version, -versions
  Display the version number of the grid-cert-info command. The second form includes more details.

- file CERTIFICATE-FILE
  Display information about the first certificate contained in the file named by CERTIFICATE-FILE instead of the default user certificate.

- rfc2253
  Display X.509 distinguished names using the string representation defined in RFC 2253 instead of the default OpenSSL oneline format.

- all
  Display the text representation of the entire certificate in addition to any other facts requested by command-line options. This is the default if no fact-specific command-line options are used.

- subject, -s
  Display the subject name of the X.509 certificate.

- issuer, -i
  Display the issuer name of the X.509 certificate.

- issuerhash, -ih
  Display the default hash of the issuer name of the X.509 certificate. This can be used to locate which CA certificate in the trusted certificate directory issued the certificate being inspected.

- startdate, -sd
  Display a string representation of the date and time when the certificate is valid from. This is displayed in the format used by the OpenSSL x509 command.

- enddate, -ed
  Display a string representation of the date and time when the certificate is valid until. This is displayed in the format used by the OpenSSL x509 command.
Examples

Display the validity times for the default certificate

```
grid-cert-info -sd -ed
Aug 31 12:33:47 2009 GMT
Aug 31 12:33:47 2010 GMT
```

Display the same information about a different certificate specified on the command-line

```
grid-cert-info -sd -ed -f /etc/grid-security/hostcert.pem
Jan 21 12:24:48 2003 GMT
Jul 15 11:30:57 2020 GMT
```

Display the subject of a certificate in both the default and the RFC 2253 forms.

```
grid-cert-info -subject
/DC=org/DC=example/DC=grid/CN=Joe User
grid-cert-info -subject -rfc2253
CN=Joe User,DC=grid,DC=example,DC=org
```

Environment Variables

The following environment variables affect the execution of `grid-cert-info`:

- `X509_USER_CERT` Path to the default certificate file to inspect.
Name

grid-cert-request -- Generate a X.509 certificate request and corresponding private key

grid-cert-request [-help] [-h] [-?] [-usage]
[ -version] [-versions]
grid-cert-request [-cn NAME] [-commonname NAME]
[ -dir DIRECTORY] [-prefix PREFIX]
[ -nopw | -nodes | -nopassphrase ]
[ -nopw | -nodes | -nopassphrase ]
[ -ca [HASH]] [-verbose] [ -interactive | -int ] [-force]
grid-cert-request -host FQDN [-service SERVICE] [-dns FQDN...] [-ip IP-ADDRESS...]
[ -dir DIRECTORY] [-prefix PREFIX]
[ -ca [HASH]] [-verbose] [ -interactive | -int ] [-force]

Description

The grid-cert-request program generates an X.509 Certificate Request and corresponding private key for the specified name, host, or service. It is intended to be used with a CA implemented using the globus_simple_ca package.

The default behavior of grid-cert-request is to generate a certificate request and private key for the user running the command. The subject name is derived from the gecos information in the local system's password database, unless the -commonname, -cn, or -host command-line options are used.

By default, grid-cert-request writes user certificate requests and keys to the $HOME/.globus directory, and host and service certificate requests and keys to /etc/grid-security. This can be overridden by using the -dir command-line option.

The full set of command-line options to grid-cert-request are:

- help, -h, -?, -usage Display the command-line options to grid-cert-request and exit.
- version, -versions Display the version number of the grid-cert-request command. The second form includes more details.
- cn NAME, -commonname NAME Create a certificate request with the common name component of the subject set to NAME. This is used to create user identity certificates.
- dir DIRECTORY Write the certificate request and key to files in the directory specified by DIRECTORY.
- prefix PREFIX Use the string PREFIX as the base name of the certificate, certificate_request, and key files instead of the default. For a user certificate request, this would mean creating files $HOME/.globus/PREFIXcert_request.pem, $HOME/.globus/PREFIXcert.pem, and $HOME/.globus/PREFIXkey.pem.
- ca CA-HASH Use the certificate request configuration for the CA with the name hash CA-HASH instead of the default CA chosen by running grid-default-ca.
- verbose Keep the output from the OpenSSL certificate request command visible after it completes, instead of clearing the screen.
- interactive, -int Prompt for each component of the subject name of the request, instead of generating the common name from other command-line options. Note that CAs may not sign certificates for subject names that don't match their signing policies.
grid-cert-request

-force Overwrite any existing certificate request and private key with a new one.

-nopw, -nodes, -no-passphrase Create an unencrypted private key for the certificate instead of prompting for a passphrase. This is the default behavior for host or service certificates, but not recommended for user certificates.

-host FQDN Create a certificate request for use on a particular host. This option also causes the private key associated with the certificate request to be unencrypted. The FQDN argument to this option should be the fully qualified domain name of the host that will use this certificate. The subject name of the certificate will be derived from the FQDN and the service option if specified by the -service command-line option. If the host for the certificate has multiple names, then use either the -dns or -ip command-line options to add alternate names or addresses to the certificates.

-service SERVICE Create a certificate request for a particular service on a host. The subject name of the certificate will be derived from the FQDN passed as the argument to the -host command-line option and the SERVICE string.

-dns FQDN, ... Create a certificate request containing a subjectAltName extension containing one or more host names. This is used when a certificate may be used by multiple virtual servers or if a host has different names when contacted within or outside a private network. Multiple DNS names can be included in the extension by separating them with a comma.

-ip IP-ADDRESS, ... Create a certificate request containing a subjectAltName extension containing the IP addresses named by the IP-ADDRESS strings. This is used when a certificate may be used by services listening on multiple networks. Multiple IP addresses can be included in the extension by separating them with a comma.

Examples

Create a user certificate request:

% grid-cert-request
A certificate request and private key is being created.
You will be asked to enter a PEM pass phrase.
This pass phrase is akin to your account password,
and is used to protect your key file.
If you forget your pass phrase, you will need to
obtain a new certificate.
A private key and a certificate request has been generated with the subject:

/O=org/OU=example/OU=grid/CN=Joe User

If the CN=Joe User is not appropriate, rerun this
script with the -force -cn "Common Name" options.

Your private key is stored in /home/juser/.globus/userkey.pem
Your request is stored in /home/juser/.globus/usercert_request.pem

Please e-mail the request to the Example CA ca@grid.example.org
You may use a command similar to the following:

    cat /home/juser/.globus/usercert_request.pem | mail ca@grid.example.org
Only use the above if this machine can send AND receive e-mail. if not, please mail using some other method.

Your certificate will be mailed to you within two working days.
If you receive no response, contact Example CA at ca@grid.example.org

Create a host certificate for a host with two names.

```
% grid-cert-request -host grid.example.org -dns grid.example.org,grid-internal.example.org
```

A private host key and a certificate request has been generated
with the subject:

```
/O=org/OU=example/OU=grid/CN=host/grid.example.org
```

The private key is stored in /etc/grid-security/hostkey.pem
The request is stored in /etc/grid-security/hostcert_request.pem

Please e-mail the request to the Example CA ca@grid.example.org
You may use a command similar to the following:

```
cat /etc/grid-security/hostcert_request.pem | mail ca@grid.example.org
```

Only use the above if this machine can send AND receive e-mail. if not, please mail using some other method.

Your certificate will be mailed to you within two working days.
If you receive no response, contact Example CA at ca@grid.example.org

**Environment Variables**

The following environment variables affect the execution of `grid-cert-request`:

- `X509_CERT_DIR` Path to the directory containing SSL configuration files for generating certificate requests.
- `GRID_SECURITY_DIR` Path to the directory containing SSL configuration files for generating certificate requests. This value is used if `X509_CERT_DIR` is not set.
- `GLOBUS_LOCATION` Path to the directory containing the Globus Toolkit. This is searched if neither the `X509_CERT_DIR` nor the `GRID_SECURITY_DIR` environment variables are set.

**Files**

- `$HOME/.globus/usercert_request.pem` Default path to write a user certificate request.
- `$HOME/.globus/usercert.pem` Default path to write a user certificate.
- `$HOME/.globus/userkey.pem` Default path to write a user private key.
<table>
<thead>
<tr>
<th>Path</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>/etc/grid-security/host-cert_request.pem</code></td>
<td>Default path to write a host certificate request.</td>
</tr>
<tr>
<td><code>/etc/grid-security/host-cert.pem</code></td>
<td>Default path to write a host certificate.</td>
</tr>
<tr>
<td><code>/etc/grid-security/hostkey.pem</code></td>
<td>Default path to write a host private key.</td>
</tr>
<tr>
<td><code>TRUSTED-CERT-DIR/globus-user-ssl.conf</code></td>
<td>SSL configuration file for requesting a user certificate. The first form is the default location, the second form is used when the <code>–ca</code> command-line option is specified.</td>
</tr>
<tr>
<td><code>TRUSTED-CERT-DIR/globus-host-ssl.conf</code></td>
<td>SSL configuration file for requesting a host or service certificate. The first form is the default location, the second form is used when the <code>–ca</code> command-line option is specified.</td>
</tr>
</tbody>
</table>
Name

grid-default-ca -- Select default CA for certificate requests

grid-default-ca -list [-dir CA-DIRECTORY]
grid-default-ca [-ca CA-HASH] [-dir CA-DIRECTORY]

Description

The grid-default-ca program sets the default certificate authority to use when the grid-cert-request script is run. The CA's certificate, configuration, and signing policy must be installed in the trusted certificate directory to be able to request certificates from that CA. Note that some CAs have different policies and use other tools to handle certificate requests. Please consult your CA's support staff if you unsure. The grid-default-ca is designed to work with CAs implemented using the globus_simple_ca package.

By default, the grid-default-ca program displays a list of installed CA certificates and the prompts the user for which one to set as the default. If invoked with the -list command-line option, grid-default-ca will print the list and not prompt nor set the default CA. If invoked with the -ca option, it will not list or prompt, but set the default CA to the one with the hash that matches the CA-HASH argument to that option. If grid-default-ca is used to set the default CA, the caller of this program must have write permissions to the trusted certificate directory.

The grid-default-ca program sets the CA in the one of the grid security directories. It looks in the directory named by the GRID_SECURITY_DIR environment, the X509_CERT_DIR, /etc/grid-security, and $GLOBUS_LOCATION/share/certificates.

The full set of command-line options to grid-default-ca are:

- -help, -h, -usage, -u
  Display the command-line options to grid-default-ca and exit.

- -version, -versions
  Display the version number of the grid-default-ca command. The second form includes more details.

- -dir CA-DIRECTORY
  Use the trusted certificate directory named by CA-DIRECTORY instead of the default.

- -list
  Instead of changing the default CA, print out a list of all available CA certificates in the trusted certificate directory

- -ca CA-HASH
  Set the default CA without displaying the list of choices or prompting. The CA file named by CA-HASH must exist.

Examples

List the contents of the trusted certificate directory that contain the string Example:

% grid-default-ca | grep Example
15) cd1186ff - /DC=org/DC=Example/DC=Grid/CN=Example CA

Choose that CA as the default:

% grid-default-ca -ca cd1186ff

setting the default CA to: /DC=org/DC=Example/DC=Grid/CN=Example CA
linking /etc/grid-security/certificates/grid-security.conf.cd1186ff to 
/etc/grid-security/certificates/grid-security.conf

linking /etc/grid-security/certificates/grid-host-ssl.conf.cd1186ff to 
/etc/grid-security/certificates/grid-host-ssl.conf

linking /etc/grid-security/certificates/grid-user-ssl.conf.cd1186ff to 
/etc/grid-security/certificates/grid-user-ssl.conf

...done.

**Environment Variables**

The following environment variables affect the execution of `grid-default-ca`:

GRID_SECURITY_DIRECTORY  Path to the default trusted certificate directory.
X509_CERT_DIR  Path to the default trusted certificate directory.
GLOBUS_LOCATION  Path to the Globus Toolkit installation directory.

**Bugs**

The `grid-default-ca` program displays CAs from all of the directories in its search list; however, `grid-cert-request` only uses the first which contains a grid security configuration.

The `grid-default-ca` program may display the same CA multiple times if it is located in multiple directories in its search path. However, it does not provide any information about which one would actually be used by the `grid-cert-request` command.

**See Also**

`grid-cert-request(1)`
Name

grid-change-pass-phrase -- Change the passphrase of a private key

grid-change-pass-phrase [-file PRIVATE-KEY]

Description

The grid-change-pass-phrase program changes the passphrase protecting a private key or PKCS12 bundle containing a private key and certificate. By default, grid-change-pass-phrase uses the X509_USER_KEY environment variable to locate the private key. If that is not set, then it looks for $HOME/.globus/userkey.pem and $HOME/.globus/usercred.p12 in succession. The path to a key can be specified by using the -file command-line option.

The full set of command-line options to grid-change-pass-phrase are:

- -help, -usage Display the command-line options to grid-change-pass-phrase and exit.
- -version, -versions Display the version number of the grid-change-pass-phrase command. The second form includes more details.
- -file PRIVATE-KEY Change the passphrase of the private key named by PRIVATE-KEY instead of the default.

Examples

Change the passphrase of the default private key:

% grid-change-pass-phrase

Enter pass phrase for /home/juser/.globus/userkey.pem:
writing RSA key
Enter PEM pass phrase:
Verifying - Enter PEM pass phrase:

Environment Variables

The following environment variables affect the execution of grid-change-pass-phrase:

X509_USER_KEY Path to the default private key file.
Name

grid-proxy-init -- Generate a new proxy certificate

grid-proxy-init [-help] [-usage] [-version]
grid-proxy-init [-debug] [-q] [-verify]
[-policy POLICYFILE]
([-draft] | [-old] | [-rfc])

Description

The grid-proxy-init program generates X.509 proxy certificates derived from the currently available certificate files. By default, this command generates a RFC 3820 Proxy Certificate with a 512 bit key valid for 12 hours in a file named /tmp/x509up_uUID. Command-line options and variables can modify the format, strength, lifetime, and location of the generated proxy certificate.

X.509 proxy certificates are short-lived certificates, signed usually by a user's identity certificate or another proxy certificate. The key associated with a proxy certificate is unencrypted, so applications can authenticate using a proxy identity without providing a passphrase.

Proxy certificates provide a convenient alternative to constantly entering passwords, but are also less secure than the user's normal security credential. Therefore, they should always be user-readable only (this is enforced by the GSI libraries), and should be deleted after they are no longer needed.

This version of grid-proxy-init supports three different proxy formats: the old proxy format used in early releases of the Globus Toolkit up to version 2.4.x, an IETF draft version of X.509 Proxy Certificate profile used in Globus Toolkit 3.0.x and 3.2.x, and the RFC 3820 profile used in Globus Toolkit Version 4.0.x and 4.2.x. By default, this version of grid-proxy-init creates an RFC 3820 compliant proxy. To create a proxy compatible with older versions of the Globus Toolkit, use the -old or -draft command-line options.

The full set of command-line options to grid-proxy-init are:

- help, -usage  Display the command-line options to grid-proxy-init.
- version       Display the version number of the grid-proxy-init command
- debug         Display information about the path to the certificate and key used to generate the proxy certificate, the path to the trusted certificate directory, and verbose error messages
- q             Suppress all output from grid-proxy-init except for passphrase prompts.
- verify        Perform certificate chain validity checks on the generated proxy.
- valid HOURS:MINUTES, -hours HOURS  Create a certificate that is valid for HOURS hours and MINUTES minutes. If not specified, the default of twelve hours and no minutes is used.

1 http://www.ietf.org/rfc/rfc3820.txt
-cert CERTFILE, -key KEYFILE
Create a proxy certificate signed by the certificate located in CERTFILE using the key located in KEYFILE. If not specified the default certificate and key will be used. This overrides the values of environment variables described below.

-certdir CERTDIR
Search CERTDIR for trusted certificates if verifying the proxy certificate. If not specified, the default trusted certificate search path is used. This overrides the value of the X509_CERT_DIR environment variable.

-out PROXYPATH
Write the generated proxy certificate file to PROXYPATH instead of the default path of /tmp/x509up_uUID.

-bits BITS
When creating the proxy certificate, use a BITS bit key instead of the default 512 bit keys.

-policy POLICYFILE
Add the certificate policy data described in POLICYFILE as the ProxyCertInfo X.509 extension to the generated proxy certificate.

-pl POLICY-OID, -policy-language POLICY-OID
Set the policy language identifier of the policy data specified by the -policy command-line option to the oid specified by the POLICY-OID string.

-path-length MAXIMUM
Set the maximum length of the chain of proxies that can be created by the generated proxy to MAXIMUM. If not set, the default of an unlimited proxy chain length is used.

-pwstdin
Read the private key's passphrase from stdin instead of reading input from the controlling tty. This is useful when scripting grid-proxy-init.

-limited
Create a limited proxy. Limited proxies are generally refused by process-creating services, but may be used to authorize with other services.

-independent
Create an independent proxy. An independent proxy is not treated as an impersonation proxy but as a separate identity for authorization purposes.

-draft
Create a IETF draft proxy instead of the default RFC 3280-compliant proxy. This type of proxy uses a non-standard proxy policy identifier. This might be useful for authenticating with older versions of the Globus Toolkit.

-old
Create a legacy proxy instead of the default RFC 3280-compliant proxy. This type of proxy uses a non-standard method of indicating that the certificate is a proxy and whether it is limited. This might be useful for authenticating with older versions of the Globus Toolkit.

-rfc
Create an RFC 3820-compliant proxy certificate. This is the default for this version of grid-proxy-init.

Examples

To create a proxy with the default lifetime and format, run the grid-proxy-init program with no arguments. For example:

% grid-proxy-init
Your identity: /DC=org/DC=example/CN=Joe User
Enter GRID pass phrase for this identity:
Creating proxy ............................... Done
Your proxy is valid until: Thu Mar 18 03:48:05 2010
To create a stronger proxy that lasts for only 8 hours, use the `-hours` and `-bits` command-line options to `grid-proxy-init`. For example:

```
% grid-proxy-init -hours 8 -bits 1024
Your identity: /DC=org/DC=example/CN=Joe User
Enter GRID pass phrase for this identity:
Creating proxy .................................. Done
Your proxy is valid until: Thu Mar 17 23:48:05 2010
```

## Environment Variables

The following environment variables affect the execution of `grid-proxy-init`:

- `X509_USER_CERT`: Path to the certificate to use as issuer of the new proxy.
- `X509_USER_KEY`: Path to the key to use to sign the new proxy.
- `X509_CERT_DIR`: Path to the directory containing trusted certificate certificates and signing policies.

## Files

The following files affect the execution of `grid-proxy-init`:

- `$HOME/.globus/user-cert.pem`: Default path to the certificate to use as issuer of the new proxy.
- `$HOME/.globus/userkey.pem`: Default path to the key to use to sign the new proxy.

## Compatibility

For more information about proxy certificate types and their compatibility in GT, see [http://dev.globus.org/wiki/Security/ProxyCertTypes](http://dev.globus.org/wiki/Security/ProxyCertTypes).

## See Also

`grid-proxy-destroy(1)`, `grid-proxy-info(1)`
Name

grid-proxy-destroy -- Destroy the default proxy certificate

grid-proxy-destroy [-help] [-usage] [-version]

Description

The grid-proxy-destroy program removes X.509 proxy files from the local filesystem. It overwrites the data in the files and removes the files from the filesystem. By default, it removes the current user's default proxy (either /tmp/x509up_uUID where UID is the current POSIX user id, or the file pointed to by the X509_USER_PROXY environment variable) unless a list of proxy file paths are included as part of the command line.

Use the -- command-line option to separate a list of proxy paths from command line options if the proxy file begins with the -- character.

The full list of command-line options to grid-proxy-destroy are:

- `-help`, `-usage` Display the command-line options to grid-proxy-destroy.
- `-version` Display the version number of the grid-proxy-destroy command
- `-debug` Display verbose error messages.
- `-dryrun` Do not remove the proxy, but display the path of the files that would have been removed, or the directory where they would have been removed from if the -all command-line option is used.
- `-default` Remove the default proxy in addition to the files included on the command-line. Only needed if other paths are included on the command-line.
- `-all` Remove the default proxy and all delegated proxies in the temporary file directory.

Environment Variables

The following environment variables affect the execution of grid-proxy-destroy:

X509_USER_PROXY Path to the default user proxy.

See Also

grid-proxy-init(1), grid-proxy-info(1)
Name

grid-proxy-info -- Display information about a proxy certificate

grid-proxy-info [-help] [-usage] [-version]
grid-proxy-info [[-subject] [-s]]
[-issuer] [-i]
[-exists | -e]
([-valid HOURS:MINUTES] | [-v HOURS:MINUTES])
([-hours HOURS] | [-h HOURS])
([-bits BITS] | [-b BITS]])

Description

The grid-proxy-info program extracts information from an X.509 proxy certificates, and optionally displays or returns an exit code based on that information.

The default mode of operation is to print the following facts about the current user's default proxy: subject, issuer, identity, type, strength, path, and time left. If the command-line option -exists or -e is included in the command-line, nothing is printed unless one of the print options is specified. Instead, grid-proxy-info determines if a valid proxy exists and, if so, exits with the exit code 0; if a proxy does not exist or is not valid, grid-proxy-info exits with the exit code 1. Additional validity criteria can be added by using the -valid, -v, -hours, -h, -bits, or -b command-line options. If used, these options must occur after the -e or -exists command-line options. Those options are only valid if one of the -e or -exists command-line options is used.

The complete set of command-line options to grid-proxy-info are:

- -help, -usage Display the command-line options to grid-proxy-info.
- -version Display the version number of the grid-proxy-info command
- -debug Display verbose error messages.
- -file PROXYFILE, -f PROXYFILE Read the proxy located in the file PROXYFILE instead of using the default proxy.
- -subject, -s Display the proxy certificate's subject distinguished name.
- -issuer, -i Display the proxy certificate issuer's distinguished name.
- -identity Display the proxy certificate's identity. For non-independent proxies, the identity is the subject of the certificate which issued the first proxy in the proxy chain.
- -type Display the type of proxy certificate. The type string includes the format ("legacy", "draft", or RFC 3280 compliant), identity type ("impersonation" or "independent"), and policy ("limited" or "full"). See grid-proxy-init(1) for information about how to create different types of proxies.
- -timeleft Display the number of seconds remaining until the proxy certificate expires.
- -strength Display the strength (in bits) of the key associated with the proxy certificate.
- -all Display the default information for the proxy when also using the -e or -exists command-line option.
grid-proxy-info

-text Display the proxy certificate contents to standard output, including policy information, issuer, public key, and modulus.
-path Display the path to the file containing the default proxy certificate.
-rfc2253 Display distinguished names for the subject, issuer, and identity using the string representation described in RFC 2253, instead of the legacy format.
-exists, -e Perform an existence and validity check for the proxy. If a valid proxy exists and matches the criteria described by other command-line options (if any), exit with 0; otherwise, exit with 1. This option must be before other validity check predicates in the command-line options. If this option is specified, the output of the default facts about the proxy is disabled. Use the -all option to have the information displayed as well as the exit code set.

-valid HOURS:MINUTES, -v HOURS:MINUTES, -hours HOURS, -h HOURS Check that the proxy certificate is valid for at least HOURS hours and MINUTES minutes. If it is not, grid-proxy-info will exit with exit code 1.
-bits BITS, -b BITS Check that the proxy certificate key strength is at least BITS bits.

Environment Variables

The following environment variables affect the execution of grid-proxy-info:

X509_USER_PROXY Path to the default user proxy.

See Also

grid-proxy-init(1), grid-proxy-destroy(1)
Name

grid-mapfile-add-entry -- Add an entry to a gridmap file

grid-mapfile-add-entry [-dn DISTINGUISHED-NAME] [-ln LOCAL-NAME...]  
[[-d] | [-dryrun]]  
[[-mapfile MAPFILE] | [-f MAPFILE]]

Description

The **grid-mapfile-add-entry** program adds a new mapping from an X.509 distinguished name to a local POSIX user name to a gridmap file. Gridmap files are used as a simple authorization method for services such as GRAM5 or GridFTP.

The **grid-mapfile-add-entry** program verifies that the LOCAL-NAME is a valid user name on the system on which it was run, and that the mapping between DISTINGUISHED-NAME and LOCAL-NAME does not already exist in the gridmap file.

By default, **grid-mapfile-add-entry** will modify the gridmap file named by the GRIDMAP environment variable if present, or the file /etc/grid-security/grid-mapfile if not. This can be changed by the use of the `-mapfile` or `-f` command-line options.

If the gridmap file does not exist, **grid-mapfile-add-entry** will create it. If it already exists, **grid-mapfile-add-entry** will save the current contents of the file to a new file with the string `.old` appended to the file name.

The full set of command-line options to **grid-mapfile-add-entry** are:

- **-help, -usage** Display the command-line options to **grid-mapfile-add-entry**.
- **-version, -versions** Display the version number of the **grid-mapfile-add-entry** command. The second form includes more details.
- **-dn DISTINGUISHED-NAME** The X.509 distinguished name to add a mapping for. The name should be in OpenSSL's one line format.
- **-ln LOCAL-NAME...** The POSIX user name to map the distinguished name to. This name must be a valid username. Add multiple LOCAL-NAME strings after the `-ln` command-line option. If any of the local names are invalid, no changes will be made to the gridmap file. Note that if multiple occurrences of the `-ln` command-line option are present, only the the last one will be added.
- **-d, -dryrun** Verify local names and display diagnostics about what would be added to the gridmap file, but don't actually modify the file.
- **-mapfile MAPFILE, -f MAPFILE** Modify the gridmap file named by MAPFILE instead of the default.

Examples

Add a mapping between the current user's certificate to the current user id to a gridmap file in $HOME/.gridmap:

```
% grid-mapfile-add-entry -f $HOME/.gridmap -dn `grid-cert-info -subject` -ln `id -un`
Modifying /home/juser/.gridmap ...
```
Add a mapping between the distinguished name and multiple local names:

```
% grid-mapfile-add-entry -dn "/DC=org/DC=example/DC=grid/CN=Joe User" juser local1 local2
```

New entry:

"/DC=org/DC=example/DC=grid/CN=Joe User" local1,local2

(1) entry added

Environment Variables

The following environment variables affect the execution of `grid-mapfile-add-entry`:

GRIDMAP Path to the default gridmap to modify.

Files

The following files affect the execution of `grid-mapfile-add-entry`:

```
/etc/grid-security/grid-mapfile
```

Path to the default gridmap to modify if GRIDMAP environment variable is not set.

See Also

grid-mapfile-check-consistency(8), grid-mapfile-delete-entry(8)
Name

grid-mapfile-check-consistency -- Add an entry to a grid map file

grid-mapfile-check-consistency [-mapfile MAPFILE] [-f MAPFILE]

Description

The grid-mapfile-check-consistency program performs basic checks for validity of a gridmap file. These checks include checks for existence, duplication of entries, and valid local user names. If the gridmap file is valid, grid-mapfile-check-consistency exits with a zero exit code, otherwise it exits with a non-zero exit code. In either case, it displays information about its progress as it parses and validates the gridmap file.

By default, grid-mapfile-check-consistency will check the gridmap file named by the GRIDMAP environment variable if present. If that variable is not set, it will check the file $HOME/.gridmap for non-root users if present. If that doesn’t exist or grid-mapfile-check-consistency is run as root, it will then check /etc/grid-security/grid-mapfile. This can be changed by the use of the -mapfile or -f command-line options.

The full set of command-line options to grid-mapfile-check-consistency are:

- help, -h, -usage Display the command-line options to grid-mapfile-check-consistency.
- version Display the version number of the grid-mapfile-check-consistency command.
- mapfile MAPFILE, -f MAPFILE Check the gridmap file named by MAPFILE instead of the default.

Examples

Check that the gridmap file in /etc/grid-security is valid:

% grid-mapfile-check-consistency -f /etc/grid-security/grid-mapfile
Checking /etc/grid-security/grid-mapfile
Verifying grid mapfile existence...OK
Checking for duplicate entries...OK
Checking for valid user names...OK

Check a gridmap file that has an invalid local user name:

% grid-mapfile-check-consistency -f /etc/grid-security/grid-mapfile
Checking /etc/grid-security/grid-mapfile
Verifying grid mapfile existence...OK
Checking for duplicate entries...OK
ERROR: baduser is not a valid local username
ERROR: Found 1 invalid username(s)

Environment Variables

The following environment variables affect the execution of grid-mapfile-check-consistency:

GRIDMAP Path to the default gridmap to check.
Files

The following files affect the execution of `grid-mapfile-check-consistency`:

- **$HOME/.gridmap**  
  Path to the default gridmap to check if the GRIDMAP environment variable is not set for non-root users.

- **/etc/grid-security/grid-mapfile**  
  Path to the default gridmap to check if GRIDMAP environment variable is not set and the above file does not exist.

See Also

`grid-mapfile-add-entry(8), grid-mapfile-delete-entry(8)`
Name

grid-mapfile-delete-entry -- Remove entries from a gridmap file

grid-mapfile-delete-entry {-dn DISTINGUISHED-NAME} {-ln LOCAL-NAME...}
[[-d] [-dryrun]]
[[-mapfile MAPFILE] [-f MAPFILE]]

Description

The **grid-mapfile-delete-entry** program deletes mappings from a gridmap file. If both the `-dn` and `-ln` options are specified, **grid-mapfile-delete-entry** removes entries which meet both criteria (remove entries mapping DISTINGUISHED-NAME to LOCAL-NAME for each LOCAL-NAME specified). If only `-dn` or `-ln` is specified all entries for that DISTINGUISHED-NAME or LOCAL-NAME are removed.

By default, **grid-mapfile-delete-entry** will modify the gridmap file named by the GRIDMAP environment variable if present, or the file `/etc/grid-security/grid-mapfile` if not. This can be changed by the use of the `-mapfile` or `-f` command-line options.

Prior to modifying a gridmap file, **grid-mapfile-delete-entry** saves its current contents to a file with the string `.old` appended to the original file name.

The full set of command-line options to **grid-mapfile-delete-entry** are:

- `-help`, `-usage` — Display the command-line options to **grid-mapfile-delete-entry**.
- `-version`, `-versions` — Display the version number of the **grid-mapfile-delete-entry** command. The second form includes more details.
- `-dn DISTINGUISHED-NAME` — The X.509 distinguished name to remove from the gridmap file. If the `-ln` option is not specified, remove all entries for this name; otherwise, remove entries that match both this name and the local name. The name should be in OpenSSL’s oneline format.
- `-ln LOCAL-NAME...` — The POSIX user name to remove from the gridmap file. Include multiple LOCAL-NAME strings after the `-ln` command-line option to remove multiple names from the gridmap. If the `-dn` option is not specified, remove all entries for these names; otherwise, remove entries that match the DISTINGUISHED-NAME and any of the LOCAL-NAME values.
- `-d`, `-dryrun` — Display diagnostics about what would be removed from the gridmap file, but don’t actually modify the file.
- `-mapfile MAPFILE`, `-f MAPFILE` — Modify the gridmap file named by MAPFILE instead of the default.

Examples

Remove all mappings for a distinguished name:

```
% grid-mapfile-delete-entry "/DC=org/DC=example/DC=grid/CN=Joe User"
Modifying /etc/grid-security/grid-mapfile ...
```
Deleting entry: "/DC=org/DC=example/DC=grid/CN=Joe User" juser, juser2
(1) entry deleted

Remove the mapping between a distinguished name and a single local username:

% grid-mapfile-delete-entry "/DC=org/DC=example/DC=grid/CN=Joe User" -ln juser2
Modifying /etc/grid-security/grid-mapfile ...
Current entry: "/DC=org/DC=example/DC=grid/CN=Joe User" juser
(1) mapping removed: (juser2), (0) not present and ignored
(0) entries deleted

Environment Variables

The following environment variables affect the execution of grid-mapfile-delete-entry:

GRIDMAP Path to the default gridmap to modify.

Files

The following files affect the execution of grid-mapfile-delete-entry:

/etc/grid-security/grid-mapfile Path to the default gridmap to modify if GRIDMAP environment variable is not set.

See Also

grid-mapfile-add-entry(8), grid-mapfile-check-consistency(8)
GridFTP Commands
Name
globus-url-copy -- Multi-protocol data movement
globus-url-copy

Tool description

globus-url-copy is a scriptable command line tool that can do multi-protocol data movement. It supports gsiftp:// (GridFTP), ftp://, http://, https://, and file:/// protocol specifiers in the URL. For GridFTP, globus-url-copy supports all implemented functionality. Versions from GT 3.2 and later support file globbing and directory moves.

• Before you begin
• Command syntax
• Command line options
  • Informational options
  • Utility options
  • Reliability options
  • Performance options
  • Security-related options
• Default usage
• MODES in GridFTP
• If you run a GridFTP server by hand
• How do I choose a value for the TCP buffer size (-tcp-bs) option?
• How do I choose a value for the parallelism (-p) option?
• Limitations
• Interactive clients for GridFTP

Before you begin

⚠️ Important

To use gsiftp:// and https:// protocols in globus-url-copy, you must have a certificate. However, you may use ftp://, http:// or sshftp:// protocols without a certificate.

1. First, as with all things Grid, you must have a valid proxy certificate to run globus-url-copy in certain protocols (gsiftp:// and https://, as noted above). If you are using ftp://, http:// or sshftp:// protocols, you may skip ahead to Command syntax

   If you do not have a certificate, you must obtain one.
If you are doing this for testing in your own environment, the SimpleCA provided with the Globus Toolkit should suffice.

If not, you must contact the Virtual Organization (VO) with which you are associated to find out whom to ask for a certificate.

One common source is the DOE Science Grid CA\(^1\), although you must confirm whether or not the resources you wish to access will accept their certificates.

Instructions for proper installation of the certificate should be provided from the source of the certificate.

Please note when your certificates expire; they will need to be renewed or you may lose access to your resources.

2. Now that you have a certificate, you must generate a temporary proxy. Do this by running:

```bash
grid-proxy-init
```

Further documentation for `grid-proxy-init` can be found [here](http://www.doegrids.org/pages/cert-request.htm).

3. You are now ready to use `globus-url-copy`! See the following sections for syntax and command line options and other considerations.

## Command syntax

The basic syntax for `globus-url-copy` is:

```bash
globus-url-copy [optional command line switches] Source_URL Destination_URL
```

where:

<table>
<thead>
<tr>
<th>[optional command line switches]</th>
<th>See Command line options below for a list of available options.</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>Source_URL</code></td>
<td>Specifies the original URL of the file(s) to be copied.</td>
</tr>
<tr>
<td></td>
<td>If this is a directory, all files within that directory will be copied.</td>
</tr>
<tr>
<td><code>Destination_URL</code></td>
<td>Specifies the URL where you want to copy the files.</td>
</tr>
<tr>
<td></td>
<td>If you want to copy multiple files, this must be a directory.</td>
</tr>
</tbody>
</table>

⚠️ **Note**

Any url specifying a directory must end with `/`.

## URL prefixes

Versions from GT 3.2 and later support the following URL prefixes:

- `file://` (on a local machine only)
- `ftp://`
- `gsiftp://`
- `http://`

\(^1\) [http://www.doegrids.org/pages/cert-request.htm](http://www.doegrids.org/pages/cert-request.htm)
• https://

Versions from GT 4.2 and later support the following URL prefix (in addition to the above-mentioned URL prefixes):

• sshftp://

🎉 Note

We do not provide an interactive client similar to the generic FTP client provided with Linux. See the Interactive Clients section below for information on an interactive client developed by NCSA/NMI/TeraGrid.

**URL formats**

URLs can be any valid URL as defined by RFC 1738 that have a protocol we support. In general, they have the following format: `protocol://host:port/path`.

🎉 Note

If the path ends with a trailing `/` (i.e. `/path/to/directory/`) it will be considered to be a directory and all files in that directory will be moved. If you want a recursive directory move, you need to add the `-r/-recurse` switch described below.

**Table 1. URL formats**

<table>
<thead>
<tr>
<th>URL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://myhost.mydomain.com/mywebpage/default.html">http://myhost.mydomain.com/mywebpage/default.html</a></td>
<td>Port is not specified; therefore, GridFTP uses protocol default (in this case, 80).</td>
</tr>
<tr>
<td>file:///foo.dat</td>
<td>Host is not specified; therefore, GridFTP uses your local host.</td>
</tr>
<tr>
<td>file://foo.dat</td>
<td>Port is not specified; therefore, GridFTP uses protocol default (in this case, 80).</td>
</tr>
<tr>
<td></td>
<td>This is also valid but is not recommended because, while many servers (including ours) accept this format, it is not RFC conformance and is not recommended.</td>
</tr>
</tbody>
</table>

⚠️ Important

For GridFTP (`gsiftp://`) and FTP (`ftp://`), it is legal to specify a user name and password in the URL as follows:

`gsiftp://myname:[mypassword]@myhost.mydomain.com/foo.dat`

If you are using GSI security, then you may specify the username (but you may not include the : or the password) and the grid-mapfile will be searched to see if that is a valid account mapping for your distinguished name (DN). If it is found, the server will setuid to that account. If not, it will fail. It will NOT fail back to your default account.

If you are using anonymous FTP, the username **must** be one of the usernames listed as a valid anonymous name and the password can be anything.
If you are using password authentication, you must specify both your username and password. THIS IS HIGHLY DISCOURAGED, AS YOU ARE SENDING YOUR PASSWORD IN THE CLEAR ON THE NETWORK. This is worse than no security; it is a false illusion of security.

Command line options

Informational Options

-`-help | -usage`  Prints help.

-`-version`  Prints the version of this program.

-`-versions`  Prints the versions of all modules that this program uses.

-`-q | -quiet`  Suppresses all output for successful operation.

-`-vb | -verbose`  During the transfer, displays:

  • number of bytes transferred,

  • performance since the last update (currently every 5 seconds), and

  • average performance for the whole transfer.

-`-dbg | -debugftp`  Debugs FTP connections and prints the entire control channel protocol exchange to STDERR.

  Very useful for debugging. Please provide this any time you are requesting assistance with a `globus-url-copy` problem.

-`-list <url>`  This option will display a directory listing for the given url.

-`-nl-bottleneck | -nlb`  This option uses NetLogger to estimate speeds of disk and network read/write system calls, and attempt to determine the bottleneck component.

Note

In order to use this, the server must be configured to enable netlogger bottleneck detection.

Utility Ease of Use Options

-`-a | -ascii`  Converts the file to/from ASCII format to/from local file format.

-`-b | -binary`  Does not apply any conversion to the files. This option is turned on by default.

-`-cd | -create-dest`  Create destination directories, if needed

-`-f filename`  Reads a list of URL pairs from a filename.

  Each line should contain:

  \[sourceURL destURL\]

---

Enclose URLs with spaces in double quotes ("). Blank lines and lines beginning
with the hash sign (#) will be ignored.

- r | - recurse
Copies files in subdirectories.

-rp | -relative-paths
The path portion of ftp urls will be interpreted as relative to the user’s starting
directory on the server. By default, all paths are root-relative. When this flag is
set, the path portion of the ftp url must start with %2F if it designates a root-re-

tative path.

-notpt | - no-third-party-transfers
 Turns third-party transfers off (on by default).

Site firewall and/or software configuration may prevent a connection between
the two servers (a third party transfer). If this is the case, globus-url-copy will
"relay" the data. It will do a GET from the source and a PUT to the destination.

This obviously causes a performance penalty but will allow you to complete a
transfer you otherwise could not do.

Reliability Options

-rst | -restart
Restarts failed FTP operations.

-rst-retries <retries>
Specifies the maximum number of times to retry the operation before giving up
on the transfer.

Use 0 for infinite.

The default value is 5.

-rst-interval <seconds>
Specifies the interval in seconds to wait after a failure before retrying the
transfer.

Use 0 for an exponential backoff.

The default value is 0.

-rst-timeout <seconds>
Specifies the maximum time after a failure to keep retrying.

Use 0 for no timeout.

The default value is 0.

-df <filename> | -dumpfile <file-
name>
Specifies path to the file where untransferred urls will be saved for later restart-
ing. The resulting file is the same format as the -f input file. If the file exists,
it will be read and all other url input will be ignored.

-do <filename> | -dump-only <file-
name>
Perform no write operations on the destination. Instead, all files that would be
transferred are enumerated and dumped to the specified file. Resulting file is
the same format as the -f input file. Note: if you intend to use this file as input
for a future transfer, the -create-dest option will be required if any destination
directories do not already exist.

-stall-timeout | -st <seconds>
Specifies how long before cancelling/restarting a transfer with no data movement.
Set to 0 to disable. Default is 600 seconds.
Performance Options

-tcp-bs <size> | -tcp-buffer-size <size>

Specifies the size (in bytes) of the TCP buffer to be used by the underlying ftp data channels.

⚠️ Important

This is critical to good performance over the WAN.

How do I pick a value?

-p <parallelism> | -parallel <parallelism>

Specifies the number of parallel data connections that should be used.

話し

This is one of the most commonly used options.

How do I pick a value?

-bs <block size> | -block-size <block size>

 Specifies the size (in bytes) of the buffer to be used by the underlying transfer methods.

-pp

Allows pipelining. GridFTP is a command response protocol. A client sends one command and then waits for a "Finished response" before sending another. Adding this overhead on a per-file basis for a large data set partitioned into many small files makes the performance suffer. Pipelining allows the client to have many outstanding, unacknowledged transfer commands at once. Instead of being forced to wait for the "Finished response" message, the client is free to send transfer commands at any time.

-mc filename source_url

Transfers a single file to many destinations. Filename is a line-separated list of destination urls. For more information on this option, click here.

Multicasting must be enabled for use on the server side.

⚠️ Warning

This option is EXPERIMENTAL.

-concurrency | -cc

Specifies the number of concurrent FTP connections to use for multiple transfers.

-udt

Uses UDT, a reliable UDP-based transport protocol, for data transfers.

-fast

Recommended when using GridFTP servers. Use MODE E for all data transfers, including reusing data channels between list and transfer operations.

Note: In order to use this option, the server must be configured to use UDT. For third party transfers, no change is required on the client side. For client-server transfers, you need the threaded flavor of the client. Refer to Switching between threaded and non-threaded flavors for information on how to switch between threaded and non-threaded flavors of globus-url-copy.
Security Related Options

-s <subject> | -subject <subject>

Specifies a subject to match with both the source and destination servers.

Note

Used when the server does not have access to the host certificate (usually when you are running the server as a user). See the section called “If you run a GridFTP server by hand...”.

-ss <subject> | -source-subject <subject>

Specifies a subject to match with the source server.

Note

Used when the server does not have access to the host certificate (usually when you are running the server as a user). See the section called “If you run a GridFTP server by hand...”.

-ds <subject> | -dest-subject <subject>

Specifies a subject to match with the destination server.

Note

Used when the server does not have access to the host certificate (usually when you are running the server as a user). See the section called “If you run a GridFTP server by hand...”.

-nodcau | -no-data-channel-authentication

Turns off data channel authentication for FTP transfers (the default is to authenticate the data channel).

Warning

We do not recommend this option, as it is a security risk.

-dcsafe | -data-channel-safe

Sets data channel protection mode to SAFE.

Otherwise known as integrity or checksumming.

Guarantees that the data channel has not been altered, though a malicious party may have observed the data.

Warning

Rarely used as there is a substantial performance penalty.

-dcpriv | -data-channel-private

Sets data channel protection mode to PRIVATE.

The data channel is encrypted and checksummed.

Guarantees that the data channel has not been altered and, if observed, it won't be understandable.
Warning

VERY rarely used due to the VERY substantial performance penalty.

Advanced Options

- stripe
  Enables striped transfers on supported servers.

- striped-block-size | -sbs
  Sets layout mode and blocksize for striped transfers.
  If not set, the server defaults will be used.
  If set to 0, partitioned mode will be used.
  If set to >0, blocked mode will be used, with this setting used as the blocksize.

- t <transfer time in seconds>
  Runs the transfer for the specified number of seconds and then ends. Useful for performance testing or forced restart loops.

- ipv6
  Uses ipv6 when available.

Warning

This option is EXPERIMENTAL. Use at your own risk.

- dp | -delayed-pasv
  Enables delayed passive.

- g2 | -gridftp2
  Uses GridFTP v2 protocol enhancements when possible.

- mn | -module-name <gridftp storage module name>
  Specifies the backend storage module to use for both the source and destination in a GridFTP transfer.

- mp | -module-parameters <gridftp storage module parameters>
  Specifies the backend storage module arguments to use for both the source and destination in a GridFTP transfer.

- smn | -src-module-name <gridftp storage module name>
  Specifies the backend storage module to use for the source file in a GridFTP transfer.

- smp | -src-module-parameters <gridftp storage module parameters>
  Specifies the backend storage module arguments to use for the source file in a GridFTP transfer.

- dmn | -dst-module-name <gridftp storage module name>
  Specifies the backend storage module to use for the destination file in a GridFTP transfer.

- dmp | -dst-module-parameters <gridftp storage module parameters>
  Specifies the backend storage module arguments to use for the destination file in a GridFTP transfer.

- aa | -authz-assert <authorization assertion file>
  Uses the assertions in the specified file to authorize access to both the source and destination servers.

- saa | -src-authz-assert <authorization assertion file>
  Uses the assertions in the specified file to authorize access to the source server.
-daa | -dst-authz-assert <authorization assertion file>  Uses the assertions in the specified file to authorize access to the destination server.

-cache-aa | -cache-authz-assert  Caches the authorization assertion for subsequent transfers.

-cache-saa | -cache-src-authz-assert  Caches the source authorization assertion for subsequent transfers.

-cache-daa | -cache-dst-authz-assert  Caches the destination authorization assertion for subsequent transfers.

-nl-bottleneck | -nlb  Uses NetLogger to estimate speeds of disk and network read/write system calls, and attempt to determine the bottleneck component.

Note: In order to use this, the server must be configured to enable netlogger bottleneck detection.

-src-pipe | -SP <command line>  Sets the source end of a remote transfer to use piped-in input with the given command line.

⚠️ Warning

Do not use with the -fsstack option.

-dst-pipe | -DP <command line>  Sets the destination end of a remote transfer to write data to then standard input of the program run via the given command line.

⚠️ Warning

Do not use with the -fsstack option.

-pipe <command line>  Sets both -src-pipe and -dst-pipe to the same value.

dcstack | -data-channel-stack  Specifies the XIO driver stack for the network on both the source and the destination. Both must be GridFTP servers.

-fsstack | -file-system-stack  Specifies the XIO driver stack for the disk on both the source and the destination. Both must be GridFTP servers.

-src-dcstack | -source-data-channel-stack  Specifies the XIO driver stack for the network on the source GridFTP server.

-src-fsstack | -source-file-system-stack  Specifies the XIO driver stack for the disk on the source GridFTP server.

-dst-dcstack | -dest-data-channel-stack  Specifies the XIO driver stack for the network on the destination GridFTP server.

-dst-fsstack | -dest-file-system-stack  Specifies the XIO driver stack for the disk on the destination GridFTP server.

-cred <path to credentials or proxy file>, -src-cred | -sc <path to credentials or proxy file>, -dst-cred | -dc <path to credentials or proxy file>  Specifies the credentials to use for source, destination, or both FTP connections.
-af <filename> | -alias-file <filename>

Specifies a file that maps logical host aliases to lists of physical hosts. When used with multiple concurrent connections, each connection uses the next host in the list. Each line should either be an alias (noted with the @ symbol), or a hostname[:port]. Currently, only the aliases @source and @destination are valid, and they are used for every source or destination url.

Synchronization Options

-sync

Only transfer files where the destination does not exist or differs from the source. -sync-level controls how to determine if files differ.

-sync-level <number>

Choose criteria for determining if files differ when performing a sync transfer. Level 0 will only transfer if the destination does not exist. Level 1 will transfer if the size of the destination does not match the size of the source. Level 2 will transfer if the timestamp of the destination is older than the timestamp of the source. Level 3 will perform a checksum of the source and destination and transfer if the checksums do not match. The default sync level is 2.

Default globus-url-copy usage

A globus-url-copy invocation using the gsiftp protocol with no options (i.e., using all the defaults) will perform a transfer with the following characteristics:

- binary
- stream mode (which implies no parallelism)
- host default TCP buffer size
- encrypted and checksummed control channel
- an authenticated data channel

MODES in GridFTP

GridFTP (as well as normal FTP) defines multiple wire protocols, or MODES, for the data channel.

Most normal FTP servers only implement stream mode (MODE S), i.e. the bytes flow in order over a single TCP connection. GridFTP defaults to this mode so that it is compatible with normal FTP servers.

However, GridFTP has another MODE, called Extended Block Mode, or MODE E. This mode sends the data over the data channel in blocks. Each block consists of 8 bits of flags, a 64 bit integer indicating the offset from the start of the transfer, and a 64 bit integer indicating the length of the block in bytes, followed by a payload of length bytes. Because the offset and length are provided, out of order arrival is acceptable, i.e. the 10th block could arrive before the 9th because you know explicitly where it belongs. This allows us to use multiple TCP channels. If you use the -p | -parallelism option, globus-url-copy automatically puts the servers into MODE E.

Note

Putting -p 1 is not the same as no -p at all. Both will use a single stream, but the default will use stream mode and -p 1 will use MODE E.
If you run a GridFTP server by hand...

If you run a GridFTP server by hand, you will need to explicitly specify the subject name to expect. The subject option provides `globus-url-copy` with a way to validate the remote servers with which it is communicating. Not only must the server trust `globus-url-copy`, but `globus-url-copy` must trust that it is talking to the correct server. The validation is done by comparing host DNs or subjects.

If the GridFTP server in question is running under a host certificate then the client assumes a subject name based on the server's canonical DNS name. However, if it was started under a user certificate, as is the case when a server is started by hand, then the expected subject name must be explicitly stated. This is done with the `-ss`, `-sd`, and `-s` options.

- `-ss` Sets the sourceURL subject.
- `-ds` Sets the destURL subject.
- `-s` If you use this option alone, it will set both urls to be the same. You can see an example of this usage under the Troubleshooting section.

Note

This is an unusual use of the client. Most times you need to specify both URLs.

How do I choose a value?

How do I choose a value for the TCP buffer size (`-tcp-bs`) option?

The value you should pick for the TCP buffer size (`-tcp-bs`) depends on how fast you want to go (your bandwidth) and how far you are moving the data (as measured by the Round Trip Time (RTT) or the time it takes a packet to get to the destination and back).

To calculate the value for `-tcp-bs`, use the following formula (this assumes that Mega means 1000^2 rather than 1024^2, which is typical for bandwidth):

```
-tcp-bs = bandwidth in Megabits per second (Mbs) * RTT in milliseconds (ms) * 1000 / 8
```

As an example, if you are using fast ethernet (100 Mbs) and the RTT was 50 ms it would be:

```
-tcp-bs = 100 * 50 * 1000 / 8 = 625,000 bytes.
```

So, how do you come up with values for bandwidth and RTT? To determine RTT, use either ping or traceroute. They both list RTT values.

Note

You must be on one end of the transfer and ping the other end. This means that if you are doing a third party transfer you have to run the ping or traceroute between the two server hosts, not from your client.

The bandwidth is a little trickier. Any point in the network can be the bottleneck, so you either need to talk with your network engineers to find out what the bottleneck link is or just assume that your host is the bottleneck and use the speed of your network interface card (NIC).
Note

The value you pick for \(-tcp-bs\) limits the top speed you can achieve. You will NOT get bandwidth any higher than what you used in the calculation (assuming the RTT is actually what you specified; it varies a little with network conditions). So, if for some reason you want to limit the bandwidth you get, you can do that by judicious choice of \(-tcp-bs\) values.

So where does this formula come from? Because it uses the bandwidth and the RTT (also known as the latency or delay) it is called the bandwidth delay product. The very simple explanation is this: TCP is a reliable protocol. It must save a copy of everything it sends out over the network until the other end acknowledges that it has been received.

As a simple example, if I can put one byte per second onto the network, and it takes 10 seconds for that byte to get there, and 10 seconds for the acknowledgment to get back (RTT = 20 seconds), then I would need at least 20 bytes of storage. Then, hopefully, by the time I am ready to send byte 21, I have received an acknowledgement for byte 1 and I can free that space in my buffer. If you want a more detailed explanation, try the following links on TCP tuning:

- [http://www.psc.edu/networking/perf_tune.html](http://www.psc.edu/networking/perf_tune.html)
- [http://www.ncne.nlanr.net/research/tcp/](http://www.ncne.nlanr.net/research/tcp/)

How do I choose a value for the parallelism (\(-p\)) option?

For most instances, using 4 streams is a very good rule of thumb. Unfortunately, there is not a good formula for picking an exact answer. The shape of the graph shown here is very characteristic.

**Figure 1. Effect of Parallel Streams in GridFTP**

You get a strong increase in bandwidth, then a sharp knee, after which additional streams have very little impact. Where this knee is depends on many things, but it is generally between 2 and 10 streams. Higher bandwidth, longer round trip times, and more congestion in the network (which you usually can only guess at based on how applications are behaving) will move the knee higher (more streams needed).

In practice, between 4 and 8 streams are usually sufficient. If things look really bad, try 16 and see how much difference that makes over 8. However, anything above 16, other than for academic interest, is basically wasting resources.

Limitations

There are no limitations for **globus-url-copy** in GT 5.0.3.
Interactive clients for GridFTP

The Globus Project does not provide an interactive client for GridFTP. Any normal FTP client will work with a GridFTP server, but it cannot take advantage of the advanced features of GridFTP. The interactive clients listed below take advantage of the advanced features of GridFTP.

There is no endorsement implied by their presence here. We make no assertion as to the quality or appropriateness of these tools, we simply provide this for your convenience. We will not answer questions, accept bugs, or in any way shape or form be responsible for these tools, although they should have mechanisms of their own for such things.

UberFTP was developed at the NCSA under the auspices of NMI and TeraGrid:

- NCSA Uberftp only download: http://dims.ncsa.uiuc.edu/set/uberftp/download.html
Name

globus-url-sync -- Used in conjunction with globus-url-copy to synchronize directories.

globus-url-sync

Tool description

globus-url-sync is a command line tool which provides a list of files to be transferred, in order to synchronize two directories. It currently supports gsiftp:// (GridFTP) and sshftp:// protocol specifiers in the URL.

The program globus-url-sync compares two endpoints, using GridFTP, and prints a list of GSI file transfers that should be performed using globus-url-copy.

The current implementation of globus-url-sync supports very basic features for directory synchronization. It includes comparators for existence checks, file size checks, modification timestamp checks, but not checksum comparison.

Before you begin

1. First, as with globus-url-copy, you must have a valid proxy certificate to run globus-url-sync using protocol "gsiftp://".

   If you do not have a certificate, you must obtain one.

   If you are doing this for testing in your own environment, the SimpleCA provided with the Globus Toolkit should suffice.

   If not, you must contact the Virtual Organization (VO) with which you are associated to find out whom to ask for a certificate.

   One common source is the DOE Science Grid CA\(^1\), although you must confirm whether or not the resources you wish to access will accept their certificates.

   Instructions for proper installation of the certificate should be provided from the source of the certificate.

   Please note when your certificates expire; they will need to be renewed or you may lose access to your resources.

2. Now that you have a certificate, you must generate a temporary proxy. Do this by running:

   grid-proxy-init

   Further documentation for grid-proxy-init can be found here.

\(^1:\) http://www.doegrids.org/pages/cert-request.htm
Command syntax

The basic syntax for globus-url-sync is:

```
globus-url-sync [optional command line switches] Source_URL Destination_URL
```

where:

<table>
<thead>
<tr>
<th>[optional command line switches]</th>
<th>See Command line options below for a list of available options.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source_URL</td>
<td>Specifies the original URL of the file(s) to be copied.</td>
</tr>
<tr>
<td></td>
<td>If this is a directory, all files within that directory that need to be synchronized will be listed.</td>
</tr>
<tr>
<td>Destination_URL</td>
<td>Specifies the URL where you want to copy the files.</td>
</tr>
<tr>
<td></td>
<td>The types of the source and the destination must match. In other words, if the source is a file, the destination must be a file, and if the source is a directory, the destination must be a directory.</td>
</tr>
</tbody>
</table>

⚠️  Note

Any url specifying a directory must end with `/`.

URL prefixes

The following URL prefixes are supported:

- `gsiftp://`
- `sshftp://`

URL formats

URLs can be any valid URL as defined by RFC 1738 that have a protocol we support. In general, they have the following format: `protocol://host:port/path`.

⚠️  Note

If the path ends with a trailing `/` (i.e. `/path/to/directory/`) it will be considered to be a directory and all files in that directory that are not synchronized will be listed.

Table 2. URL formats

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gsiftp://myhost.mydomain.com/tmp/file1</td>
<td>File name, absolute path.</td>
</tr>
<tr>
<td>gsiftp://myhost.mydomain.com/~file1</td>
<td>File name, absolute path.</td>
</tr>
<tr>
<td>gsiftp://myhost.mydomain.com/file1</td>
<td>File name, relative path.</td>
</tr>
</tbody>
</table>
**Command line options**

- **-help | -usage**  
  Print help text.

- **-version**  
  Print the version of this program.

- **-d | -debug | -v | -verbose**  
  Print additional detail.

- **-g | -globus-endpoints**  
  Output endpoints, in place of the host portion of source and destination URLs.

- **-r | -recursive-dir-copy**  
  Output directory names, when an entire directory is to be copied recursively.

- **-n | -newer**  
  File is to be transferred, if the source timestamp is newer than the destination timestamp.

- **-o | -older**  
  File is to be transferred, if the source timestamp is older than the destination timestamp.

- **-s | -size**  
  File is to be transferred, if the sizes of the source and the destination are not the same.

**Limitations**

- This is an early version of **globus-url-sync**. In the event that unexpected results are returned, please re-run the command with the **-verbose** option.

- **globus-url-copy** should be invoked with the **-r** (copy files in subdirectories) **-cd** (create directory) options, so that directories can be copied recursively (for "globus-url-sync -r"), and so that directories at the destination can be created.

- Authentication errors may be erroneously be reported as though a file is missing.

- Order of options does not currently effect order in which matching criteria are evaluated.
Name
globus-gridftp-server -- Configures the GridFTP Server
globus-gridftp-server

Tool description
globus-gridftp-server configures the GridFTP server using a config file and/or commandline options.

Note
Command line options and configuration file options may both be used, but the command line overrides the config file.

The configuration file for the GridFTP server is read from the following locations, in the given order. Only the first file found will be loaded:

- Path specified with the -c <configfile> command line option.
- $GLOBUS_LOCATION/etc/gridftp.conf
- /etc/grid-security/gridftp.conf

Options are one per line, with the format:

<option> <value>

If the value contains spaces, they should be enclosed in double-quotes ("). Flags or boolean options should only have a value of 0 or 1. Blank lines and lines beginning with # are ignored.

For example:

    port 5000
    allow_anonymous 1
    anonymous_user bob
    banner "Welcome!"

Developer notes

The Globus implementation of the GridFTP server draws on:

- three IETF RFCs:
  - RFC 959
  - RFC 2228
  - RFC 2389
- an IETF Draft: MLST-16
- the GridFTP protocol specification, which is Global Grid Forum (GGF) Standard GFD.020.
The command line tools and the client library completely hide the details of the protocol from the user and the developer. Unless you choose to use the control library, it is not necessary to have a detailed knowledge of the protocol.

**Command syntax**

The basic syntax for `globus-gridftp-server` is:

```
globus-gridftp-server [optional command line switches]
```

To use `globus-gridftp-server` with a config file, make sure to use the `-c <configfile>` option.

**Command line options**

The table below lists config file options, associated command line options (if available) and descriptions.

> **Note**
> Any boolean option can be negated on the command line by preceding the specified option with `-no-` or `-n`. example: `-no-cas` or `-nf`.

### Informational Options

- **help** <0|1>, `-h`, `-help`  
  Show usage information and exit.
  Default value: FALSE

- **version** <0|1>, `-v`, `-version`  
  Show version information for the server and exit.
  Default value: FALSE

- **versions** <0|1>, `-V`, `-versions`  
  Show version information for all loaded globus libraries and exit.
  Default value: FALSE

### Modes of Operation

- **inetd** <0|1>, `-i`, `-inetd`  
  Run under an inetd service.
  Default value: FALSE

- **daemon** <0|1>, `-s`, `-daemon`  
  Run as a daemon. All connections will fork off a new process and setuid if allowed. See Section 4.4.1, "Running in daemon mode" for more information.
  Default value: TRUE

- **detach** <0|1>, `-S`, `-detach`  
  Run as a background daemon detached from any controlling terminals. See Section 4.4.1, "Running in daemon mode" for more information.
  Default value: FALSE

- **ssh** , `-ssh`  
  Run over a connected ssh session.
  Default value: not set
**globus-gridftp-server**

exec <string>, -exec <string>
For statically compiled or non-GLOBUS_LOCATION standard binary locations, specify the full path of the server binary here. Only needed when run in daemon mode.
Default value: not set

chdir <0|1>, -chdir
Change directory when the server starts. This will change directory to the dir specified by the chdir_to option.
Default value: TRUE

chdir_to <string>, -chdir-to <string>
Directory to chdir to after starting. Will use / if not set.
Default value: not set

fork <0|1>, -f, -fork
Server will fork for each new connection. Disabling this option is only recommended when debugging. Note that non-forked servers running as 'root' will only accept a single connection and then exit.
Default value: TRUE

single <0|1>, -1, -single
Exit after a single connection.
Default value: FALSE

chroot_path <string>, -chroot-path <string>
Path to become the new root after authentication. This path must contain a valid certificate structure, /etc/passwd, and /etc/groups. The command globus-gridftp-server-setup-chroot can help create a suitable directory structure.
Default value: not set

**Authentication, Authorization, and Security Options**

auth_level <number>, -auth-level <number>
• 0 = Disables all authorization checks.
• 1 = Authorize identity only.
• 2 = Authorize all file/resource accesses.

If not set, the GridFTP Server uses level 2 for front ends and level 1 for data nodes.
Default value: not set

ipc_allow_from <string>, -ipc-allow-from <string>
Only allow IPC connections (applicable for backend servers in a striped configuration) from these source IP addresses. Specify a comma-separated list of IP address fragments. A match is any IP address that starts with the specified fragment. Example: ‘192.168.1.’ will match and allow a connection from 192.168.1.45. Note that if this option is used, any address not specifically allowed will be denied.
Default value: not set

ipc_deny_from <string>, -ipc-deny-from <string>
Deny IPC connections (applicable for backend servers in a striped configuration) from these source IP addresses. Specify a comma-separated list of IP address fragments. A match is any IP address that starts with the specified fragment. Example: ‘192.168.2.’ will match and deny a connection from 192.168.2.45.
allow_from <string>, -allow-from <string>
- Default value: not set
- Only allow connections from these source IP addresses. Specify a comma-separated list of IP address fragments. A match is any IP address that starts with the specified fragment. Example: '192.168.1.' will match and allow a connection from 192.168.1.45. Note that if this option is used, any address not specifically allowed will be denied.

deny_from <string>, -deny-from <string>
- Default value: not set
- Deny connections from these source IP addresses. Specify a comma-separated list of IP address fragments. A match is any IP address that starts with the specified fragment. Example: '192.168.2.' will match and deny a connection from 192.168.2.45.

secure_ipc <0|1>, -si, -secure-ipc
- Default value: TRUE
- Use GSI security on the IPC channel.

ipc_auth_mode <string>, -ia, -ipc-auth-mode <string>
- Default value: host
- Set GSI authorization mode for the IPC connection. Options are one of the following:
  - none
  - host
  - self
  - subject:[subject]

allow_anonymous <0|1>, -aa, -allow-anonymous
- Default value: FALSE
- Allow cleartext anonymous access. If server is running as root, anonymous_user must also be set. Disables IPC security.

anonymous_names_allowed <string>, -anonymous-names-allowed <string>
- Default value: not set
- Comma-separated list of names to treat as anonymous users when allowing anonymous access. If not set, the default names of 'anonymous' and 'ftp' will be allowed. Use '*' to allow any username.

anonymous_user <string>, -anonymous-user <string>
- Default value: not set
- User to setuid to for an anonymous connection. Only applies when running as root.

anonymous_group <string>, -anonymous-group <string>
- Default value: not set
- Group to setgid to for an anonymous connection. If not set, the default group of anonymous_user will be used.
<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pw_file, <code>-password-file &lt;string&gt;</code></td>
<td>Enable cleartext access and authenticate users against this <code>/etc/passwd</code> formatted file. Default value: not set</td>
</tr>
<tr>
<td>connections_max, <code>-connections-max &lt;number&gt;</code></td>
<td>Maximum concurrent connections allowed. Only applies when running in daemon mode. Unlimited if not set. Default value: not set</td>
</tr>
<tr>
<td>connections_disabled, `-connections-disabled &lt;0</td>
<td>1&gt;`</td>
</tr>
<tr>
<td>offline_msg, <code>-offline-msg &lt;string&gt;</code></td>
<td>Custom message to be displayed to clients when the server is offline via the connections_disabled or connections_max = 0 options. Default value: not set</td>
</tr>
<tr>
<td>disable_command_list, <code>-disable-command-list &lt;string&gt;</code></td>
<td>A comma separated list of client commands that will be disabled. Default value: not set</td>
</tr>
<tr>
<td>authz_cal-louts, <code>-cas, -authz-cal-louts</code></td>
<td>Enable the GSI authorization callout framework, for callouts such as CAS. Default value: TRUE</td>
</tr>
<tr>
<td>acl, <code>-em, -acl</code></td>
<td>A comma seperated list of ACL or event modules to load. Default value: not set</td>
</tr>
</tbody>
</table>

**Logging Options**

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>log_level, <code>-d &lt;string&gt;, -log-level &lt;string&gt;</code></td>
<td>Log level. A comma-separated list of levels from the following:</td>
</tr>
<tr>
<td></td>
<td>- ERROR</td>
</tr>
<tr>
<td></td>
<td>- WARN</td>
</tr>
<tr>
<td></td>
<td>- INFO</td>
</tr>
<tr>
<td></td>
<td>- DUMP</td>
</tr>
<tr>
<td></td>
<td>- ALL</td>
</tr>
<tr>
<td></td>
<td>For example:</td>
</tr>
<tr>
<td></td>
<td><code>globus-gridftp-server -d error,warn,info</code></td>
</tr>
<tr>
<td></td>
<td>You may also specify a numeric level of 1-255.</td>
</tr>
</tbody>
</table>
Default value: ERROR

**log_module**

Indicates the globus_logging module that will be loaded. If not set, the default stdio module will be used and the logfile options (see next option) will apply.

Built-in modules are stdio and syslog. Log module options may be set by specifying module:opt1=val1:opt2=val2. Available options for the built-in modules are:

- **interval** - Indicates buffer flush interval. Default is 5 seconds. A 0 second flush interval will disable periodic flushing, and the buffer will only flush when it is full.

- **buffer** - Indicates buffer size. Default is 64k. A value of 0k will disable buffering and all messages will be written immediately.

Example:

```-log-module stdio:buffer=4096:interval=10```

Default value: not set

**log_single**

Indicates the path of a single file to which you want to log all activity. If neither this option nor log_unique is set, logs will be written to stderr, unless the execution mode is detached, or inetd, in which case logging will be disabled.

Note

You have to provide full path

Default value: not set

**log_unique**

Partial path to which gridftp.(pid).log will be appended to construct the log filename. Example:

```-L /var/log/gridftp/```

will create a separate log (/var/log/gridftp/gridftp.xxxx.log) for each process (which is normally each new client session). If neither this option nor log_single is set, logs will be written to stderr, unless the execution mode is detached, or inetd, in which case logging will be disabled.

Note

You have to provide full path

Default value: not set

**log_transfer**

Log NetLogger-style info for each transfer into this file.

Note

You have to provide full path

Default value: not set

Example: DATE=20050520163008.306532 HOST=localhost PROG=globus-gridftp-server NL.EVNT=FTP_INFO START=20050520163008.305913 USER=ftp FILE=/etc/group BUF-
FER=0 BLOCK=262144 NBYTES=542 VOLUME=/ STREAMS=1 STRIPES=1
DEST=[127.0.0.1] TYPE=RETR CODE=226

Time format is YYYYMMDDHHMMSS.UUUUUU (microsecs).

- DATE: time the transfer completed.
- START: time the transfer started.
- HOST: hostname of the server.
- USER: username on the host that transferred the file.
- BUFFER: tcp buffer size (if 0 system defaults were used).
- BLOCK: the size of the data block read from the disk and posted to the network.
- NBYTES: the total number of bytes transferred.
- VOLUME: the disk partition where the transfer file is stored.
- STREAMS: the number of parallel TCP streams used in the transfer.
- STRIPES: the number of stripes used on this end of the transfer.
- DEST: the destination host.
- TYPE: the transfer type, RETR is a send and STOR is a receive (ftp 959 commands).
- CODE: the FTP rfc959 completion code of the transfer. 226 indicates success, 5xx or 4xx are failure codes.

log_filemode <string>, -log-filemode <string>

File access permissions of log files. Should be an octal number such as 0644 (the leading 0 is required).

Default value: not set

disable_usage_stats <0|1>, -disable-usage-stats

Disable transmission of per-transfer usage statistics. See the Usage Statistics\(^1\) section in the online documentation for more information.

Default value: FALSE

usage_stats_target <string>, -usage-stats-target <string>

Comma-separated list of contact strings for usage statistics listeners. The format of <string> is host:port.

Default value: usage-stats.globus.org:4810

Example:


In this example, the usage statistics will be transmitted to the default Globus target (usage-stats.globus.org:4810) and another target (usage-stats.uc.teragrid.org:5920).

\(^1\)../../../Usage_Stats.html
The usage stats sent to a particular receiver may be customized by configuring it with a taglist (host:port!taglist) The taglist is a list of characters that each correspond to a usage stats tag. When this option is unset, stats are reported to usage-stats.globus.org:4810. If you set your own receiver, and wish to continue reporting to the Globus receiver, you will need to add it manually. The list of available tags follow. Tags marked * are reported by default.

- *(e) START - start time of transfer
- *(E) END - end time of transfer
- *(v) VER - version string of gridftp server
- *(b) BUFFER - tcp buffer size used for transfer
- *(B) BLOCK - disk blocksize used for transfer
- *(N) NBYTES - number of bytes transferred
- *(s) STREAMS - number of parallel streams used
- *(S) STRIPES - number of stripes used
- *(t) TYPE - transfer command: RETR, STOR, LIST, etc
- *(c) CODE - ftp result code (226 = success, 5xx = fail)
- *(D) DSI - DSI module in use
- *(A) EM - event modules in use
- *(T) SCHEME - ftp, gsiftp, sshftp, etc. (client supplied)
- *(a) APP - guc, rft, generic library app, etc. (client supplied)
- *(V) APPVER - version string of above. (client supplied)
- *(f) FILE - name of file/data transferred
- *(i) CLIENTIP - ip address of host running client (control channel)
- *(I) DATAIP - ip address of source/dest host of data (data channel)
- *(u) USER - local user name the transfer was performed as
- *(d) USERDN - DN that was mapped to user id
- *(C) CONFID - ID defined by -usage-stats-id config option
- *(U) SESSID - unique id that can be used to match transfers in a session and transfers across source/dest of a third party transfer. (client supplied)

Usage stats id
Identifying tag to include in usage statistics data.
Default value: not set
Single and Striped Remote Data Node Options

remote_nodes <string>, -r <string>, -remote-nodes <string>

Comma-separated list of remote node contact strings. See Remote data-nodes and striped operations and Separation of processes for higher security for examples of using this option.

Default value: not set

data_node <0|1>, -dn, -data-node

This server is a back end data node. See Separation of processes for higher security for an example of using this option.

Default value: FALSE

stripe_blocksize <number>, -sbs <number>, -stripe-blocksize <number>

Size in bytes of sequential data that each stripe will transfer.

Default value: 1048576

stripe_count <number>, -stripe-count <number>

Number of stripes to use per transfer when this server controls that number. If remote nodes are statically configured (via -r or remote_nodes), this will be set to that number of nodes, otherwise the default is 1.

Default value: not set

stripe_layout <number>, -sl <number>, -stripe-layout <number>

Stripe layout. 1 = Partitioned, 2 = Blocked.

Default value: 2

stripe_blocksize_locked <0|1>, -stripe-blocksize-locked; stripe_layout_locked <0|1>, -stripe-layout-locked

Do not allow client to override stripe blocksize with the OPTS RETR command.

Default value: FALSE

Do not allow client to override stripe layout with the OPTS RETR command.

Default value: FALSE

Disk Options

blocksize <number>, -bs <number>, -blocksize <number>

Size in bytes of data blocks to read from disk before posting to the network.

Default value: 262144

sync_writes <0|1>, -sync-writes

Flush disk writes before sending a restart marker. This attempts to ensure that the range specified in the restart marker has actually been committed to disk. This option will probably impact per-
formance and may result in different behavior on different storage systems. See the man page for `sync()` for more information.

Default value: FALSE

**use_home_dirs**, `-use-home-dirs`  
Set the startup directory to the authenticated users home dir.  
Default value: TRUE

**perms**, `<string>`, `-perms`  
Set the default permissions for created files. Should be an octal number such as 0644. The default is 0644. Note: If umask is set it will affect this setting -- i.e. if the umask is 0002 and this setting is 0666, the resulting files will be created with permissions of 0664.  
Default value: not set

**file_timeout**, `<number>`, `-file-timeout`  
Timeout in seconds for all disk accesses. A value of 0 disables the timeout.  
Default value: not set

---

**Network Options**

**port**, `<number>`, `-p <number>`, `-port <number>`  
Port on which a front end will listen for client control channel connections or on which a data node will listen for connections from a front end. If not set, a random port will be chosen and printed via the logging mechanism. See Remote data-nodes and striped operations and Separation of processes for higher security for examples of using this option.  
Default value: not set

**control_interface**, `<string>`, `-control-interface`  
Hostname or IP address of the interface to listen for control connections on. If not set, will listen on all interfaces.  
Default value: not set

**data_interface**, `<string>`, `-data-interface`  
Hostname or IP address of the interface to use for data connections. If not set will use the current control interface.  
Default value: not set

**ipc_interface**, `<string>`, `-ipc-interface`  
Hostname or IP address of the interface to use for IPC connections. If not set, will listen on all interfaces.  
Default value: not set

**hostname**, `<string>`, `-hostname`  
Effectively sets the above control_interface, data_interface and ipc_interface options.  
Default value: not set

**ipc_port**, `<number>`, `-ipc-port`  
Port on which the front end will listen for data node connections.  
Default value: not set
Timeouts

control_preauth_timeout <number>,
- control-preauth-timeout <number>

Time in seconds to allow a client to remain connected to the control channel without activity before authenticating.

Default value: 120

control_idle_timeout <number>,
- control-idle-timeout <number>

Time in seconds to allow a client to remain connected to the control channel without activity.

Default value: 600

ipc_idle_timeout <number>,
- ipc-idle-timeout <number>

Idle time in seconds before an unused IPC connection will close.

Default value: 600

ipc_connect_timeout <number>,
- ipc-connect-timeout <number>

Time in seconds before cancelling an attempted IPC connection.

Default value: 60

User Messages

banner <string>,
- banner <string>

Message that is displayed to the client before authentication.

Default value: not set

banner_file <string>,
- banner-file <string>

Read banner message from this file.

Default value: not set

banner_terse <0|1>, -banner-terse

When this is set, the minimum allowed banner message will be displayed to unauthenticated clients.

Default value: FALSE

banner_append <0|1>, -banner-append

When this is set, the message set in the 'banner' or 'banner_file' option will be appended to the default banner message rather than replacing it.

Default value: FALSE

login_msg <string>, -login-msg <string>

Message that is displayed to the client after authentication.

Default value: not set
lo-
gin_msg_file
<string>, -lo-
gin-msg-file
<string>
Read login message from this file.
Default value: not set

Module Options

load_dsi_mod-
ule <string>,
-dsi <string>
Load this Data Storage Interface module. File and remote modules are defined by the server. If
not set, the file module is loaded, unless the remote option is specified, in which case the remote
module is loaded. An additional configuration string can be passed to the DSI using the format
[module name]:[configuration string]. The format of the configuration string is
defined by the DSI being loaded.
Default value: not set

allowed_mod-
ules <string>,
-allowed-
modules
<string>
Comma-separated list of ERET/ESTO modules to allow and, optionally, specify an alias for.
Example:
-allowed-modules module1,alias2:module2,module3
(module2 will be loaded when a client asks for alias2).
Default value: not set

dc_whitelist
<string>,
-dc-whitelist
<string>
A comma separated list of drivers allowed on the network stack.
Default value: not set

fs_whitelist
<string>,
-fs-whitelist
<string>
A comma separated list of drivers allowed on the disk stack.
Default value: not set

popen_whitel-
ist <string>,
-popen-whitel-
ist <string>
A comma separated list of programs that the popen driver is allowed to execute, when used on
the network or disk stack. An alias may also be specified, so that a client does not need to specify
the full path. Format is [alias:][prog],[alias:][prog]. example: /bin/gzip.tar:/bin/tar
Default value: not set

Other Options

configfile
<string>, -c
<string>
Path to configuration file that should be loaded. Otherwise will attempt to load $GLOBUS_LOC-
ATION/etc/gridftp.conf and /etc/grid-security/gridftp.conf.

Note
You have to provide full path
Default value: not set

debug <0|1>,
-debug
Set options that make the server easier to debug. Forces no-fork, no-chdir, and allows core dumps
on bad signals instead of exiting cleanly. Not recommended for production servers. Note that
non-forked servers running as root will only accept a single connection and then exit.
Default value: FALSE

**Warning**

Any FLAG can be negated by prepending `-no` or `-n` to the command line option.

**Limitations**

For transfers using parallel data transport streams and for transfers using multiple computers at each end, the direction of the connection on the data channels must go from the sending to the receiving side. For more information about this limitations see [http://www.ogf.org/documents/GFD.20.pdf](http://www.ogf.org/documents/GFD.20.pdf).

Globus GridFTP server does not run on windows
Replica Location Service (RLS) Commands
Name

globus-rls-admin -- RLS administration tool
globus-rls-admin

Tool description

Performs administrative operations on an RLS server.

Synopsis

-A|--a|--C option value|--c option|--d|--e|--p|--q|s|--t timeout|--u|--v
   [ rli ] [ pattern ] [ server ]
# Options

## Table 3. Options for globus-rls-admin

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
</table>
| `-A`   | Adds `rli` to the list of RLI servers updated by an LRC server using Bloom filters.  

*Note:* Partitions are not supported with Bloom filters. The LRC server maintains one Bloom filter for all LFNs in its database, which is sent to all RLI servers configured to receive Bloom filter updates with this option. |
| `-a`   | Adds `rli` and optionally `pattern` to the list of RLI servers that the LRC server sends updates to (using a list of LFNs).  

If `pattern` is specified, then only LFNs matching it will be sent to `rli`.  

If `rli` is added with no patterns, then it is sent all updates. Pattern matching is done using standard Unix file globbing. |
| `-C option value` | Sets server `option` to `value`.  

*Important:* This does not update the configuration file. The next time the server is restarted, the configuration change will be lost. |
| `-c option` | Retrieves the configuration value for the specified option from the server.  

If `option` is set to `all`, then all options are retrieved. |
| `-d`   | Removes `rli` and `pattern` from the list of RLI servers that the LRC server sends updates to.  

If `pattern` is not specified, then all entries for `rli` are removed.  

*Note:* If all patterns are removed separately, then `rli` is sent all updates. To stop any updates from being sent to `rli`, do not specify `pattern`. |
| `-e`   | Clears the LRC database. Removes all lfn, *pfn* mappings. |
| `-p`   | Verifies that the server is responding. |
| `-q`   | Causes the RLS server to exit. |
| `-S`   | Shows statistics and other information gathered by the RLS server.  

This is intended to be input into GRIS. |
| `-s`   | Shows the list of RLI servers and patterns being sent updates by the LRC server.  

If `rli` or `pattern` are not specified, they are considered wildcards. |
| `-t timeout` | Sets timeout (in seconds) for RLS server requests.  

The default value is **30**. |
| `-u`   | Causes the LRC server to immediately start full soft state updates to any RLI servers previously added with the `-a` option. |
| `-v`   | Shows the version and exits. |
Name
globus-rls-cli -- RLS client tool
globus-rls-cli

Tool description
Provides a command line interface to some of the functions supported by RLS. It also supports an interactive interface (if command is not specified). In interactive mode, double quotes may be used to encode an argument that contains white space.

Synopsis
command [ -c ] [ -h ] [ -l reslimit ] [ -s ] [ -t timeout ] [ -u ] [ command ] rls-server

Options
The client command tool uses getopt for command line parsing.

Note: Some versions will continue scanning for options (works that begin with a hyphen) for the entire command line, which makes it impossible to specify negative integer or floating point value for an attribute. The workaround for this problem is to tell getopt() that there are no more options by including 2 hyphens. For example, to specify the value -2 you must enter -- -2.

Table 4. Options for globus-rls-cli

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-c</td>
<td>Sets &quot;clearvalues&quot; flag when deleting an attribute (will remove any attribute value records when an attribute is deleted).</td>
</tr>
<tr>
<td>-h</td>
<td>Shows usage.</td>
</tr>
<tr>
<td>-l reslimit</td>
<td>Sets an incremental limit on the number of results returned by a wildcard query at a time. Note that all results will be returned by the client. This parameter only limits the number of results incrementally retrieved by the client during a single internal communication call. For instance, if the wildcard query produces 1000 results and the reslimit is set to 100, the client will internally make 10 calls to the server. From the user's perspective the client will simply return all 1000 results. Zero means no limit.</td>
</tr>
<tr>
<td>-s</td>
<td>Uses SQL style wildcards (% and _).</td>
</tr>
<tr>
<td>-t timeout</td>
<td>Sets timeout (in seconds) for RLS server requests. The default is 30 seconds.</td>
</tr>
<tr>
<td>-u</td>
<td>Uses Unix style wildcards (* and ?).</td>
</tr>
<tr>
<td>-v</td>
<td>Shows version.</td>
</tr>
</tbody>
</table>
Commands
Table 5. Commands for globus-rls-cli

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>add &lt;lfn&gt; &lt;pfn&gt;</td>
<td>Adds pfn to mappings of lfn in an LRC catalog.</td>
</tr>
<tr>
<td>attribute add &lt;object&gt; &lt;attr&gt; &lt;obj-type&gt; &lt;attr-type&gt;</td>
<td>Adds an attribute to an object, where object should be the lfn or pfn name. obj-type should be one of lfn or pfn. attr-type should be one of date, float, int, or string. If &lt;value&gt; is of type date then it should be in the form &quot;YYYY-MM-DD HH:MM:DD&quot;.</td>
</tr>
<tr>
<td>attribute bulk add &lt;object&gt; &lt;attr&gt; &lt;obj-type&gt;</td>
<td>Bulk adds attribute values.</td>
</tr>
<tr>
<td>attribute bulk delete &lt;object&gt; &lt;attr&gt; &lt;obj-type&gt;</td>
<td>Bulk deletes attributes.</td>
</tr>
<tr>
<td>attribute bulk query &lt;attr&gt; &lt;obj-type&gt;</td>
<td>Bulk queries attributes.</td>
</tr>
<tr>
<td>attribute define &lt;attr&gt; &lt;obj-type&gt; &lt;attr-type&gt;</td>
<td>Defines a new attribute.</td>
</tr>
<tr>
<td>attribute delete &lt;object&gt; &lt;attr&gt; &lt;obj-type&gt;</td>
<td>Removes attribute from object.</td>
</tr>
<tr>
<td>attribute modify &lt;object&gt; &lt;attr&gt; &lt;obj-type&gt; &lt;attr-type&gt;</td>
<td>Modifies the value of an attribute.</td>
</tr>
<tr>
<td>attribute query &lt;object&gt; &lt;attr&gt; &lt;obj-type&gt;</td>
<td>Retrieves the value of the specified attribute for object.</td>
</tr>
<tr>
<td>attribute search &lt;attr&gt; &lt;obj-type&gt; &lt;operator&gt; &lt;attr-type&gt;</td>
<td>Searches for objects which have the specified attribute matching operator and value. operator should be one of =, !=, &gt;, &gt;=, &lt;, or &lt;=.</td>
</tr>
<tr>
<td>attribute show &lt;attr&gt; &lt;obj-type&gt;</td>
<td>Shows an attribute definition. If attr is a hyphen (-) then all attributes are shown.</td>
</tr>
<tr>
<td>attribute undefine &lt;attr&gt; &lt;obj-type&gt;</td>
<td>Deletes an attribute definition. Will return an error if any objects possess this attribute.</td>
</tr>
<tr>
<td>bulk add &lt;lfn&gt; &lt;pfn&gt; [ &lt;lfn&gt; &lt;pfn&gt; ]</td>
<td>Bulk adds lfn, pfn mappings.</td>
</tr>
<tr>
<td>bulk create &lt;lfn&gt; &lt;pfn&gt; [ &lt;lfn&gt; &lt;pfn&gt; ]</td>
<td>Bulk creates lfn, pfn mappings.</td>
</tr>
<tr>
<td>bulk delete &lt;lfn&gt; &lt;pfn&gt; [ &lt;lfn&gt; &lt;pfn&gt; ]</td>
<td>Bulk deletes lfn, pfn mappings.</td>
</tr>
<tr>
<td>bulk query lrc lfn [ &lt;lfn&gt; ... ]</td>
<td>Bulk queries the LRC for lfns.</td>
</tr>
<tr>
<td>bulk query lrc pfn [ &lt;pfn&gt; ... ]</td>
<td>Bulk queries the LRC for pfns.</td>
</tr>
<tr>
<td>bulk query rli lfn [ &lt;lfn&gt; ... ]</td>
<td>Bulk queries the RLI for lfns.</td>
</tr>
<tr>
<td>create &lt;lfn&gt; &lt;pfn&gt;</td>
<td>Creates a new lfn, pfn mapping in an LRC catalog.</td>
</tr>
<tr>
<td>delete &lt;lfn&gt; &lt;pfn&gt;</td>
<td>Deletes a lfn, pfn mapping from an LRC catalog.</td>
</tr>
<tr>
<td>exit</td>
<td>Exits the interactive session.</td>
</tr>
<tr>
<td>help</td>
<td>Prints a help message.</td>
</tr>
<tr>
<td>query lrc lfn &lt;lfn&gt;</td>
<td>Queries an LRC server for mappings of lfn.</td>
</tr>
<tr>
<td>query lrc pfn &lt;pfn&gt;</td>
<td>Queries an LRC server for mappings to pfn.</td>
</tr>
<tr>
<td>query rli lfn &lt;lfn&gt;</td>
<td>Queries an RLI server for mappings of lfn.</td>
</tr>
<tr>
<td>Command</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>query wildcard lrc lfn &lt;lfn-pattern&gt;</strong></td>
<td>Performs a wildcarded query of an LRC server for mappings of <code>lfn-pattern</code>. Patterns use the standard Unix wildcard characters: an asterisk (*) matches 0 or more characters, and a question mark (?) matches any single character.</td>
</tr>
<tr>
<td><strong>query wildcard lrc pfn &lt;pfn-pattern&gt;</strong></td>
<td>Queries an LRC server for mappings to <code>pfn-pattern</code>. Patterns use the standard Unix wildcard characters: an asterisk (*) matches 0 or more characters, and a question mark (?) matches any single character.</td>
</tr>
<tr>
<td><strong>query wildcard rli lfn &lt;lfn-pattern&gt;</strong></td>
<td>Queries an RLI server for mappings of <code>lfn-pattern</code>. Patterns use the standard Unix wildcard characters: an asterisk (*) matches 0 or more characters, and a question mark (?) matches any single character.</td>
</tr>
<tr>
<td><strong>set reslimit &lt;limit&gt;</strong></td>
<td>Sets an incremental limit on the number of results returned by a wildcard query at a time. Note that <em>all results will be returned</em> by the client. This parameter only limits the number of results <em>incrementally retrieved</em> by the client during a single internal communication call. For instance, if the wildcard query produces 1000 results and the reslimit is set to 100, the client will internally make 10 calls to the server. From the user’s perspective the client will simply return all 1000 results.</td>
</tr>
<tr>
<td><strong>set timeout &lt;timeout&gt;</strong></td>
<td>Sets the timeout (in seconds) on calls to the RLS server. The default value is <strong>30</strong>.</td>
</tr>
<tr>
<td><strong>version</strong></td>
<td>Shows the version and exits.</td>
</tr>
</tbody>
</table>
Name
globus-rls-server -- RLS server tool
globus-rls-server

Tool description
The RLS server (globus-rls-server) can be configured as either one or both of the following:

• Location Replica Catalog (LRC) server, which manages Logical FileName (LFN) to Physical FileName (PFN) mappings in a database. Note: If globus-rls-server is configured as an LRC server, the RLI servers that it sends updates to should be added to the database using globus-rls-admin.

• Replica Location Index (RLI) server, which manages mappings of LFNs to LRC servers.

Clients wishing to locate one or more physical filenames associated with a logical filename should first contact an RLI server, which will return a list of LRCs that may know about the LFN. The LRC servers are then contacted in turn to find the physical filenames.

Note: RLI information may be out of date, so clients should be prepared to get a negative response when contacting an LRC (or no response at all if the LRC server is unavailable).

Synopsis


LRC to RLI Updates

Two methods exist for LRC servers to inform RLI servers of their LFNs.

• By default, the LFNs are sent from the LRC to the RLI. This can be time consuming if the number of LFNs is large, but it does give the RLI an exact list of the LFNs known to the LRC, and it allows wildcard searching of the RLI.

• Alternatively, Bloom filters may be sent, which are highly compressed summaries of the LFNs. However, they do not allow wildcard searching and will generate more "false positives" when querying an RLI.

Please see below for more on Bloom filters.

globus-rls-admin can be used to manage the list of RLIs that an LRC server updates. This includes partitioning LFNs among multiple RLI servers.

A softstate algorithm is used for updates, periodically the source server sends its state (LFN information) to the RLI servers it updates. The RLI servers add these LFNs to their index, or update a timestamp if the LFNs were already known. RLI servers expire information about LFN,LRC mappings if they haven’t been updated for a period longer than the softstate update interval.

Options that can be configured to control the softstate algorithm when a source server updates an RLI by sending LFNs include:
• **rli_expire_int** (seconds)

  How often an RLI server will check for stale entries in its database.

• **rli_expire_stale** (seconds)

  How old an entry must be in an RLI database before it's considered stale. This value should be no smaller than `update_ll_int`. Note if the LRC server is responding this value is not used, instead the value of `update_ll_int` or `update_bf_int` is retrieved from the LRC server, multiplied by 1.2, and used as the value for `rli_expire_stale`.

• **update_bf_int** (seconds)

  Interval between RLI updates when using Bloom filters.

• **update_ll_int** (seconds)

  Interval between RLI updates when using LFN lists for softstate updates.

Updates to an LRC (new LFNs or deleted LFNs) normally don't propagate to RLI servers until the next softstate update (controlled by `update_ll_int` and `update_bf_int`). However by enabling "immediate update" mode an LRC will send updates to an RLI within `update_buftime` seconds. Immediate updates are enabled by setting `update_immediate` to true. If updates are done with LFN lists then only the LFNs that have been added or deleted to the source server are sent, if Bloom filters are used then the entire Bloom filter is sent.

When immediate updates are enabled, the interval between softstate updates is multiplied by `update_factor`, so long as no updates have failed (source and RLI are considered to be in sync). This can greatly reduce the number of softstate updates a source needs to send to an RLI. Incremental updates are buffered by the source server until either 100 updates have accumulated (when LFN lists are used), or `update_buftime` seconds have passed since the last update.

## Bloom filter updates

A Bloom filter is an array of bits. Each LFN is hashed multiple times and the corresponding bits in the Bloom filter are set.

Querying an RLI to verify if an LFN exists is done by performing the same hashes and checking if the bits in the filter are on. If not, then the LFN is known not to exist. If they're all on, then all that's known is that the LFN probably exists.

The size of the Bloom filter (as a multiple of the number of LFNs) and the number of hash functions control the false positive rate. The default values of 10 and 3 give a false positive rate of approximately 1%.

The advantage of Bloom filters is their efficiency. For example, if the LRC has 1,000,000 LFNs in its database, with an average length of 20 bytes, then 20,000,000 bytes must be sent to an RLI during a soft state update (assuming no partitioning). The RLI server must perform 1,000,000 updates to its database to create new LFN, LRC mappings or update timestamps on existing entries. With Bloom filters only 1,250,000 bytes are sent (10 x 1,000,000 bits / 8), and there are no database operations on the RLI (Bloom filters are maintained entirely in memory). A comparison of the time to perform a 1,000,000 LFN update: it took 20 minutes sending all the LFNs and less than 1 second using a Bloom filter. However as noted before, Bloom filters do not support wild card searches of an RLI.

*Note:* An LRC server can update some RLIs with Bloom filters and others with LFNs. However, an RLI server can only be updated using one method.

The following options in the `Configuration` file control Bloom filter updates:

• **rli_bloomfilter true|false**

  RLI servers must have this set to accept Bloom filter updates.
• **rli_bloomfilter_dir none|default|pathname**
  
  Bloom filters saved in this directory and read at start time if not "none". See CONFIGURATION for details.

• **lrc_bloomfilter_numhash N**
  
  Number of hash functions, an integer from 1 to 8. The default is 3.

• **lrc_bloomfilter_ratio N**
  
  Size of the Bloom filter as a multiple of the number of LFNs in the LRC database. Too small a value will generate too many false positives, too large wastes memory and network bandwidth.

Note: An LRC server can update some RLIs with Bloom filters, and others with LFNs. However an RLI server can only be updated using one method, and an RLI acting as a source for updates can only send the type of updates that it receives.

**Log Messages**

*globus-rls-server* uses syslog to log errors and other information (facility LOG_DAEMON) when it's running in normal (daemon) mode.

If the `-d` option (debug) is specified, then log messages are written to stdout.

**Signals**

The server will reread its configuration file if it receives a HUP signal. It will wait for all current requests to complete and shut down cleanly if sent any of the following signals: INT, QUIT or TERM.

**Options (globus-rls-server)**

The following table describes the command line options available for globus-rls-server:
Table 6. Options for globus-rls-server

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-B update_bf_int</td>
<td>Interval between RLI updates when using Bloom filters.</td>
</tr>
<tr>
<td>-b maxbackoff</td>
<td>Maximum time (in seconds) that globus-rls-server will attempt to reopen the socket it listens on after an I/O error.</td>
</tr>
<tr>
<td>-C rlscertfile</td>
<td>Name of the X.509 certificate file that identifies the server; sets environment variable X509_USER_CERT.</td>
</tr>
<tr>
<td>-c conffile</td>
<td>Name of the configuration file for the server. The default is $GLOBUS_LOCATION/etc/globus-rls-server.conf if the environment variable GLOBUS_LOCATION is set; else, /usr/local/etc/globus-rls-server.conf.</td>
</tr>
<tr>
<td>-d</td>
<td>Enables debugging. The server will not detach from the controlling terminal, and log messages will be written to stdout rather than syslog. For additional logging verbosity set the loglevel (see the -L option) to higher values.</td>
</tr>
<tr>
<td>-e rli_expire_int</td>
<td>Interval (seconds) at which an RLI server should expire stale entries.</td>
</tr>
<tr>
<td>-F lrc_update_factor</td>
<td>If lrc_update Immediate mode is on, and the LRC server is in sync with an RLI server (an LRC and RLI are synced if there have been no failed updates since the last full soft state update), then the interval between RLI updates for this server (update_ll_int) is multiplied by lrc_update_factor.</td>
</tr>
<tr>
<td>-f maxfreethreads</td>
<td>Maximum number of idle threads the server will leave running. Excess threads are terminated.</td>
</tr>
<tr>
<td>-I true</td>
<td>false</td>
</tr>
<tr>
<td>-i idletimeout</td>
<td>Seconds after which idle client connections are timed out.</td>
</tr>
<tr>
<td>-K rlskeyfile</td>
<td>Name of the X.509 key file. Sets environment variable X509_USER_KEY.</td>
</tr>
<tr>
<td>-L loglevel</td>
<td>Sets the log level. By default this is 0, which means only errors will be logged. Higher values mean more verbose logging.</td>
</tr>
<tr>
<td>-l true</td>
<td>false</td>
</tr>
<tr>
<td>-M maxconnections</td>
<td>Maximum number of active connections. It should be small enough to prevent the server from running out of open file descriptors. The default value is 100.</td>
</tr>
<tr>
<td>-m maxthreads</td>
<td>Maximum number of threads server will start up to support simultaneous requests.</td>
</tr>
<tr>
<td>-N</td>
<td>Disables authentication checking. This option is intended for debugging. Clients should use the URL RLSN://host to disable authentication on the client side.</td>
</tr>
<tr>
<td>-o update_buftime</td>
<td>LRC to RLI updates are buffered until either the buffer is full or this much time (in seconds) has elapsed since the last update. The default value is 30.</td>
</tr>
<tr>
<td>-p pidfiledir</td>
<td>Directory where PID files should be written.</td>
</tr>
<tr>
<td>Option</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td><code>-r</code></td>
<td>Configures whether the server is an RLI server. The default value is <code>false</code>.</td>
</tr>
<tr>
<td><code>-S rli_expire_stale</code></td>
<td>Interval (in seconds) after which entries in the RLI database are considered stale (presumably because they were deleted in the LRC). Stale entries are not returned in queries.</td>
</tr>
<tr>
<td><code>-s startthreads</code></td>
<td>Number of threads to start up initially.</td>
</tr>
<tr>
<td><code>-t timeout</code></td>
<td>Timeout (in seconds) for calls to other RLS servers (in other words, for LRC calls to send an update to an RLI). A value of 0 disables timeouts. The default value is 30.</td>
</tr>
<tr>
<td><code>-U myurl</code></td>
<td>URL for this server.</td>
</tr>
<tr>
<td><code>-u update_ll_int</code></td>
<td>Interval (in seconds) between lfn-list LRC to RLI updates.</td>
</tr>
<tr>
<td><code>-v</code></td>
<td>Shows version and exits.</td>
</tr>
</tbody>
</table>
GRAM5 Commands
Name

globusrun -- Execute and manage jobs via GRAM

globusrun [-p | -parse]
{ -f RSL_FILENAME | -file RSL_FILENAME | RSL_SPECIFICATION }
globusrun [-n] [-no-interrupt]
{ -r RESOURCE_CONTACT | -resource RESOURCE_CONTACT }
{ -a | -authenticate-only }
globusrun [-n] [-no-interrupt]
{ -r RESOURCE_CONTACT | -resource RESOURCE_CONTACT }
{ -j | -jobmanager-version }
globusrun [-n] [-no-interrupt] { -k | -kill } { JOB_ID }
globusrun [-status] { JOB_ID }
{ -r RESOURCE_CONTACT | -resource RESOURCE_CONTACT }
{ -f RSL_FILENAME | -file RSL_FILENAME | RSL_SPECIFICATION }

Description

The globusrun program for submits and manages jobs run on a local or remote job host. The jobs are controlled by the globus-job-manager program which interfaces with a local resource manager that schedules and executes the job.

The globusrun program can be run in a number of different modes chosen by command-line options.

When -help, -usage, -version, or -versions command-line options are used, globusrun will print out diagnostic information and then exit.

When the -p or -parse command-line option is present, globusrun will verify the syntax of the RSL specification and then terminate. If the syntax is valid, globusrun will print out the string "RSL Parsed Successfully..." and exit with a zero exit code; otherwise, it will print an error message and terminate with a non-zero exit code.

When the -a or -authenticate-only command-line option is present, globusrun will verify that the service named by RESOURCE_CONTACT exists and the client's credentials are granted permission to access that service. If authentication is successful, globusrun will display the string "GRAM Authentication test successful" and exit with a zero exit code; otherwise, it will print an explanation of the problem and will with a non-zero exit code.

When the -j or -jobmanager-version command-line option is present, globusrun will attempt to determine the software version that the service named by RESOURCE_CONTACT is running. If successful, it will display both the Toolkit version and the Job Manager package version and exit with a zero exit code; otherwise, it will print an explanation of the problem and exit with a non-zero exit code.

When the -k or -kill command-line option is present, globusrun will attempt to terminate the job named by JOB_ID. If successful, globusrun will exit with zero; otherwise, it will display an explanation of the problem and exit with a non-zero exit code.

When the -y or -refresh-proxy command-line option is present, globusrun will attempt to delegate a new X.509 proxy to the job manager which is managing the job named by JOB_ID. If successful, globusrun will exit with zero; otherwise, it will display an explanation of the problem and exit with a non-zero exit code. This behavior can be modified by the -full-proxy or -D command-line options to enable full proxy delegation. The default is limited proxy delegation.
When the `-status` command-line option is present, `globusrun` will attempt to determine the current state of the job. If successful, the state will be printed to standard output and `globusrun` will exit with a zero exit code; otherwise, a description of the error will be displayed and it will exit with a non-zero exit code.

Otherwise, `globusrun` will submit the job to a GRAM service. By default, `globusrun` waits until the job has terminated or failed before exiting, displaying information about job state changes and at exit time, the job exit code if it is provided by the GRAM service.

The `globusrun` program can also function as a GASS file server to allow the `globus-job-manager` program to stage files to and from the machine on which `globusrun` is executed to the GRAM service node. This behavior is controlled by the `-s`, `-o`, and `-w` command-line options.

Jobs submitted by `globusrun` can be monitored interactively or detached. To have `globusrun` detach from the GRAM service after submitting the job, use the `-b` or `-F` command-line options.

**Options**

The full set of options to `globusrun` consist of:

- `-help` Display a help message to standard error and exit.
- `-usage` Display a one-line usage summary to standard error and exit.
- `-version` Display the software version of `globusrun` to standard error and exit.
- `-versions` Display the software version of all modules used by `globusrun` (including DiRT information) to standard error and then exit.
- `-p`, `-parse` Do a parse check on the job specification and print diagnostics. If a parse error occurs, `globusrun` exits with a non-zero exit code.
- `-f RSL_FILENAME`, `-file RSL_FILENAME` Read job specification from the file named by `RSL_FILENAME`.
- `-n`, `-no-interrupt` Disable handling of the SIGINT signal, so that the interrupt character (typically `Control-C`) causes `globusrun` to terminate without canceling the job.
- `-r RESOURCE_CONTACT`, `-resource RESOURCE_CONTACT` Submit the request to the resource specified by `RESOURCE_CONTACT`. A resource may be specified in the following ways:
  - `HOST`
  - `HOST:PORT`
  - `HOST:PORT/SERVICE`
  - `HOST/SERVICE`
  - `HOST/:SERVICE`
  - `HOST::SUBJECT`
  - `HOST:PORT:SUBJECT`
  - `HOST/SERVICE:SUBJECT`
If any of PORT, SERVICE, or SUBJECT is omitted, the defaults of 2811, jobmanager, and host@HOST are used respectively.

-j, -jobmanager-version
Print the software version being run by the service running at RESOURCE_CONTACT.

-k JOB_ID, -kill JOB_ID
Kill the job named by JOB_ID.

-D, -full-proxy
Delegate a full impersonation proxy to the service. By default, a limited proxy is delegated when needed.

-y, -refresh-proxy
Delegate a new proxy to the service processing JOB_ID.

-status
Display the current status of the job named by JOB_ID.

-q, -quiet
Do not display job state change or exit code information.

-o, -output-enable
Start a GASS server within the globusrun application that allows access to its standard output and standard error streams only. Also, augment the RSL_SPECIFICATION with a definition of the GLOBUSRUN_GASS_URL RSL substitution and add stdout and stderr clauses which redirect the output and error streams of the job to the output and error streams of the interactive globusrun command. If this is specified, then globusrun acts as though the -q were also specified.

-s, -server
Start a GASS server within the globusrun application that allows access to its standard output and standard error streams for writing and any file local the the globusrun invocation for reading. Also, augment the RSL_SPECIFICATION with a definition of the GLOBUSRUN_GASS_URL RSL substitution and add stdout and stderr clauses which redirect the output and error streams of the job to the output and error streams of the interactive globusrun command. If this is specified, then globusrun acts as though the -q were also specified.

-w, -write-allow
Start a GASS server within the globusrun application that allows access to its standard output and standard error streams for writing and any file local the the globusrun invocation for reading or writing. Also, augment the RSL_SPECIFICATION with a definition of the GLOBUSRUN_GASS_URL RSL substitution and add stdout and stderr clauses which redirect the output and error streams of the job to the output and error streams of the interactive globusrun command. If this is specified, then globusrun acts as though the -q were also specified.

-b, -batch
Terminate after submitting the job to the GRAM service. The globusrun program will exit after the job hits any of the following states: PENDING, ACTIVE, FAILED, or DONE. The GASS-related options can be used to stage input files, but standard output, standard error, and file staging after the job completes will not be processed.

-F, -fast-batch
Terminate after submitting the job to the GRAM service. The globusrun program will exit after it receives a reply from the service. The JOB_ID will be displayed to standard output before terminating so that the job can be checked with the

76
globusrun

- status command-line option or modified by the -refresh-proxy or
  -kill command-line options.

- d, -dryrun
  Submit the job with the dryrun attribute set to true. When this is done, the job
  manager will prepare to start the job but start short of submitting it to the service.
  This can be used to detect problems with the RSL_SPECIFICATION.

Environment

If the following variables affect the execution of globusrun

X509_USER_PROXY Path to proxy credential.

X509_CERT_DIR Path to trusted certificate directory.

Bugs

The globusrun program assumes any failure to contact the job means the job has terminated. In fact, this may be due
to the globus-job-manager program exiting after all jobs it is managing have reached the DONE or FAILED states.
In order to reliably detect job termination, the two_phase RSL attribute should be used.

See Also

globus-job-submit(1), globus-job-run(1), globus-job-clean(1), globus-job-get-output(1), globus-job-cancel(1)
Name

globus-job-cancel -- Cancel a GRAM batch job

globus-job-cancel [-f | -force] [ -q | -quiet ] JOBID


Description

The globus-job-cancel program cancels the job named by JOBID. Any cached files associated with the job will remain until globus-job-clean is executed for the job.

By default, globus-job-cancel prompts the user prior to canceling the job. This behavior can be overridden by specifying the -f or -force command-line options.

Options

The full set of options to globus-job-cancel are:

- help, -usage
  Display a help message to standard error and exit.

- version
  Display the software version of the globus-job-cancel program to standard output.

- version
  Display the software version of the globus-job-cancel program including DiRT information to standard output.

- force, -f
  Do not prompt to confirm job cancel and clean-up.

- quiet, -q
  Do not print diagnostics for succesful cancel. Implies -f

ENVIRONMENT

If the following variables affect the execution of globus-job-cancel.

X509_USER_PROXY Path to proxy credential.

X509_CERT_DIR Path to trusted certificate directory.
Name

globus-job-clean -- Cancel and clean up a GRAM batch job

globus-job-clean [-r RESOURCE] [-resource RESOURCE]


Description

The globus-job-clean program cancels the job named by JOBID if it is still running, and then removes any cached files on the GRAM service node related to that job. In order to do the file clean up, it submits a job which removes the cache files. By default this cleanup job is submitted to the default GRAM resource running on the same host as the job. This behavior can be controlled by specifying a resource manager contact string as the parameter to the -r or -resource option.

By default, globus-job-clean prompts the user prior to canceling the job. This behavior can be overridden by specifying the -f or -force command-line options.

Options

The full set of options to globus-job-clean are:

- help, -usage Display a help message to standard error and exit.
- version Display the software version of the globus-job-clean program to standard output.
- version Display the software version of the globus-job-clean program including DiRT information to standard output.

- resource RESOURCE, -r RESOURCE Submit the clean-up job to the resource named by RESOURCE instead of the default GRAM service on the same host as the job contact.

- force, -f Do not prompt to confirm job cancel and clean-up.
- quiet, -q Do not print diagnostics for succesful clean-up. Implies -f

ENVIRONMENT

If the following variables affect the execution of globus-job-clean.

X509_USER_PROXY Path to proxy credential.
X509_CERT_DIR Path to trusted certificate directory.
Name

globus-job-get-output -- Retrieve the output and error streams from a GRAM job

globus-job-get-output [-r RESOURCE | -resource RESOURCE ]
[-out | -err] [-t LINES | -tail LINES ] [-follow LINES | -f LINES ] JOBID

Description

The globus-job-get-output program retrieves the output and error streams of the job named by JOBID. By default, globus-job-get-output will retrieve all output and error data from the job and display them to its own output and error streams. Other behavior can be controlled by using command-line options. The data retrieval is implemented by submitting another job which simply displays the contents of the first job's output and error streams. By default this retrieval job is submitted to the default GRAM resource running on the same host as the job. This behavior can be controlled by specifying a particular resource manager contact string as the RESOURCE parameter to the -r or -resource option.

Options

The full set of options to globus-job-get-output are:

- -help, -usage
  Display a help message to standard error and exit.

- -version
  Display the software version of the globus-job-get-output program to standard output.

- -version
  Display the software version of the globus-job-get-output program including DiRT information to standard output.

- -resource RESOURCE,
  -r RESOURCE
  Submit the retrieval job to the resource named by RESOURCE instead of the default GRAM service on the same host as the job contact.

- -out
  Retrieve only the standard output stream of the job. The default is to retrieve both standard output and standard error.

- -err
  Retrieve only the standard error stream of the job. The default is to retrieve both standard output and standard error.

- -tail LINES, -t LINES
  Print only the last LINES count lines of output from the data streams being retrieved. By default, the entire output and error file data is retrieved. This option can not be used along with the -f or -follow options.

- -follow LINES, -f LINES
  Print the last LINES count lines of output from the data streams being retrieved and then wait until canceled, printing any subsequent job output that occurs. By default, the entire output and error file data is retrieved. This option can not be used along with the -t or -tail options.

ENVIRONMENT

If the following variables affect the execution of globus-job-get-output.

X509_USER_PROXY Path to proxy credential.
X509_CERT_DIR   Path to trusted certificate directory.
Name

globus-job-run -- Execute a job using GRAM

globus-job-run [-dumprsl] [-dryrun] [-verify]
[-file ARGUMENT_FILE]
SERVICE_CONTACT
[-np PROCESSES | -count PROCESSES ]
[-m MAX_TIME | -maxtime MAX_TIME ]
[-p PROJECT | -project PROJECT ]
[-q QUEUE | -queue QUEUE ]
[-d DIRECTORY | -directory DIRECTORY ] [-env NAME=VALUE]...
[-x RSL_CLAUSE]
[-l | -s ] EXECUTABLE [ARGUMENT...]

Description

The globus-job-run program constructs a job description from its command-line options and then submits the job to
the GRAM service running at SERVICE_CONTACT. The executable and arguments to the executable are provided
on the command-line after all other options. Note that the -dumprsl, -dryrun, -verify, and -file command-
line options must occur before the first non-option argument, the SERVICE_CONTACT.

The globus-job-run provides similar functionality to globusrun in that it allows interactive start-up of GRAM jobs.
However, unlike globusrun, it uses command-line parameters to define the job instead of RSL expressions.

Options

The full set of options to globus-job-run are:

- **help, -usage** Display a help message to standard error and exit.
- **version** Display the software version of the globus-job-run program to standard output.
- **version** Display the software version of the globus-job-run program including DiRT
information to standard output.
- **dumprsl** Translate the command-line options to globus-job-run into an RSL expression
that can be used with tools such as globusrun.
- **dryrun** Submit the job request to the GRAM service with the dryrun option enabled.
When this option is used, the GRAM service prepares to execute the job but
stops before submitting the job to the LRM. This can be used to diagnose some
problems such as missing files.
- **verify** Submit the job request to the GRAM service with the dryrun option enabled
and then without it enabled if the dryrun is successful.
- **file ARGUMENT_FILE** Read additional command-line options from ARGUMENT_FILE.
- **np PROCESSES, -count PROCESSES** Start PROCESSES instances of the executable as a single job.
Schedule the job to run for a maximum of \texttt{MAX_TIME} minutes.

Request that the job use the allocation \texttt{PROJECT} when submitting the job to the LRM.

Request that the job be submitted to the LRM using the named \texttt{QUEUE}.

Run the job in the directory named by \texttt{DIRECTORY}. Input and output files will be interpreted relative to this directory. This directory must exist on the file system on the LRM-managed resource. If not specified, the job will run in the home directory of the user the job is running as.

Define an environment variable named by \texttt{NAME} with the value \texttt{VALUE} in the job environment. This option may be specified multiple times to define multiple environment variables.

Use the file named by \texttt{STDIN_FILE} as the standard input of the job. If the \texttt{-l} option is specified, then this file is interpreted to be on a file system local to the LRM. If the \texttt{-s} option is specified, then this file is interpreted to be on the file system where \texttt{globus-job-run} is being executed, and the file will be staged via GASS. If neither is specified, the local behavior is assumed.

Use the file named by \texttt{STDOUT_FILE} as the destination for the standard output of the job. If the \texttt{-l} option is specified, then this file is interpreted to be on a file system local to the LRM. If the \texttt{-s} option is specified, then this file is interpreted to be on the file system where \texttt{globus-job-run} is being executed, and the file will be staged via GASS. If neither is specified, the local behavior is assumed.

Use the file named by \texttt{STDERR_FILE} as the destination for the standard error of the job. If the \texttt{-l} option is specified, then this file is interpreted to be on a file system local to the LRM. If the \texttt{-s} option is specified, then this file is interpreted to be on the file system where \texttt{globus-job-run} is being executed, and the file will be staged via GASS. If neither is specified, the local behavior is assumed.

Add a set of custom RSL attributes described by \texttt{RSL_CLAUSE} to the job description. The clause must be an RSL conjunction and may contain one or more attributes. This can be used to include attributes which cannot be defined by other command-line options of \texttt{globus-job-run}.

When included outside the context of \texttt{-stdin}, \texttt{-stdout}, or \texttt{-stderr} command-line options, \texttt{-l} option alters the interpretation of the executable path. If the \texttt{-l} option is specified, then the executable is interpreted to be on a file system local to the LRM.

When included outside the context of \texttt{-stdin}, \texttt{-stdout}, or \texttt{-stderr} command-line options, \texttt{-s} option alters the interpretation of the executable path. If the \texttt{-s} option is specified, then the executable is interpreted to be on the file system where \texttt{globus-job-run} is being executed, and the file will be staged via GASS. If neither is specified, the local behavior is assumed.

\textbf{ENVIRONMENT}

If the following variables affect the execution of \texttt{globus-job-run}.
globus-job-run

X509_USER_PROXY  Path to proxy credential.

X509_CERT_DIR  Path to trusted certificate directory.

See Also

globusrun(1), globus-job-submit(1), globus-job-clean(1), globus-job-get-output(1), globus-job-cancel(1)
Name

globus-job-status -- Check the status of a GRAM5 job

globus-job-status JOBID

Description

The globus-job-status program checks the status of a GRAM job by sending a status request to the job manager contact for that job specified by the JOBID parameter. If successful, it will print the job status to standard output. The states supported by globus-job-status are:

**PENDING** The job has been submitted to the LRM but has not yet begun execution.
**ACTIVE** The job has begun execution.
**FAILED** The job has failed.
**SUSPENDED** The job is currently suspended by the LRM.
**DONE** The job has completed.
**UNSUBMITTED** The job has been accepted by GRAM, but not yet submitted to the LRM.
**STAGE_IN** The job has been accepted by GRAM and is currently staging files prior to being submitted to the LRM.
**STAGE_OUT** The job has completed execution and is currently staging files from the service node to other http, GASS, or GridFTP servers.

Options

The full set of options to globus-job-status are:

- **-help, -usage** Display a help message to standard error and exit.
- **-version** Display the software version of the globus-job-status program to standard output.
- **-versions** Display the software version of the globus-job-status program including DiRT information to standard output.

ENVIRONMENT

If the following variables affect the execution of globus-job-status.

- **X509_USER_PROXY** Path to proxy credential.
- **X509_CERT_DIR** Path to trusted certificate directory.
Bugs

The `globus-job-status` program cannot distinguish between the case of the job manager terminating for any reason and the job being in the `DONE` state.

See Also

globusrun(1)
Name

globus-job-submit -- Submit a batch job using GRAM

globus-job-submit [-dumprsl] [-dryrun] [-verify]
[-file ARGUMENT_FILE]
SERVICE_CONTACT
[-np PROCESSES | -count PROCESSES ]
[-m MAX_TIME | -maxtime MAX_TIME ]
[-p PROJECT | -project PROJECT ]
[-q QUEUE | -queue QUEUE ]
[-d DIRECTORY | -directory DIRECTORY ] [-env NAME=VALUE]...
[-x RSL_CLAUSE]
[-l | -s ] EXECUTABLE [ARGUMENT...]  

Description

The globus-job-submit program constructs a job description from its command-line options and then submits the job to the GRAM service running at SERVICE_CONTACT. The executable and arguments to the executable are provided on the command-line after all other options. Note that the -dumprsl, -dryrun, -verify, and -file command-line options must occur before the first non-option argument, the SERVICE_CONTACT.

The globus-job-submit provides similar functionality to globusrun in that it allows batch submission of GRAM jobs. However, unlike globusrun, it uses command-line parameters to define the job instead of RSL expressions.

To retrieve the output and error streams of the job, use the program globus-job-get-output. To reclaim resources used by the job by deleting cached files and job state, use the program globus-job-clean. To cancel a batch job submitted by globus-job-submit, use the program globus-job-cancel.

Options

The full set of options to globus-job-submit are:

- -help, -usage              Display a help message to standard error and exit.
- -version                  Display the software version of the globus-job-submit program to standard output.
- -versions                 Display the software version of the globus-job-submit program including DiRT information to standard output.
- -dumprsl                  Translate the command-line options to globus-job-submit into an RSL expression that can be used with tools such as globusrun.
- -dryrun                   Submit the job request to the GRAM service with the dryrun option enabled. When this option is used, the GRAM service prepares to execute the job but stops before submitting the job to the LRM. This can be used to diagnose some problems such as missing files.
- -verify                   Submit the job request to the GRAM service with the dryrun option enabled and then without it enabled if the dryrun is successful.
-file ARGUMENT_FILE Read additional command-line options from ARGUMENT_FILE.

-np PROCESSES,-count PROCESSES Start PROCESSES instances of the executable as a single job.

-m MAX_TIME,-maxtime MAX_TIME Schedule the job to run for a maximum of MAX_TIME minutes.

-p PROJECT,-project PROJECT Request that the job use the allocation PROJECT when submitting the job to the LRM.

-q QUEUE,-queue QUEUE Request that the job be submitted to the LRM using the named QUEUE.

-d DIRECTORY,-directory DIRECTORY Run the job in the directory named by DIRECTORY. Input and output files will be interpreted relative to this directory. This directory must exist on the file system on the LRM-managed resource. If not specified, the job will run in the home directory of the user the job is running as.

-env NAME=VALUE Define an environment variable named by NAME with the value VALUE in the job environment. This option may be specified multiple times to define multiple environment variables.

-stdin [-l | -s] STDIN_FILE Use the file named by STDIN_FILE as the standard input of the job. If the -l option is specified, then this file is interpreted to be on a file system local to the LRM. If the -s option is specified, then this file is interpreted to be on the file system where globus-job-submit is being executed, and the file will be staged via GASS. If neither is specified, the local behavior is assumed.

-stdout [-l | -s] STDOUT_FILE Use the file named by STDOUT_FILE as the destination for the standard output of the job. If the -l option is specified, then this file is interpreted to be on a file system local to the LRM. If the -s option is specified, then this file is interpreted to be on the file system where globus-job-submit is being executed, and the file will be staged via GASS. If neither is specified, the local behavior is assumed.

-stderr [-l | -s] STDERR_FILE Use the file named by STDERR_FILE as the destination for the standard error of the job. If the -l option is specified, then this file is interpreted to be on a file system local to the LRM. If the -s option is specified, then this file is interpreted to be on the file system where globus-job-submit is being executed, and the file will be staged via GASS. If neither is specified, the local behavior is assumed.

-x RSL_CLAUSE Add a set of custom RSL attributes described by RSL_CLAUSE to the job description. The clause must be an RSL conjunction and may contain one or more attributes. This can be used to include attributes which can not be defined by other command-line options of globus-job-submit.

-l When included outside the context of -stdin, -stdout, or -stderr command-line options, -l option alters the interpretation of the executable path. If the -l option is specified, then the executable is interpreted to be on a file system local to the LRM.

-s When included outside the context of -stdin, -stdout, or -stderr command-line options, -s option alters the interpretation of the executable path. If the -s option is specified, then the executable is interpreted to be on the file system where globus-job-submit is being executed, and the file will be staged via GASS. If neither is specified, the local behavior is assumed.
system where `globus-job-run` is being executed, and the file will be staged via GASS. If neither is specified, the local behavior is assumed.

ENVIRONMENT

If the following variables affect the execution of `globus-job-submit`.

- **X509_USER_PROXY**: Path to proxy credential.
- **X509_CERT_DIR**: Path to trusted certificate directory.

See Also

globusrun(1), globus-job-run(1), globus-job-clean(1), globus-job-get-output(1), globus-job-cancel(1)
**Name**

globus-personal-gatekeeper -- Manage a user's personal gatekeeper daemon

[-log [=DIRECTORY]] [-seg] [-acctfile ACCOUNTING_FILE]
globus-personal-gatekeeper [-killall] [-kill]

**Description**

The `globus-personal-gatekeeper` command is a utility which manages a gatekeeper and job manager service for a single user. Depending on the command-line arguments it will operate in one of several modes. In the first set of arguments indicated in the synopsis, the program provides information about the `globus-personal-gatekeeper` command or about instances of the `globus-personal-gatekeeper` that are running currently. The second set of arguments indicated in the synopsis provide control over starting a new `globus-personal-gatekeeper` instance. The final set of arguments provide control for terminating one or more `globus-personal-gatekeeper` instances.

The `-start` mode will create a new subdirectory of `$HOME/.globus` and write the configuration files needed to start a `globus-gatekeeper` daemon which will invoke the `globus-job-manager` service when new authenticated connections are made to its service port. The `globus-personal-gatekeeper` then exits, printing the contact string for the new gatekeeper prefixed by GRAM contact: to standard output. In addition to the arguments described above, any arguments described in `globus-job-manager(8)` can be appended to the command-line and will be added to the job manager configuration for the service started by the `globus-gatekeeper`.

The new `globus-gatekeeper` will continue to run in the background until killed by invoking `globus-personal-gatekeeper` with the `-kill` or `-killall` argument. When killed, it will kill the `globus-gatekeeper` and `globus-job-manager` processes, remove state files and configuration data, and then exit. Jobs which are running when the personal gatekeeper is killed will continue to run, but their job directory will be destroyed so they may fail in the LRM.

The full set of command-line options to `globus-personal-gatekeeper` consists of:

- `--help, -help` Print command-line option summary and exit
- `--version` Print software version
- `--versions` Print software version including DiRT information
- `--list` Print a list of all currently running personal gatekeepers. These entries will be printed one per line.
- `--directory CONTACT` Print the configuration directory for the personal gatekeeper with the contact string `CONTACT`.
- `--debug` Print additional debugging information when starting a personal gatekeeper. This option is ignored in other modes.
- `--start` Start a new personal gatekeeper process.
- `--jmtype LRM` Use `LRM` as the local resource manager interface. If not provided when starting a personal gatekeeper, the job manager will use the default fork LRM.
- `--auditdir AUDIT_DIRECTORY` Write audit report files to `AUDIT_DIRECTORY`. If not provided, the job manager will not write any audit files.
-port PORT

Listen for gatekeeper TCP/IP connections on the port PORT. If not provided, the gatekeeper will let the operating system choose.

-log[=DIRECTORY]

Write job manager log files to DIRECTORY. If DIRECTORY is omitted, the default of $HOME will be used. If this option is not present, the job manager will not write any log files.

-seg

Try to use the SEG mechanism to receive job state change information, instead of polling for these. These require either the system administrator or the user to run an instance of the globus-job-manager-event-generator program for the LRM specified by the -jmtype option.

-acctfile ACCOUNTING_FILE

Write gatekeeper accounting entries to ACCOUNTING_FILE. If not provided, no accounting records are written.

Examples

This example shows the output when starting a new personal gatekeeper which will schedule jobs via the lsf LRM, with debugging enabled.

% globus-personal-gatekeeper -start -jmtype lsf

verifying setup...
done.
GRAM contact: personal-grid.example.org:57846:/DC=org/DC=example/CN=Joe User

This example shows the output when listing the current active personal gatekeepers.

% globus-personal-gatekeeper -list

personal-grid.example.org:57846:/DC=org/DC=example/CN=Joe User

This example shows the output when querying the configuration directory for the above personal gatekeeper. gatekeepers.

% globus-personal-gatekeeper -directory "personal-grid.example.org:57846:/DC=org/DC=example/CN=Joe User"

/home/juser/.globus/.personal-gatekeeper.personal-grid.example.org.1337

% globus-personal-gatekeeper -kill "personal-grid.example.org:57846:/DC=org/DC=example/CN=Joe User"

killing gatekeeper: "personal-grid.example.org:57846:/DC=org/DC=example/CN=Joe User"

See Also

globusrun(1), globus-job-manager(8), globus-gatekeeper(8)
Name

globus-gram-audit -- Load GRAM4 and GRAM5 audit records into a database

globus-gram-audit [--conf CONFIG_FILE] [--check] [--delete] [--audit-directory AUDITDIR]

Description

The globus-gram-audit program loads audit records to an SQL-based database. It reads $GLOBUS_LOCATION/etc/globus-job-manager.conf by default to determine the audit directory and then uploads all files in that directory that contain valid audit records to the database configured by the globus_gram_job_manager_auditing_setup_scripts package. If the upload completes successfully, the audit files will be removed.

The full set of command-line options to globus-gram-audit consist of:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>--conf CONFIG_FILE</td>
<td>Use CONFIG_FILE instead of the default from the configuration file for audit database configuration.</td>
</tr>
<tr>
<td>--check</td>
<td>Check whether the insertion of a record was successful by querying the database after inserting the records. This is used in tests.</td>
</tr>
<tr>
<td>--delete</td>
<td>Delete audit records from the database right after inserting them. This is used in tests to avoid filling the database with test records.</td>
</tr>
<tr>
<td>--audit-directory DIR</td>
<td>Look for audit records in DIR, instead of looking in the directory specified in the job manager configuration. This is used in tests to control which records are loaded to the database and then deleted.</td>
</tr>
<tr>
<td>--query SQL</td>
<td>Perform the given SQL query on the audit database. This uses the database information from the configuration file to determine how to contact the database.</td>
</tr>
</tbody>
</table>

FILES

The globus-gram-audit uses the following files (paths relative to $GLOBUS_LOCATION).

<table>
<thead>
<tr>
<th>File</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>etc/globus-gram-job-manager.conf</td>
<td>GRAM5 job manager configuration. It includes the default path to the audit directory</td>
</tr>
<tr>
<td>etc/globus-gram-audit.conf</td>
<td>Audit configuration. It includes the information needed to contact the audit database.</td>
</tr>
</tbody>
</table>
Name

globus-job-manager -- Execute and monitor jobs


Description

The `globus-job-manager` program is a servivce which starts and controls GRAM jobs which are executed by a local resource management system, such as LSF or Condor. The `globus-job-manager` program is typically started by the `globus-gatekeeper` program and not directly by a user. It runs until all jobs it is managing have terminated or its delegated credentials have expired.

Typically, users interact with the `globus-job-manager` program via client applications such as `globusrun`, `globus-job-submit`, or tools such as CoG jglobus or Condor-G.

The full set of command-line options to `globus-job-manager` consists of:

- `-help` Display a help message to standard error and exit
- `-type LRM` Execute jobs using the local resource manager named `LRM`.
- `-conf CONFIG_PATH` Read additional command-line arguments from the file `CONFIG_PATH`. If present, this must be the first command-line argument to the `globus-job-manager` program.
- `-globus-host-manufacturer MANUFACTURER` Indicate the manufacturer of the system which the jobs will execute on. This parameter sets the value of the `$(GLOBUS_HOST_MANUFACTURER)` RSL substitution to `MANUFACTURER`.
- `-globus-host-cputype CPU-TYPE` Indicate the CPU type of the system which the jobs will execute on. This parameter sets the value of the `$(GLOBUS_HOST_CPUTYPE)` RSL substitution to `CPU-TYPE`.
- `-globus-host-osname OSNAME` Indicate the operating system type of the system which the jobs will execute on. This parameter sets the value of the `$(GLOBUS_HOST_OSNAME)` RSL substitution to `OSNAME`.
- `-globus-host-osversion OSVERSION` Indicate the operating system version of the system which the jobs will execute on. This parameter sets the value of the `$(GLOBUS_HOST_OSVERSION)` RSL substitution to `OSVERSION`.
- `-globus-gatekeeper-host HOST` Indicate the host name of the machine which the job was submitted to. This parameter sets the value of the `$(GLOBUS_GATEKEEPER_HOST)` RSL substitution to `HOST`.
Indicate the TCP port number of gatekeeper to which jobs are submitted to. This parameter sets the value of the $\{GLOBUS\_GATEKEEPER\_PORT\}$ RSL substitution to $PORT$.

Indicate the X.509 identity of the gatekeeper to which jobs are submitted to. This parameter sets the value of the $\{GLOBUS\_GATEKEEPER\_SUBJECT\}$ RSL substitution to $SUBJECT$.

Indicate the path where the Globus Toolkit(r) is installed on the service node. This is used by the job manager to locate its support and configuration files.

Indicate the path where the Globus Toolkit(r) is installed on the execution host. If this is omitted, the value specified as a parameter to $-home$ is used. This parameter sets the value of the $\{GLOBUS\_LOCATION\}$ RSL substitution to $TARGET\_GLOBUS\_LOCATION$.

Configure the job manager to write job history files to $HISTORY\_DIRECTORY$. These files are described in the FILES section below.

Configure the job manager to use $SCRATCH\_DIRECTORY$ as the default scratch directory root if a relative path is specified in the job RSL's $scratch\_dir$ attribute.

Configure the job manager to write log messages via syslog. Logging is further controlled by the argument to the $-log\_levels$ parameter described below.

Configure the job manager to write log messages to files in the $LOG\_DIRECTORY$ directory. Files will be named $LOG\_DIRECTORY/gram_YYYYMMDD.log$. Logging is further controlled by the argument to the $-log\_levels$ parameter described below. The $LOG\_DIRECTORY$ value can include variables derived from the job manager environment using the same syntax as RSL substitutions. For example, $-stdio\_log$ $(HOME)$ would cause each user's logs to be stored in their individual home directories.

Configure the job manager to write log messages of certain levels to syslog and/or log files. The available log levels are FATAL, ERROR, WARN, INFO, DEBUG, and TRACE. Multiple values can be combined with the $|$ character. The default value of logging when enabled is FATAL | ERROR.

Configure the job manager to write state files to $STATE\_DIRECTORY$. If not specified, the job manager uses the default of $\{GLOBUS\_LOCATION\}/tmp/gram_job_state/$. This directory must be writable by all users and be on a file system which supports POSIX advisory file locks.

Configure the job manager to restrict its TCP/IP communication to use ports in the range described by $PORT\_RANGE$. This value is also made available in the job environment via the $GLOBUS\_TCP\_PORT\_RANGE$ environment variable.

Configure the job manager to search $TRUSTED\_CERTIFICATE\_DIRECTORY$ for its list of trusted CA certificates and their signing policies. This value is also made available in the job environment via the $X509\_CERT\_DIR$ environment variable.
Configure the job manager to use the path `GASS_CACHE_DIRECTORY` for its temporary GASS-cache files. This value is also made available in the job environment via the `GLOBUS_GASS_CACHE_DEFAULT` environment variable.

Configure the job manager to assume it is using Kerberos for authentication instead of X.509 certificates. This disables some certificate-specific processing in the job manager.

Configure the job manager to define a set of environment variables in the job environment beyond those defined in the base job environment. The format of the parameter to this argument is a comma-separated sequence of `VAR=VAL` pairs, where `VAR` is the variable name and `VAL` is the variables value.

Configure the job manager to use the schedule event generator module named by `SEG_MODULE` to detect job state changes events from the local resource manager, in place of the less efficient polling operations used in GT2. To use this, one instance of the `globus-job-manager-event-generator` must be running to process events for the LRM into a generic format that the job manager can parse.

Configure the job manager to write audit records to the directory named by `AUDIT_DIRECTORY`. This records can be loaded into a database using the `globus-gram-audit` program.

Configure the job manager to use `TOOLKIT_VERSION` as the version for audit and usage stats records.

Configure the job manager to use `SERVICE_TAG` as a unique identifier to allow multiple GRAM instances to use the same job state directories without interfering with each other's jobs. If not set, the value `untagged` will be used.

Configure the job manager to disable file streaming. This is propagated to the LRM script interface but has no effect in GRAM5.

Disable sending of any usage stats data, even if `-usagestats-targets` is present in the configuration.

Send usage packets to a data collection service for analysis. The `TARGET` string consists of a comma-separated list of HOST:PORT combinations, each containing an optional list of data to send. See Usage Stats Packets\(^1\) for more information about the tags. Special tag strings of all (which enables all tags) and default may be used, or a sequence of characters for the various tags. If this option is not present in the configuration, then the default of `usage-stats.globus.org:4810` is used.

Set the architecture specification for condor jobs to be `ARCH` in job classified ads generated by the GRAM5 condor LRM script. This is required for the condor LRM but ignored for all others.

Set the operating system specification for condor jobs to be `OS` in job classified ads generated by the GRAM5 condor LRM script. This is required for the condor LRM but ignored for all others.

\(^1\) http://confluence.globus.org/display/~bester/GRAM5+Usage+Stats+Packets
Environment

If the following variables affect the execution of **globus-job-manager**

- **HOME** User's home directory.
- **LOGNAME** User's name.
- **JOBMANAGER_SYSLOG_ID** String to prepend to syslog audit messages.
- **JOBMANAGER_SYSLOG_FAC** Facility to log syslog audit messages as.
- **JOBMANAGER_SYSLOG_LVL** Priority level to use for syslog audit messages.
- **GATEKEEPER_JM_ID** Job manager ID to be used in syslog audit records.
- **GATEKEEPER_PEER** Peer information to be used in syslog audit records.
- **GLOBUS_ID** Credential information to be used in syslog audit records.
- **GLOBUS_JOB_MANAGER_SLEEP** Time (in seconds) to sleep when the job manager is started. [For debugging purposes only]
- **GRID_SECURITY_HT-TP_BODY_FD** File descriptor of an open file which contains the initial job request and to which the initial job reply should be sent. This file descriptor is inherited from the **globus-gatekeeper**.
- **X509_USER_PROXY** Path to the X.509 user proxy which was delegated by the client to the **globus-gatekeeper** program to be used by the job manager.
- **GRID_SECURITY_CONTEXT_FD** File descriptor containing an exported security context that the job manager should use to reply to the client which submitted the job.
- **GLOBUS_USAGE_TARGETS** Default list of usagestats services to send usage packets to.

Files

- **$HOME/.globus/job/HOST-NAME/LRM.TAG.red** Job manager delegated user credential.
- **$HOME/.globus/job/HOST-NAME/LRM.TAG.lock** Job manager state lock file.
- **$HOME/.globus/job/HOST-NAME/LRM.TAG.pid** Job manager pid file.
- **$HOME/.globus/job/HOST-NAME/LRM.TAG.sock** Job manager socket for inter-job manager communications.
- **$HOME/.globus/job/HOST-NAME/JOB_ID/** Job-specific state directory.
- **$HOME/.globus/job/HOST-NAME/JOB_ID/stdin** Standard input which has been staged from a remote URL.
$HOME/.globus/job/HOST-NAME/JOB_ID/stdout

$HOME/.globus/job/HOST-NAME/JOB_ID/stderr

$HOME/.globus/job/HOST-NAME/JOB_ID/x509_user_proxy

$GLOBUS_LOCATION/tmp/gram_job_state/job.HOST-NAME.JOB_ID

$GLOBUS_LOCATION/tmp/gram_job_state/job.HOST-NAME.JOB_ID.lock

$GLOBUS_LOCATION/etc/globus-job-manager.conf

$GLOBUS_LOCATION/etc/grid-services/jobmanager-LRM

See Also
globusrun(1), globus-gatekeeper(8), globus-personal-gatekeeper(1), globus-gram-audit(8)
Name

globus-job-manager-event-generator -- Create LRM-independent SEG files for the job manager to use

globus-job-manager-event-generator [-help] {-scheduler LRM} [-background] [-pidfile PIDPATH]

Description

The globus-job-manager-event-generator program is a utility which uses LRM-specific SEG parsers to generate a LRM-independent log file that a job manager instance can use to process job status change events. This program runs independently of all globus-job-manager instances so that only one process needs to deal with the LRM interface. The globus-job-manager-event-generator program can be run as a privileged user if required to interface with the LRM.

The full set of command-line options to globus-job-manager-event-generator consists of:

- help
  Print command-line option summary and exit.

- scheduler LRM
  Process events for the local resource manager named by LRM.

- background
  Run globus-job-manager-event-generator as a background process. It will fork a new process, print out its process ID and then the original process will terminate.

- pidfile PIDPATH
  Write the process ID of an instance of globus-job-manager-event-generator to the file named by PIDPATH. This file can be used to kill or monitor the globus-job-manager-event-generator process.

Files

globus-job-manager-seg.conf
  Configuration file for globus-job-manager-event-generator. Each line consists of a string of the form LRM_log_path=PATH, which indicates the directory containing LRM-independent format SEG log files for the LRM. This file is created by the running the globus_scheduler_event_generator_job_manager_setup setup package.

See Also

globus-scheduler-event-generator(8), globus-job-manager(8)
GSI-OpenSSH Commands

The gsiss(1), gsiscp(1), and gsisftp(1) commands provide the same interfaces as the standard OpenSSH ssh, scp, and sftp commands, respectively, with the added ability to perform X.509 proxy credential authentication and delegation.
Name

gsissh -- Secure remote login

gsissh

Tool description

Use the gsissh command to securely login to a remote machine.

Command syntax

gsissh [-l login_name] hostname | user@hostname [command]
Name

gsiscp -- Secure remote file copy

gsiscp

Tool description

Use the gsiscp command to securely copy files to or from a remote machine.

Command syntax

```
gsiscp [-P port] [[user@]host1:]file1 [...] [[user@]host2:]destfile
```
Name

gsisftp -- Secure file transfer

gsisftp

Tool description

The gsisftp command provides an interactive interface for transferring files to and from remote machines.

Command syntax

gsisftp [[user@]host[:dir/]]
## Glossary

### B

**Bloom filter**
Compression scheme used by the Replica Location Service (RLS) that is intended to reduce the size of soft state updates between Local Replica Catalogs (LRCs) and Replica Location Index (RLI) servers. A Bloom filter is a bit map that summarizes the contents of a Local Replica Catalog (LRC). An LRC constructs the bit map by applying a series of hash functions to each logical name registered in the LRC and setting the corresponding bits.

### 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.

### E

**extended block mode (MODE E)**
MODE E is a critical GridFTP component because it allows for out of order reception of data. This in turn, means we can send the data down multiple paths and do not need to worry if one of the paths is slower than the others and the data arrives out of order. This enables parallelism and striping within GridFTP. In MODE E, a series of “blocks” are sent over the data channel. Each block consists of:

- an 8 bit flag field,
- a 64 bit field indicating the offset in the transfer,
- and a 64 bit field indicating the length of the payload,
- followed by length bytes of payload.

Note that since the offset and length are included in the block, out of order reception is possible, as long as the receiving side can handle it, either via something like a seek on a file, or via some application level buffering and ordering logic that will wait for the out of order blocks.

### L

**Local Replica Catalog (LRC)**
Stores mappings between logical names for data items and the target names (often the physical locations) of replicas of those items. Clients query the LRC to discover replicas associated with a logical name. Also may associate attributes with logical or target names. Each LRC periodically sends information about its logical name mappings to one or more RLIs.

See also RLI\(^3\).

**logical file name**
A unique identifier for the contents of a file.

---

\(^3\) #rli

XML to PDF by RenderX XEP XSL-FO Formatter, visit us at [http://www.renderx.com/](http://www.renderx.com/)
<table>
<thead>
<tr>
<th><strong>P</strong></th>
<th>physical file name</th>
<th>The address or the location of a copy of a file on a storage system.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>proxy credentials</td>
<td>The combination of a proxy certificate and its corresponding private key. GSI typically stores proxy credentials in <code>/tmp/x509up_u&lt;uid&gt;</code>, where <code>&lt;uid&gt;</code> is the user id of the proxy owner.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>R</strong></th>
<th>Replica Location Index (RLI)</th>
<th>Collects information about the logical name mappings stored in one or more Local Replica Catalogs (LRCs) and answers queries about those mappings. Each RLI periodically receives updates from one or more LRCs that summarize their contents.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RLS attribute</td>
<td>Descriptive information that may be associated with a logical or target name mapping registered in a Local Replica Catalog (LRC). Clients can query the LRC to discover logical names or target names that have specified RLS attributes.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>S</strong></th>
<th>server</th>
<th>A process that receives commands and sends responses to those commands. Since it is a server or service, and it receives commands, it must be listening on a port somewhere to receive the commands. Both FTP and GridFTP have IANA registered ports. For FTP it is port 21, for GridFTP it is port 2811. This is normally handled via inetd or xinetd on Unix variants. However, it is also possible to implement a daemon that listens on the specified port. This is described more fully in in the Architecture section of the GridFTP Developer's Guide.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>stream mode (MODE S)</td>
<td>The only mode normally implemented for FTP is MODE S. This is simply sending each byte, one after another over the socket in order, with no application level framing of any kind. This is the default and is what a standard FTP server will use. This is also the default for GridFTP.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>T</strong></th>
<th>third party transfers</th>
<th>In the simplest terms, a third party transfer moves a file between two GridFTP servers.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The following is a more detailed, programmatic description.</td>
<td>In a third party transfer, there are three entities involved. The client, who will only orchestrate, but not actually take place in the data transfer, and two servers one of which will be sending data to the other. This scenario is common in Grid applications where you may wish to stage data from a data store somewhere to a supercomputer you have reserved. The commands are quite similar to the client/server transfer. However, now the client must establish two control channels, one to each server. He will then choose one to listen, and send it the PASV command. When it responds with the IP/port it is listening on, the client will send that IP/port as part of the PORT command to the other server. This will cause the second server to connect to the first server, rather than the client. To initiate the actual movement of the data, the client then sends the RETR “filename” command to the server that</td>
</tr>
</tbody>
</table>
will read from disk and write to the network (the “sending” server) and will send the STOR “filename” command to the other server which will read from the network and write to the disk (the “receiving” server).

See Also client/server transfer.