Globus GridFTP

Transfer of bulk baseband data is most efficiently done using GridFTP (part of the Globus Toolkit). GridFTP can be used without grid certificates via the sshftp url mechanism described below. Note that this uses ssh only to instigate the transfer between the two hosts - the data are then transferred using GridFTP and are not encrypted or otherwise handled by ssh so there is no performance penalty.

Transfer data

To use the `globus-url-copy` command below you will have to set the environment variable $GLOBUS_LOCATION and modify your path appropriately.

Recommended command for copying data:

```
globus-url-copy -cd -r -sync -sync-level 1 -restart -udt -vb
file://<from_path>/ sshftp://user@magnus-data.pawsey.org.au/<to_path>/
```

where <from_path> is the directory on pbstore and <to_path> is the destination path on the cuppa node. If either path is a directory you must terminate the url with a forward slash (/).

The `-udt` flag uses udt as the protocol. This is faster than tcp (at least for long haul transfers), but is not always available. TCP transfer may be improved using parallelism (e.g. substitute `-p 4` for `-udt` in the command above).

Installing GridFTP

Binary installation (no UDT option)

The simplest option by far is a binary installation, but currently the binary installation does not include the udt libraries (see installing from source below).

Follow the procedures at: http://www.globus.org/toolkit/data/gridftp/quickstart.html

To get the server and client programs running and configured:

```
sudo apt-get install globus-ftp-client-progs  globus-gridftp
export GLOBUS_LOCATION=/usr
```

Notes for compiling from source (required for UDT)

Globus Toolkit 5.*

You can also build from source by following the procedures here:

http://www.globus.org/toolkit/data/gridftp/quickstart-source.html

The 'latest-stable' installer linked from the page above actually installs a rather old version of GridFTP, so you may wish to download a more recent version from here: http://www.globus.org/ftppub/. We have had good success with version gt5.0.5: http://www.globus.org/ftppub/gt5/5.0/5.0.5

For ubuntu 10.4 you may also need libssl-dev: i.e. the following may need to be installed

```
sudo apt-get install build-essential bzip2 autoconf libxml-parser-perl
xinetd openssl telnet libssl-dev
```

Note that to enable UDT transfers, you should use the following make command (or similar depending on the version of gridftp/OS):

```
cd gt*-all-source-installer
./configure --prefix=/path/to/install/to
make globus_libtool globus_libtool-thr udt globus-xio-extra-drivers gridftp
install
```

To configure the server and client:

```
$GLOBUS_LOCATION/setup/globus/setup-globus-common
$GLOBUS_LOCATION/setup/globus/globus-gridftp-server-enable-sshftp -nonroot
$GLOBUS_LOCATION/setup/globus/setup-globus-gridftp-sshftp
```

On some systems you may get a version conflict if you have a system installation of the Compress::Zlib perl module. The gridftp installer comes with its own version of this library but will only install it if it doesn't find a system version. You can force the installation of the gridftp version thus:

```
cd /<path_to_installer>/gpt/support/Compress-Zlib-1.21/
perl Makefile.PL PREFIX=$GLOBUS_LOCATION
make install
```

Globus Toolkit 6.*

On more recent OSes (after Debian 7.11 - wheezy) SSL library changes mean you will want Globus toolkit version 6: http://toolkit.globus.org/ftppub/gt6/installers/src/ . We have had success with version globus_toolkit-6.0.1421093009

Build instructions are here: http://toolkit.globus.org/toolkit/docs/latest-stable/gridftp/admin/index.html

The following worked on Debian 7.11:

```
apt-get install build-essential bzip2 autoconf libxml-parser-perl xinetd
openssl telnet libssl-dev gettext
configure --prefix=$GLOBUS_LOCATION
make  udt xio gridftp install
$GLOBUS_LOCATION/sbin/globus-gridftp-server-enable-sshftp -nonroot
```

With the version globus_toolkit-6.0.1421093009 I found it necessary, after installation, to edit the last line of $GLOBUS_LOCATION/libexec/gridftp-ssh to replace @@SSH@@ with the correct path to ssh (/usr/bin/ssh). This was probably due to an unseen error in the configure script.

Testing the Installation

Once you have set up ssh keys to the destination machine, a nice test that everything is working for outgoing connections (and a rough idea of transfer speed) is the following command

```
globus-url-copy -vb -p 4 /dev/zero
sshftp://username@hpc-data.pawsey.org.au/dev/null
```

If you get *error: globus_ftp_client: an invalid value for url was used* this probably means that the sshftp url wasn't understood. Double-check that the globus-ftp-client-progs were installed from apt-get. For source installations, check that the $GLOBUS_LOCATION/setup/globus/setup-globus-gridftp-sshftp command worked.

From:
https://www.atnf.csiro.au/vlbi/dokuwiki/ - **ATNF VLBI Wiki**

Permanent link:
https://www.atnf.csiro.au/vlbi/dokuwiki/doku.php/correlator/globus

Last update: **2019/03/15 14:18**