Lab 4: Data Management

This lab will cover these data management tools:

- scp
- globus-url-copy
- UberFTP
- RLS utilities

Note: because some of these command lines are longer than one line of text can hold, the beginning of each command is marked with a bullet, as shown in this entirely bogus command you shouldn't try:

- `entirely-bogus-command -pilsner_urquell=3 gsiftp://home/zippidee/parameter1 gsiftp://home/zippidee/parameter2`

Some of these commands need you to substitute your username in those places where *username* is used as a placeholder.

Set up

First, we need to shell into gk1 and then authenticate to the grid. Use grid-proxy-info to check on your proxy.

- `ssh gk1.phys.utb.edu`
- `grid-proxy-init`
- `grid-proxy-info`

Since this lab is mostly about moving files around, we need to create some files to move. This next sequence will put one file in your home directory on gk1 and another on your home directory on gk2. The two `ls` commands will let you verify that the files were created. (Here we're using the `dd` command only to create empty files of arbitrary length.)

- `dd if=/dev/zero of=./gk1yourname bs=1000 count=10`
- `ls -l`
- `ssh gk2.phys.utb.edu`
- `dd if=/dev/zero of=./gk2yourname bs=1000 count=10`
- `ls -l`
- `exit`

scp

The syntax of scp is similar to the cp command.

Moving a file from localhost to another host:

- `scp gk1yourname gk2.phys.utb.edu:gk1yourname`

Moving a file from another host to local:

- `scp gk2.phys.utb.edu:gk2yourname .`

- `ls`
- `ssh gk2.phys.utb.edu`
- `ls`

Now remove the files you've copied, because you'll copy them again in the next section.

- `rm gk1yourname`
- `exit`
- `rm gk2yourname`

globus-url-copy

The basic syntax of globus-url-copy is `globus-url-copy` <from> <to>.

Moving a file from localhost to another host:

- `globus-url-copy file:/home/trainingXX/gk1yourname gsiftp://gk2.phys.utb.edu/home/trainingXX/gk1yourname`

Moving a file from another host to local:

- `globus-url-copy gsiftp://gk2.phys.utb.edu/home/training23/gk2yourname file:/gk1.phys.utb.edu/home/training23/gk2yourname`
- `ls`
- `ssh gk2.phys.utb.edu`
- `ls`

Delete those copied files again.

- `rm gk1yourname`
- `exit`
- `rm gk2yourname`

Let's try a globus_url_copy without authorization, just to see what breaks.

First, destroy your grid proxy:

- `grid-proxy-destroy`

Now that you're a grid outlaw, try the globus_url_copy again:

- `globus-url-copy file:/home/username/gk1yourname gsiftp://gk2.phys.utb.edu/home/username/gk1yourname`

Notice the error message. Buried in it somewhere you should find "Could not find a valid proxy certificate file location." (Grid error messages are often not as helpful as you might like, although they're getting better.)

Now we'll reauthenticate and try a third-party copy, moving a file from one remote host to another remote host, in this case clu1.

- `grid-proxy-init`
- `globus_url_copy gsiftp://gk2.phys.utb.edu/home/username/gk2yourname gsiftp://clu1.phys.utb.edu/home/username/gk2yourname`

- `ssh clu1.phys.utb.edu`
- `ls`
- `rm gk2yourname`
- `exit`

UberFTP

Make sure you have access to UberFTP:

- `uberftp -v`

If you're already familiar with FTP, then UberFTP should be no problem to use. Try the following:

- `uberftp -H gk2.phys.utb.edu -a gsi`

The -H flag specifies the host you're connecting to; the -a flag specifies that you're going to use GSI to authenticate. The system should respond with a login message and an `uberftp>` prompt. Let's look around:

- `ls`

You should see a list of your files on gk2. To move gk2yourname to gk1:

- `get gk2yourname`

To move gk1yourname to gk2:

- `put gk1yourname`

To exit UberFTP:

- `quit`

Using RLS

First, let's check that the RLS services are running on gk1 and gk2:

- `globus-rls-admin -p rls://gk1`
- `globus-rls-admin -p rls://gk2`

In both cases, you should receive a reply that says the services are up and running. If you want to see more information about the RLS servers, you can use the -S flag.

- `globus-rls-admin -S rls://gk1`
- `globus-rls-admin -S rls://gk2`

Now add create a mapping for your test files in the gk1 LRC.

- `globus-rls-cli create gk1yourname gsiftp://home/yourname/gk1yourname rls://gk1`

You can create more than one mapping associated with one file. This time, we'll use a **file:** designation rather than a **gsiftp:** designation.

- `globus-rls-cli add gk1yourname file://localhost/home/yourname/gk1yourname rls://gk1`

This information will be propagated out to the RLS services on gk2 when gk1's LRI updates gk2's LRI. The update happens every fifteen minutes. You can query the RLS on gk2 to see whether it's learned of the existence of your files yet.

- globus-rls-cli query rli lfn gk1yourname rls://gk2

If you don't see it yet, keep trying now and then. Some time in the next fifteen minutes, the RLS on gk2 will get the news.

While you're waiting, you can also use a wildcard query to see who else has registered their files with the gk1 and gk2 LRC.

- globus-rls-cli query wildcard lrc lfn "*" rls://gk1
- globus-rls-cli query wildcard lrc lfn "*" rls://gk2