

B2STAGE and gridFTP

- Simple data up and download to irods
- Transfer of Collections recursively
- Synchronising local and remote data
- Third-party transfers
- GridFTP and PIDs

Managing Data with gridFTP

We will use the gridFTP client *globus-url-copy* and *UberFTP* to manage our data in iRODS. We will use the *globus-url-copy* client by default. However, *UberFTP* has some complementary features, which we will explore.

First we need to create a proxy for the gridFTP client:

```
grid-proxy-init
```

You can list your iRODS home collection with:

```
globus-url-copy -list gsiftp://eudat-training2/aliceZone/home/<user>/
```

- gsiftp: endpoint
- eudat-training2: alias for the server, check /etc/hosts
- aliceZone: iRODS zone name

Exercise

How can you connect with the icommands to the iRODS server?

Solution

The irods server is *eudat-training2*, the irods zone is *aliceZone*. You can assume that you connect through the standard control port 1247. Now you just need your username and password.

Upload data

Create a non-empty file and upload it to the iRODS server:

```
echo "My first test" > testfile1.txt
globus-url-copy testfile1.txt \
  gsiftp://eudat-training2/aliceZone/home/alice/testfile1.txt
uberftp file:///home/ubuntu/testfile1.txt \
  gsiftp://eudat-training2/aliceZone/home/alice/testfile2.txt
```

Exercise: Upload of nested collection

Inspect the help of *globus-url-copy*. How would you transfer a whole folder?

- Create a folder with non-empty data files
- Transfer the folder to iRODS
- How would you assign a different folder name on the gridFTP server?
- Change the content of some files, add some new files in your local data folder
- Explore the *-sync* and *-sync-level* options to update your data folder

Solution Create data:

```
mkdir -p Data/SubCollection

for i in {000..002}; do
echo "Data/File${i} and some text.">"Data/File${i}.txt";
done

for i in {003..010}; do
echo "Data/SubCollection/File${i} and some text.">"Data/SubCollection/File${i}";
done
```

Solution Transfer data recursively with *globus-url-copy -r* and *-cd*

```
globus-url-copy -cd -r Data/ gsiftp://eudat-training2/aliceZone/home/alice/Data1/
```

Solution Exploring the option *-sync*

- Create an extra file in the Collection

```
echo "extra file" > Data/extrafile.txt
```

- Modify an existing file

```
vim Data/File000
```

```
Data/File000 and some text.  
Some more text.
```

- Synchronise the local and the remote data

```
globus-url-copy -cd -r -sync -sync-level 0 Data/ \  
  gsiftp://eudat-training2/aliceZone/home/alice/Data1/
```

- Compare the result: The *extrafile.txt* has been created but the content of *File000.txt* did not change:

```
globus-url-copy -list gsiftp://eudat-training2/aliceZone/home/alice/Data1/  
uberftp -cat gsiftp://eudat-training2/aliceZone/home/alice/Data1/File000.txt
```

Removing data with uberftp

The *globus-url-copy* client does not offer to delete data. We will use *uberftp* instead.

- Removing several files with wildcards

```
uberftp -rm gsiftp://eudat-training2/aliceZone/home/alice/File00*.txt
```

- Removing folders recursively

Remove all folders that start with *Data*

```
uberftp -rm -r gsiftp://eudat-training2/aliceZone/home/alice/Data*
```

Watch out, data that is deleted with *uberftp* is gone forever. There is no soft and hard delete as with iRODS.

Third-party transfers

We will now transfer data between the iRODS and another gridFTP server. You all have access to the user *training* on *eudat-training3*:

```
globus-url-copy -list gsiftp://eudat-training3/home/training/
```

This gridFTP server uses the normal linux filesystem.

Resolving PIDs, the link to B2SAFE

Exercise: Data management by PIDs

- Create a PID for your data collection on aliceZone.

The URL needs to have the form

```
irods://130.186.13.14:1247/aliceZone/home/<user>/<collection>
```

- Copy the folder to your local home directory using the PID
- Example for listing of a collection by PID:

```
globus-url-copy -list \  
gsiftp://eudat-training2/21.T12995/e3d30ab9-ffc2-4b92-87d0-1926d3db9496/
```

Amongst others B2SAFE will take care of the proper generation of PIDs for files and collections in iRODS.