D4Science Infrastructure - Task #176

Configure DSS to support EUDAT-D4Science integration

May 28, 2015 04:18 PM - Pasquale Pagano

Status:	Closed	**Start date:**	May 28, 2015
Priority:	Normal	**Due date:**	Jun 08, 2015
Assignee:	Andrea Dell'Amico	**% Done:**	100%
Category:	Data Management	**Estimated time:**	0.00 hour
Target version:	EUDAT-D4Science		
Infrastructure:	Production		

Description

According to what reported here:
http://eudat.eu/User%20Documentation%20-%20Data%20Staging%20Script.html#UserDocumentation-DataStagingScript-Installationnotesforend-user the following steps should be performed to setup DSS

Installation notes for end-user

The following requirements need to be satisfied in order to install the Data Staging Script:

Python 2.7 (pip is suggested)
grid-proxy-utils (in particular grid-proxy-init)
python-m2crypto (script version > 1.1). You should also replace the file m2.py installed with the GO API with a new one, available on the SVN (see below).
or myproxy commands (in particular myproxy-init)
(optionally) icommands
The DSS uses the GlobusOnline service via its API to manage the transfers. For this reason, you also need the following provisions:

Acquire an account on GO and configure the needed GO end points.
Associate your X.509 certificate to your GO account under https://www.globusonline.org/account/ManageIdentities.
If you would like to transfer data from your laptop where no GridFTP server is available, please consider installing GlobusConnect.
Thanks to GO you will be able to look at your data transfers transfer in the GO web GUI.
The installation is quite simple:

get the Script from the GitHub repository.
install globus online api
for example: pip install globusonline-transfer-api-client
configure your irods environment in order to point to the iRODS server enabled with the needed rules
store the path of icommands; in datastagerconfig.py, iPATH stores the the path of the icommands, like in the following line:

```
        iPATH='<IRODS-PATH>/clients/icommands/bin/'
```

In order to use the DSS software, you should adapt it to the path of your iRODS installation.

The DSS software distribution consists of the following files:

datastagerconfig.py.example template for the required (cp to) datastagerconfig.py file (script >1.1)
PIDselecter.r the rule which gives you the PID of a file given its URL
URLselecter.r the rule which gives you the URL of a file given its PID

datastager.py the main script
datamover.py the script which invokes GO

example some concrete example
pid.file.example an example of PID file, i.e. a list of PIDs (to be used with -PF option)
task.file.example an example of path file, i.e. a list of file names (to be used with -pF option)
url.file.example an example of URL file, i.e. a list of URLs (to be used with -UF option)
README with some further information

History

#1 - Jun 11, 2015 07:28 PM - Andrea Dell'Amico

We had a conference call with Giovanni Morelli, he asked to wait for more precise guidelines. He will send them in the next days.

#2 - Jun 15, 2015 02:52 PM - Andrea Dell'Amico

- Status changed from New to In Progress

New document from Cineca: http://eudat.eu/User%20Documentation%20-%20B2STAGE.html

#3 - Mar 02, 2016 05:57 PM - Andrea Dell'Amico

- Status changed from In Progress to Paused

#4 - Sep 03, 2018 08:28 PM - Andrea Dell'Amico

Can we close this one?

#5 - Sep 04, 2018 11:16 AM - Pasquale Pagano

- Status changed from Paused to Closed

We had a conference call with Giovanni Morelli, he asked to wait for more precise guidelines. He will send them in the next days.

#2 - Jun 15, 2015 02:52 PM - Andrea Dell'Amico

- Status changed from New to In Progress