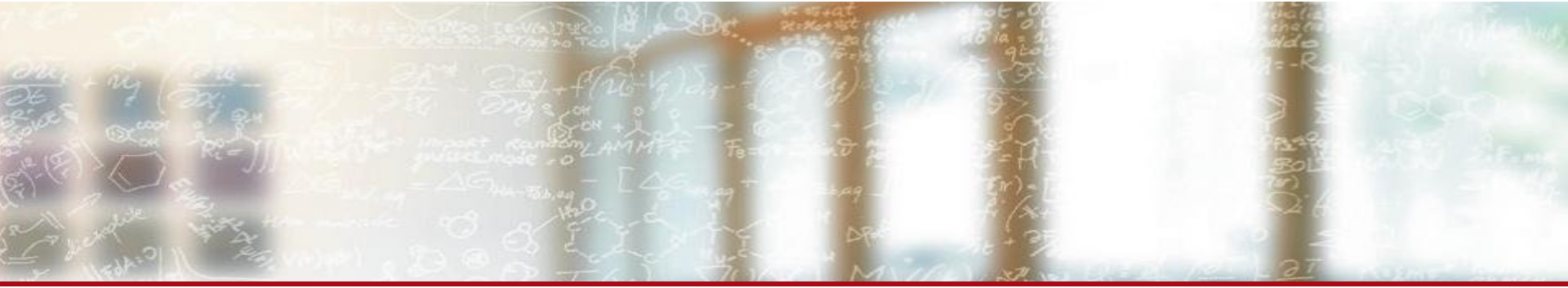
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich



Data Transfer

How to use the transfer queue and globus online

CSCS User Lab Day, August 31st 2020

Salvatore Di Nardo, CSCS

Moving data

When we talk about “moving data,” we might talk about two different scenarios:

- **EXTERNAL data transfers:** Moving data in/out CSCS (involves internet broadband usage)
- **INTERNAL data transfers:** Moving data between different storage location in CSCS network (no internet usage)

Knowing the difference it’s crucial in order to use the best tool for what you intend to do.

External data transfers

This happens when you want to upload or download data from/to CSCS environment from/to somewhere else (your own computer, your university, another research institute, and so on).

Long story short, a typical example its when you want to upload your data in order to run some jobs on them, or you want to take back the output of the jobs you ran.

For that you should use **GLOBUS ONLINE ENDPOINTS**.

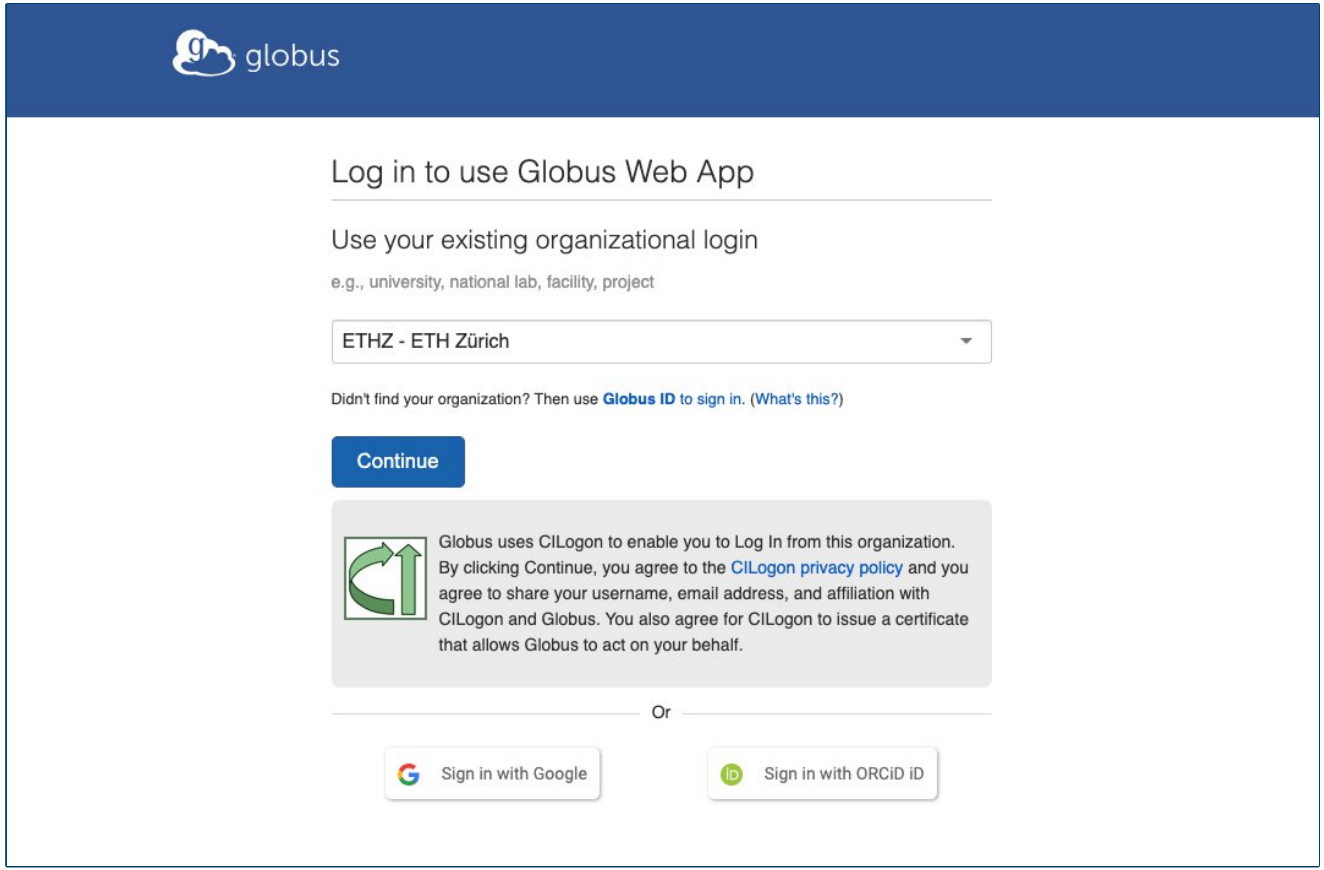
Moving data

When we talk about “moving data,” we might talk about two different scenarios:

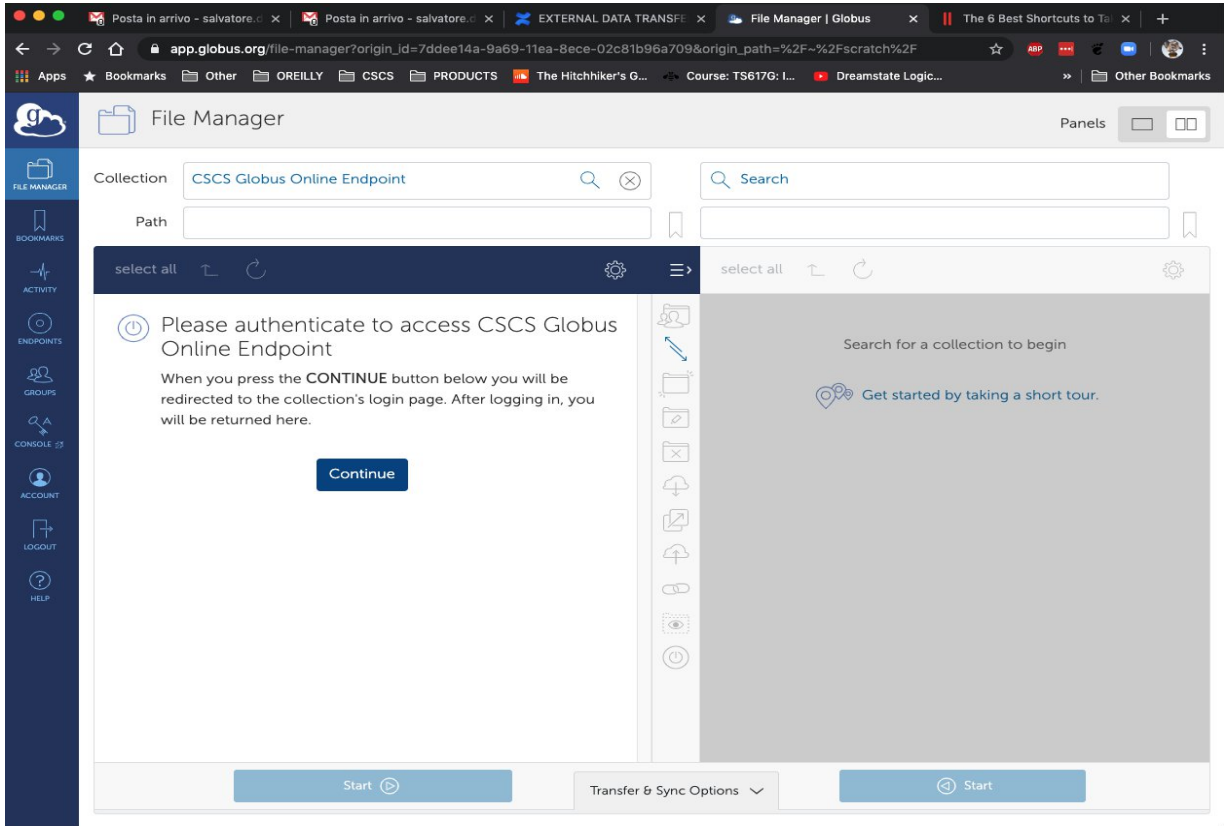
- **EXTERNAL data transfers:** Moving data in/out CSCS (involves internet broadband usage)
- **INTERNAL data transfers:** Moving data between different storage location in CSCS network (no internet usage)

Knowing the difference its crucial in order to use the best tool for what you intend to do.

www.globug.org



Search “CSCS”


Posta in arrivo - salvatore.c
Posta in arrivo - salvatore.c
EXTERNAL DATA TRANSFE
File Manager | Globus
The 6 Best Shortcuts to Ta
app.globus.org/file-manager?origin_id=7ddee14a-9a69-11ea-8ece-02c81b96a709&origin_path=%2F~%2Fscratch%2F
Apps
Bookmarks
Other
OREILLY
CSCS
PRODUCTS
The Hitchhiker's G...
Course: TS617G: I...
Dreamstate Logic...
Other Bookmarks
File Manager
Panels
FILE MANAGER
BOOKMARKS
ACTIVITY
ENDPOINTS
GROUPS
CONSOLE
ACCOUNT
LOGOUT
HELP
Collection
CSCS Globus Online Endpoint
Search
Path
select all
Please authenticate to access CSCS Globus Online Endpoint
When you press the CONTINUE button below you will be redirected to the collection's login page. After logging in, you will be returned here.
Continue
select all
Search for a collection to begin
Get started by taking a short tour.
Start
Transfer & Sync Options
Start

Select the destination

- If you want to transfer the data to another endpoint, just search for it and transfer the data.

- If you want to download the data to your PC, you will need the Globus Connect Personal client. This will turn your PC into an endpoint so you can select it and transfer the data directly in your PC.
  Download the client from the above link and follow the instructions.

Internal data transfer

If you want just mode data from one storage to another one, then the easiest way its just to submit a job that does the copy.

NOTE: make sure you submit the job to the xfer partition

Example:

```
# srun -p xfer "your command"
```

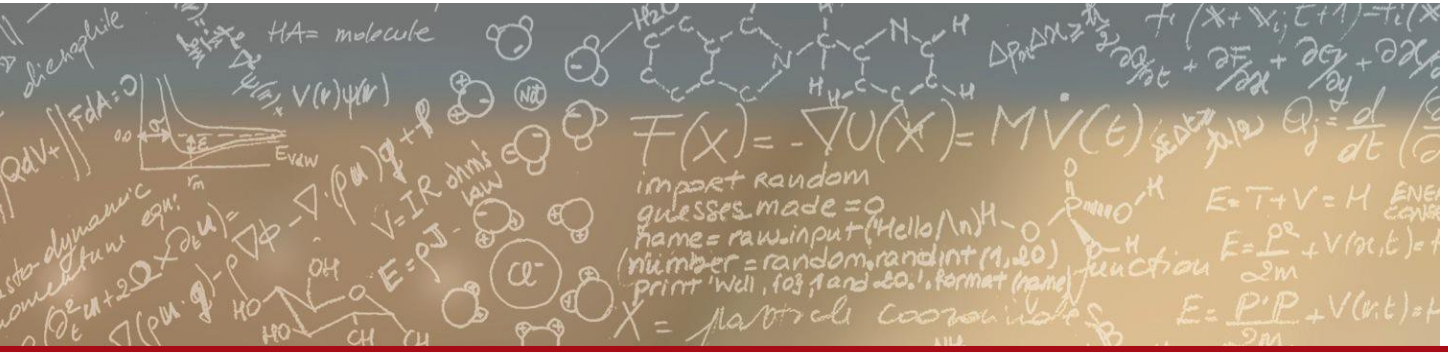
WHY the XFER partition

- That partition will execute the job in dedicated nodes called “datamovers”, optimized for data transfers.
- Jobs executed in this partition will not be counted on your job quota
- See https://user.cscs.ch/storage/transfer/internal

WHY globus online

- Secure
- Great scalability (ideal solution for big transfers)
- Parallel transfers
- resilient (if an endpoint become unreachable, the transfers pauses and automatically resumes once the endpoint is again available)
- It's a well known and used solution among research institutes
- It offers free membership

- See https://user.cscs.ch/storage/transfer/external



Thank you for your kind attention