

Step-by-step instructions for moving files with GLOBUS

GLOBUS (globus.org) is a sophisticated file transfer protocol client that you can use to move large files or lots of small files between servers without the need of your attention (as long as everything configured correctly). You will need to login in using your CWL login:



Log in

Use your organizational login

e.g., university, national lab, facility, project

By selecting Continue, you agree to Globus [terms of service](#) and [privacy policy](#).

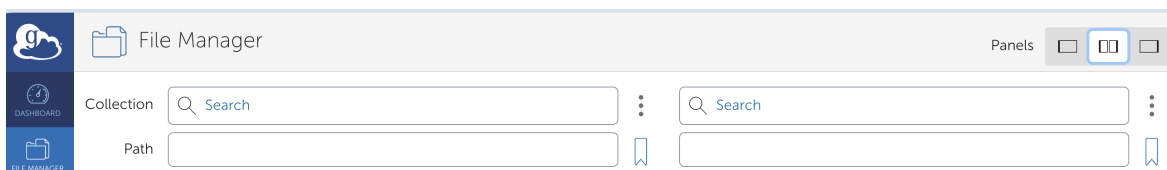
Continue

or for example your Canadian Light Source or DRAC ComputeCanada account. All will work; however, I recommend to only use the UBC CWL given that this is the way to get access to UBC ARC Sockeye and Chinook servers.

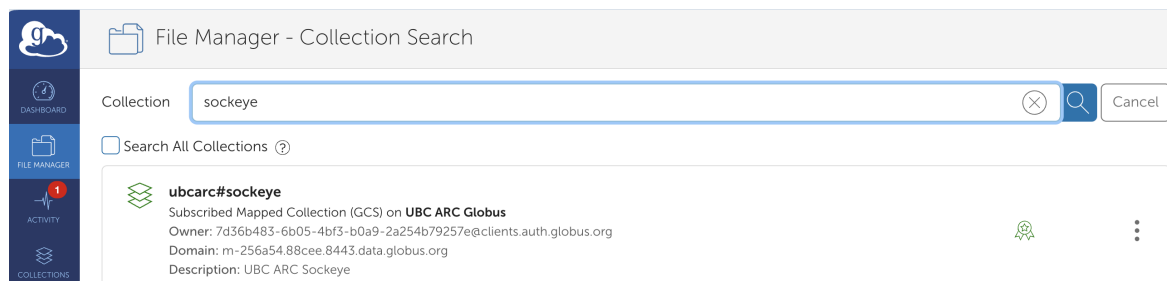
On Chinook, we have 5 TB allocation that I have started to use to archive user data who are no longer with the lab. You will need to adhere to the same process as you access Sockeye (MYVPN.OK.UBC.CA and DUO authentication).

Let's say we wish to move a large amount of files from Sockeye to M\$ OneDrive.

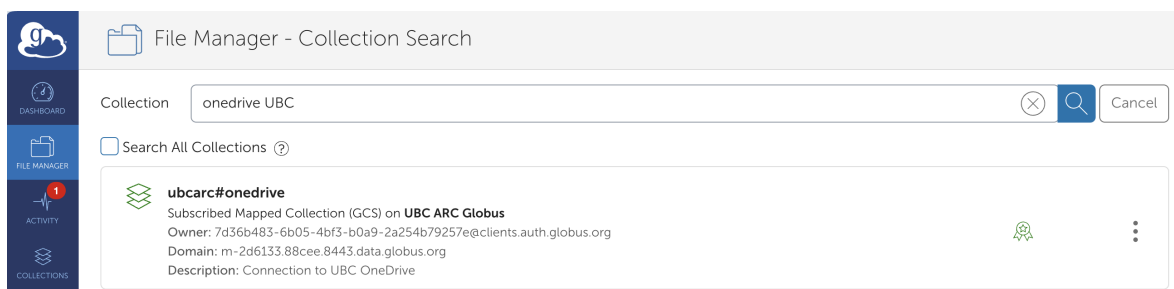
In the File manager, select the split screen view:



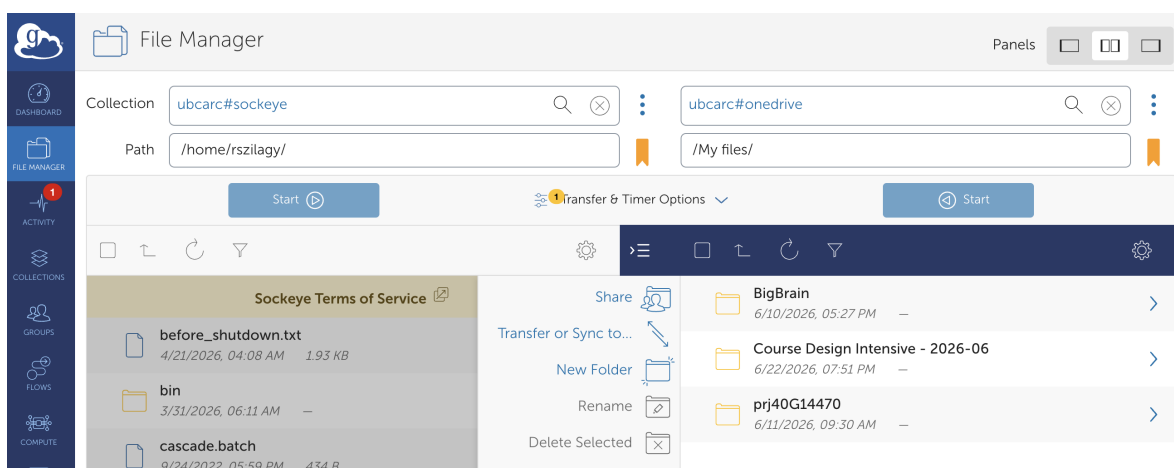
Search in the Collection field for the left-hand side “sockeye”



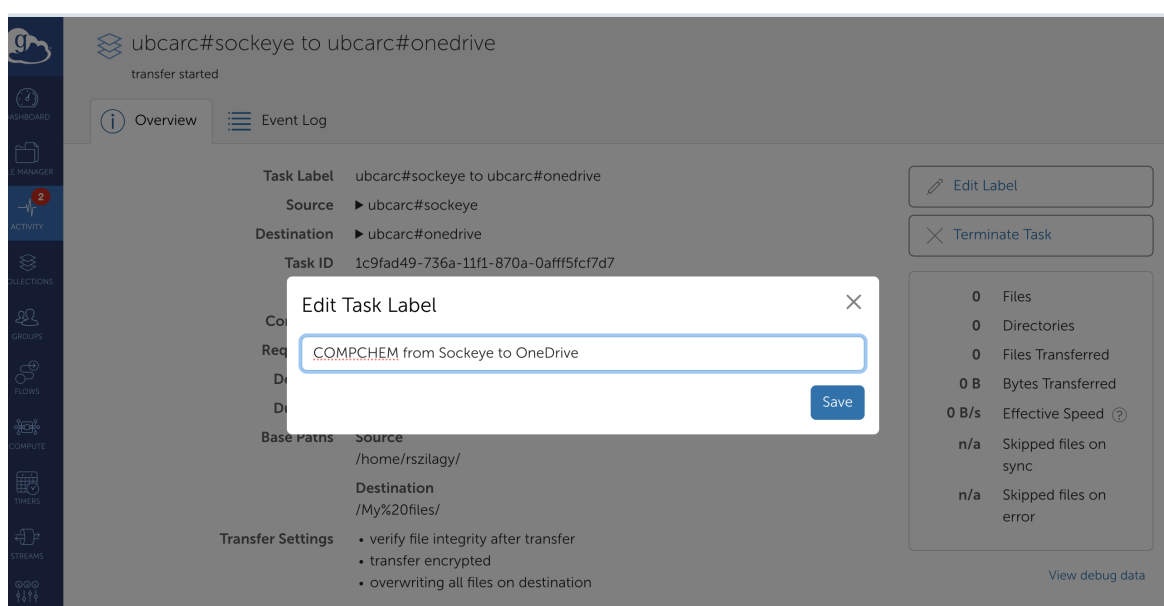
and the right-hand side for “onedrive UBC”



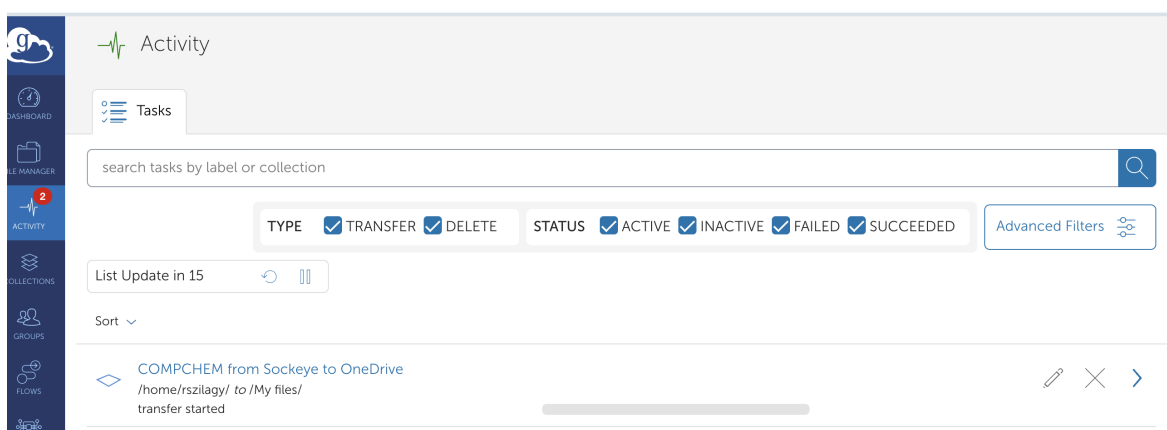
Now with the source and target destinations selected,



Navigate to the folder on Sockeye (LHS) that you wish to transfer to OneDrive (RHS). For example, I will copy the “COMPCHEM” folder to OneDrive in order to share with someone else who does not have access to Sockeye. Select the folder or a file and hit “Start” in the blue rectangle button to initiate the transfer. When the green window of Transfer Successfully Started, click on the Details and name the transfer in order to see monitor what is happening and recall what you wished to do



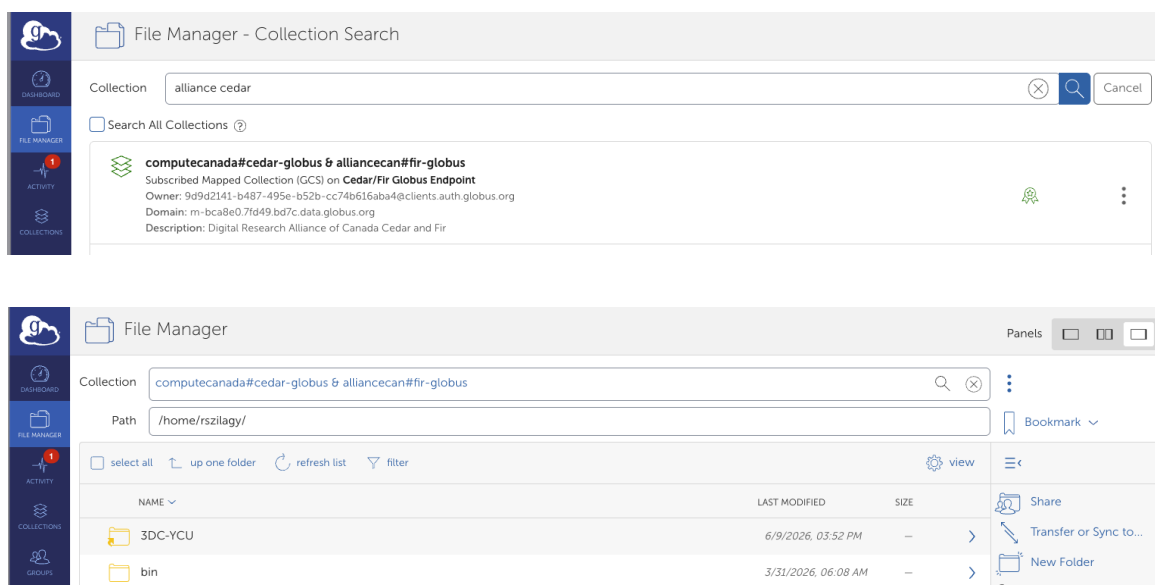
In the Activity button in the dark blue vertical bar, you can monitor what is happening:



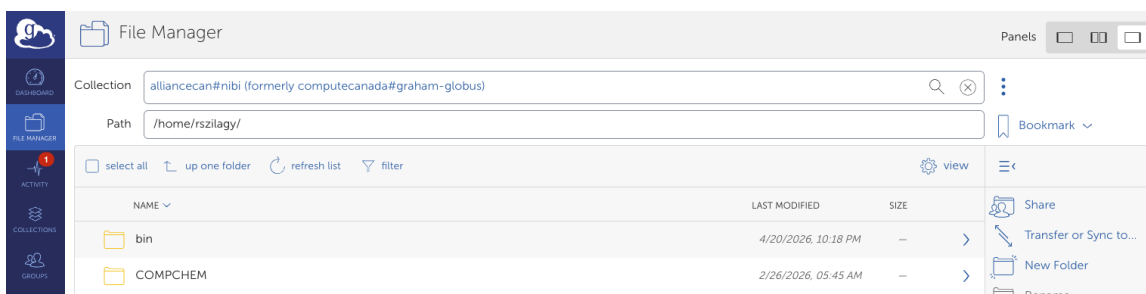
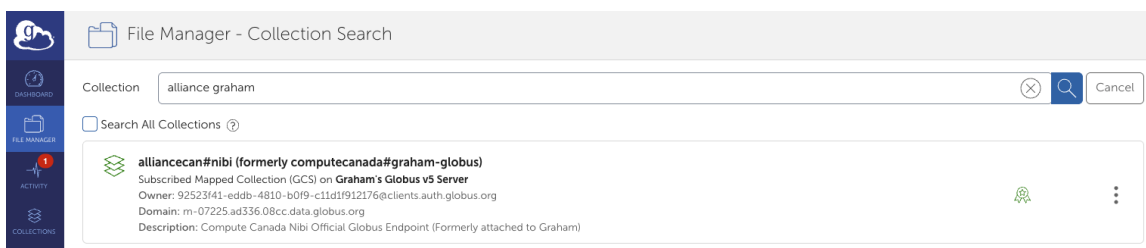
The transfer has started, but not much progress was done. When the transfer is complete, you should see a green checkmark. In any case, you will receive notification at the email address associated with your CWL account.

Using GLOBUS you can transfer also your files from DRAC or CLS or any other computer systems that are connected to GLOBUS.ORG.

Connection to FIR:

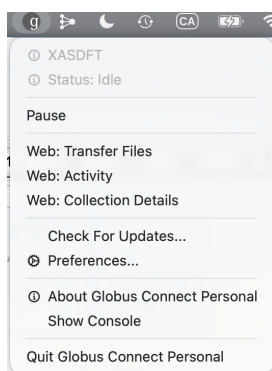


Connection to NIBI:



Conveniently, you can include your own computer as a Collection, by downloading and installing Globus Connect Personal app. With this app, you can upload and download to and from any server without needing to go through all those security steps.

On MAC, Globus Personal will appear on the top bar:



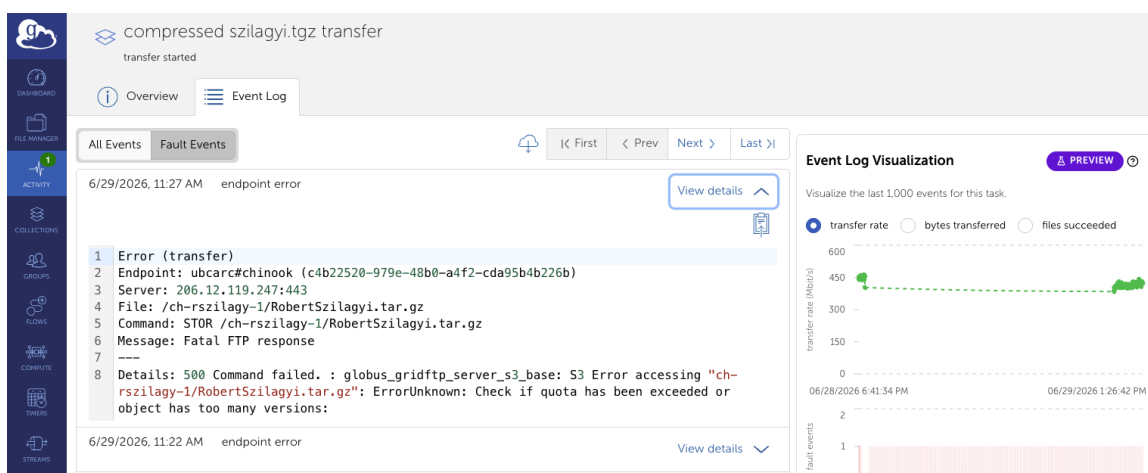
In Globus WEB app, you will need to search in the Collection:



and then it will be visible to transfer remote files to local and *vice versa*:



Last, but not least, a bit about GLOBUS' limitation. How to calculate the amount of disk space we take up. As of now (2026-06-29) we have allocated 15 TB of disk space. If we reach the quote the file transfer will give error messages such as:



In order to calculate the disk usage, you will need to access GLOBUS through the command line interface (CLI). To install GLOBUS-CLI you will need to have “pipx” on your workstation. Then issue the command “pipx install globus-cli” as described [in the notes](#). The overview of the whole process is in the [Chinook Documentation](#). Once the install is done, then issue “globus login” command on your computer, which will open an access terminal to GLOBUS. Search for our research lab's depository by

```
[UBCO-CLM-103316:~] rszilagy% globus endpoint search 'XASDFT on Chinook'
```

ID	Owner	Display Name	Domain
8c835d84-8ea5-4cb5-889c-25f72954963d	rszilagy@ubc.ca	XASDFT on Chinook	g-df9ccf.88cee.8443.data.globus.org

```
[UBCO-CLM-103316:~] rszilagy%
```

The ID is the rather unpleasant code to reference our depository. If you set the ID string to an environmental variable “CHINOOK”, you can check the connection by

```
[[UBCO-CLM-103316:~] rszilagy% globus ls $CHINOOK
JacobSchulting/
JessSamra/
JohnKVillanueva/
KayeBennett/
KevinJung/
ManjinderKour/
```

Then you can issue the “OverLord” command

```
globus ls $CHINOOK --recursive-depth-limit 10 -r --jmespath 'DATA[?
type==`file`.size' --format unix | tr '\t' '\n' | awk '{sum+=$1}END{print "Total
space used = " sum/1024/1024/1024 " GB"}'\n
```

which will calculate all the file sizes in all the folders and report it in GB unit.

```
Total space used = 1094.9 GB
```

The number of files can be generated by

```
globus ls ${GLOBUS} --recursive-depth-limit 10 -r | wc -l | awk '{print "The number
of files in the archive = " $1}'\n
```

Please be patient, the files are counted after the full list is generated () and the size determined once all the individual file sizes are known (). Not yet parallelized.

You can also limit the search to a specific folder (your data folder) and get information only for your files.

```
globus ls ${GLOBUS}:/RobertSzilagyi --recursive-depth-limit 10 -r --jmespath
'DATA[?type==`file`.size' --format unix | tr '\t' '\n' | awk '{sum+=$1}END{print
"Total space used = " sum/1024/1024/1024 " GB"}'\n
```

```
Total space used = 1894.13 GB
```

In summary, instead of using Termius and WinSCP or MOBAXterm – you can use a single universal file transfer platform that connects all your data locations at all institutions with a background queue system.

Please explore and see whether this platform calls your name.

Happy Computing !

A script for “cron” jobs to update the status of the XASDFT collection daily:

```
[UBCO-CLM-103316:~] rszilagy% cat XASDFT_on_chinook.x
#!/bin/csh

globus login
echo

globus endpoint search XASDFT
echo

# This should be "8c835d84-8ea5-4cb5-889c-25f72954963d"
setenv GLOBUS `globus endpoint search XASDFT | grep "XASDFT on Chinook" | \
    awk '{print $1}'`
echo

time globus ls ${GLOBUS} --recursive-depth-limit 10 -r \
    --jmespath 'DATA[?type==`file`].{Filename: name, Size: size}' \
    --format unix > XASDFT_on_chinook.lst
cat XASDFT_on_chinook.lst | wc -l | awk '{ print "Total number of files: " $1 }'
cat XASDFT_on_chinook.lst | awk '{sum+=$2}END{print "Total space used = " \
    sum/1024/1024/1024 " GB"}'
echo
```

and the corresponding output looks as follows

ID	Owner	Display Name	Domain
57b03cb8-8d2d-11ec-811e-493dd0cf73a1	rszilagy@ubc.ca	XASDFT	None
ca0e8b06-5dc2-11ec-b2c1-1b99bfd4976a	rszilagy@ubc.ca	XASDFT	None
69b40114-5e17-11ec-b2c1-1b99bfd4976a	rszilagy@ubc.ca	XASDFT	None
229c37f4-5c61-11ec-9b59-f9dfb1abb183	rszilagy@ubc.ca	XASDFT	None
69feeb6f-6672-11f1-995b-0afff5fcf7d7	rszilagy@ubc.ca	XASDFT	None
8c835d84-8ea5-4cb5-889c-25f72954963d	rszilagy@ubc.ca	XASDFT on Chinook	g-df9ccf.88cee.8443.data..

16.095u 3.500s 1:01:04.49 0.5% 0+0k 0+0io 625pf+0w
 Total number of files: 63973
 Total space used = 4647.77 GB