Introduction to Data Transfer using Globus

Joe Wu, PhD
NCI CCR Bioinformatics Training and Education Program
ncibtep@nih.gov

Globus

> "Globus is a high-performance data-transfer and sharing platform that allows you to move large and complex datasets directly between any two applications, systems, or local machines, eliminating the need for downloading and then uploading the data." - - https://www.cuit.columbia.edu/research-data-transfer

Why use Globus

- Recommended for transferring large quantities of data including next generation sequencing (NGS).
- "Globus will manage file transfers, monitor performance, retry failures, recover from faults automatically when possible, and report the status of your data transfer." -- Biowulf. Once a transfer is initiated, the user can walk away from the computer.
- Fast and secure way to transfer data.

Possible Globus Endpoints

Globus on NIH HPC (Biowulf)

Globus is a service that makes it easy to move, sync, and share large amounts of data. Globus will manage file transfers, monitor performance, retry failures, recover from faults automatically when possible, and report the status of your data transfer. Globus uses GridFTP for more reliable and high-performance file transfer, and will queue file transfers to be performed asynchronously in the background.

Globus was developed and is maintained at the University of Chicago and is used extensively at jor research facilities. [Globus website]

a in and out of our systems, be aware that PII and PHI data cannot be IH HPC systems.

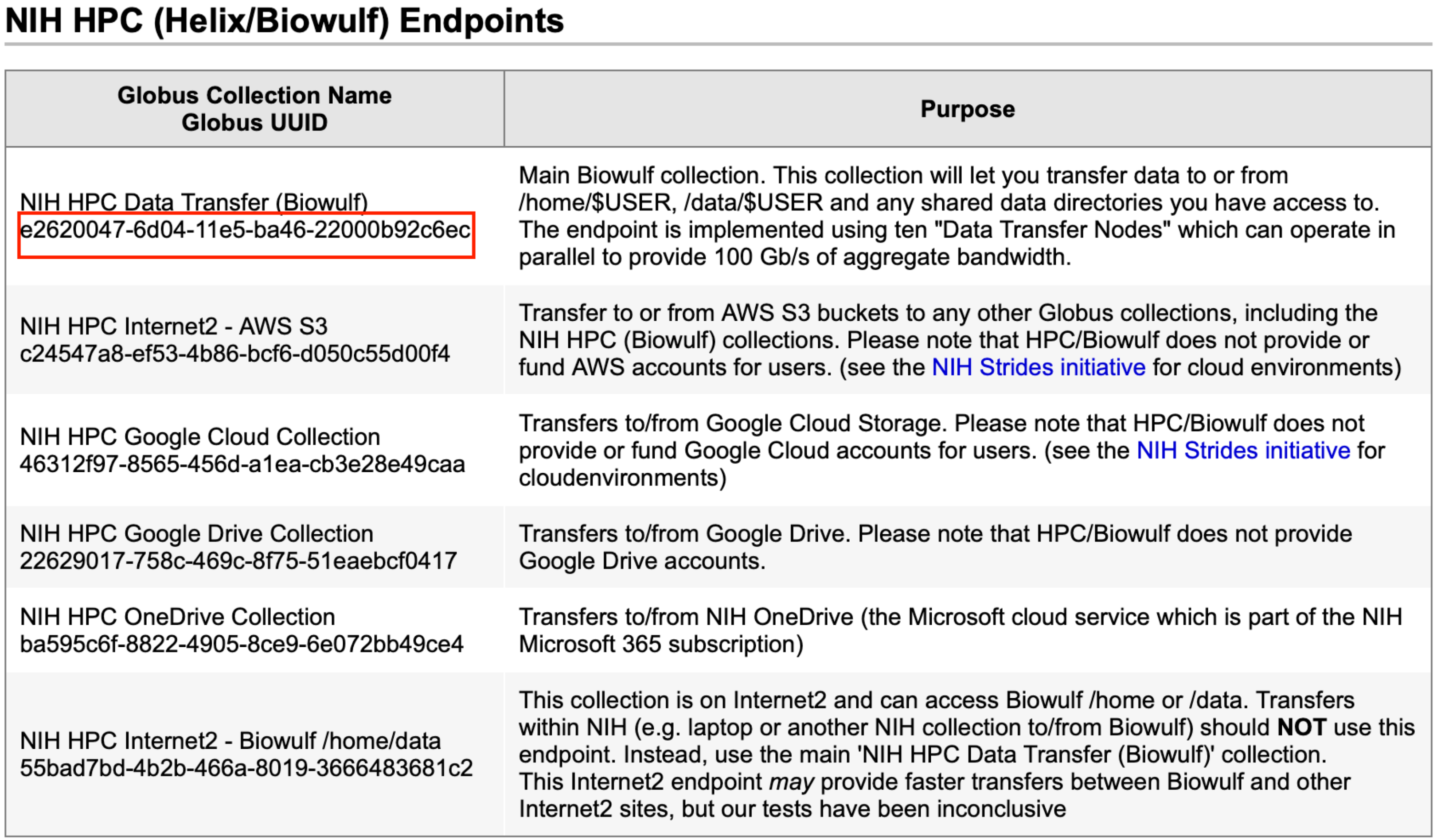
menu at left for details.

Sharing data with co

If your data is on the NIH HPC with collaborators who are at NIH or elsewhere. All they will need is a (free) Globus account. The adv s is that you do not need to transfer your data anywhere. This prevents data duplication, wastage of st have full control over which files your collaborator can access, and whether they have read-only o

Sharing data from NIH HPC s

Possible Globus Endpoints (table)



NIH HPC (Helix/Biowulf) Endpoints

Globus Collection Name **Globus UUID**	**Purpose**
NIH HPC Data Transfer (Biowulf) e2620047-6d04-11e5-ba46-22000b92c6ec	Main Biowulf collection. This collection will let you transfer data to or from /home/$USER, /data/$USER and any shared data directories you have access to. The endpoint is implemented using ten "Data Transfer Nodes" which can operate in parallel to provide 100 Gb/s of aggregate bandwidth.
NIH HPC Internet2 - AWS S3 c24547a8-ef53-4b86-bcf6-d050c55d00f4	Transfer to or from AWS S3 buckets to any other Globus collections, including the NIH HPC (Biowulf) collections. Please note that HPC/Biowulf does not provide or fund AWS accounts for users. (see the NIH Strides initiative for cloud environments)
NIH HPC Google Cloud Collection 46312f97-8565-456d-a1ea-cb3e28e49caa	Transfers to/from Google Cloud Storage. Please note that HPC/Biowulf does not provide or fund Google Cloud accounts for users. (see the NIH Strides initiative for cloudenvironments)
NIH HPC Google Drive Collection 22629017-758c-469c-8f75-51eaebcf0417	Transfers to/from Google Drive. Please note that HPC/Biowulf does not provide Google Drive accounts.
NIH HPC OneDrive Collection ba595c6f-8822-4905-8ce9-6e072bb49ce4	Transfers to/from NIH OneDrive (the Microsoft cloud service which is part of the NIH Microsoft 365 subscription)
NIH HPC Internet2 - Biowulf /home/data 55bad7bd-4b2b-466a-8019-3666483681c2	This collection is on Internet2 and can access Biowulf /home or /data. Transfers within NIH (e.g. laptop or another NIH collection to/from Biowulf) should **NOT** use this endpoint. Instead, use the main 'NIH HPC Data Transfer (Biowulf)' collection. This Internet2 endpoint *may* provide faster transfers between Biowulf and other Internet2 sites, but our tests have been inconclusive

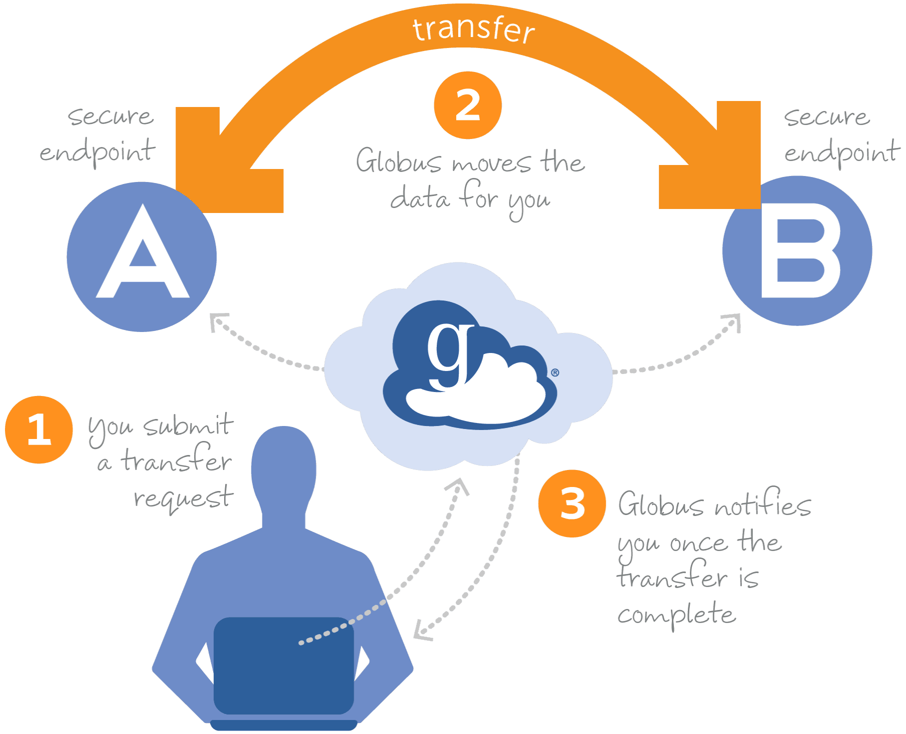
Globus Transfer Between Two Desktops

"If you need to transfer data between two Globus Connect Personal endpoints (e.g. your desktop system and your laptop, or between two desktops), you will need a Globus Plus license. Email staff@hpc.nih.gov to request one. Your Globus Plus license will be terminated when you leave NIH." -- Biowulf

General Steps to Using Globus

- Have a Biowulf account.
- Install the Globus desktop client to local computer. This enables the use of local computer as a data transfer endpoint.
    - NCI staff should submit a ticket with service.cancer.gov to get software installed on government furnished computer.
    - Staff from other ICs should contact their corresponding computing help desk for software installation.
- Setup data transfer endpoints.
- Initiate data transfer.

General Steps to Using Globus Illustrated



Source: https://www.globus.org/data-transfer

Help Resources

- Biowulf has an extensive tutorial on using Globus. See https://hpc.nih.gov/docs/globus/setup.php
- Globus Docs

Logging into Globus

Use https://www.globus.org/ to sign onto Globus. Google Chrome is recommended.

Go beyond data

Globus Compute

Reliable, distributed Function-as-a-Service

COMPUTE ANYWHERE: EDGE TO SUPERCOMPUTER

Research IT. Reimagined.

Transfer your data

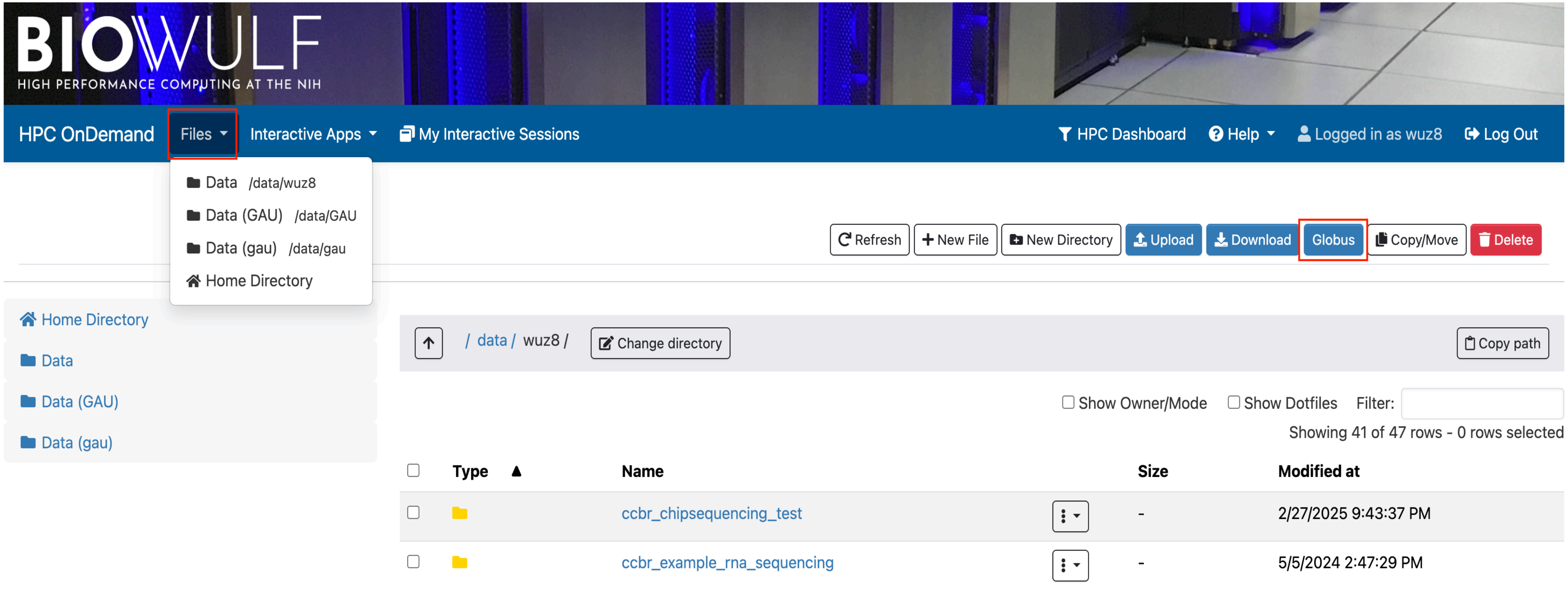
Gigabytes, terabytes, petabytes—research data is large and distributed. Globus lets you

Share your data

Globus lets you share data on your storage systems with collaborators at other

Accessing Globus from Biowulf HPC OnDemand

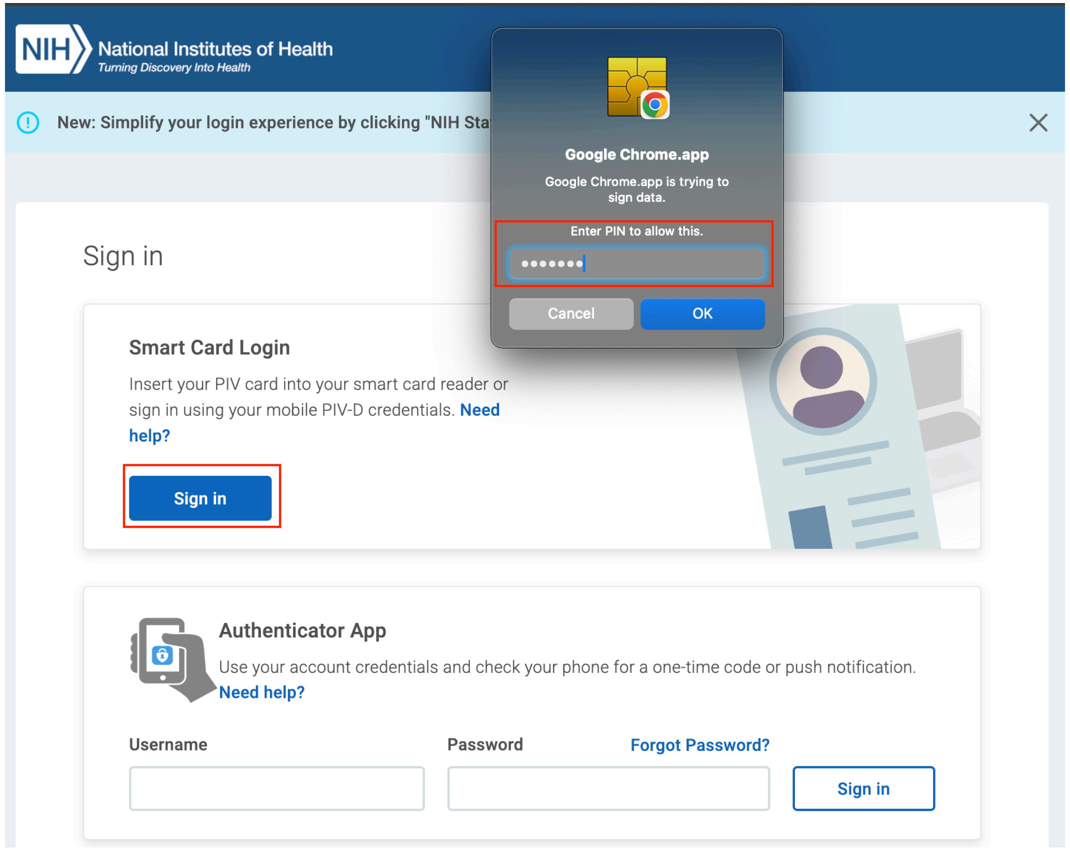
Globus can be accessed from Biowulf HPC OnDemand. Just click on any of the user's Biowulf directories under the "Files" tab and then on the "Globus" icon.



☐	Type ▲	Name	Size	Modified at
☐		ccbr_chipsequencing_test	-	2/27/2025 9:43:37 PM
☐		ccbr_example_rna_sequencing	-	5/5/2024 2:47:29 PM

Sign in with PIV Card

Subsequently, select to sign onto Globus using NIH PIV card credentials and enter PIN when prompted.



New: Simplify your login experience by clicking "NIH Sta

Google Chrome.app is trying to sign data.

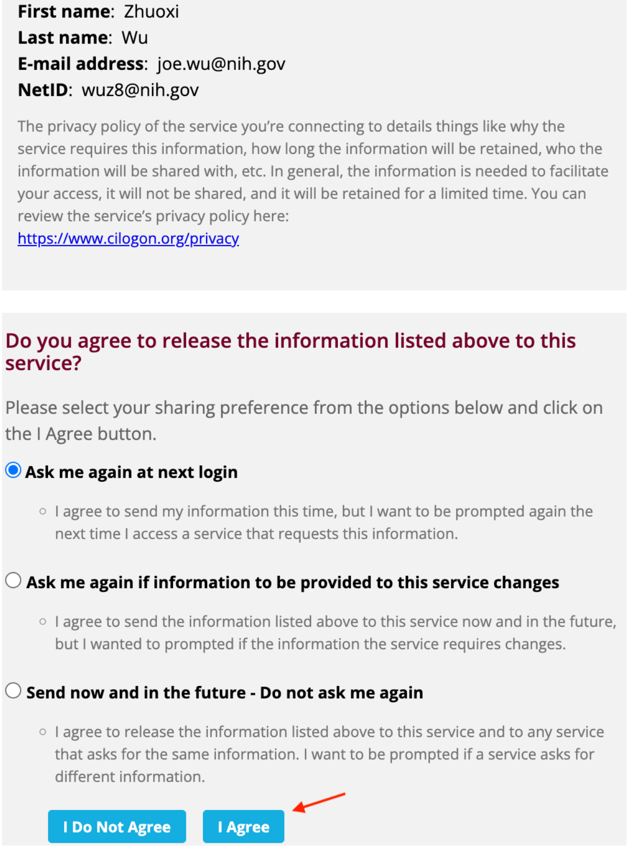
Enter PIN to allow this.

Insert your PIV card into your smart card reader or sign in using your mobile PIV-D credentials. Need help?

Use your account credentials and check your phone for a one-time code or push notification. Need help?

Agree to the Terms of Globus and Authenticate

At the next screen, click "I Agree" to accept the terms of Globus.



First name: Zhuoxi
Last name: Wu
E-mail address: joe.wu@nih.gov
NetID: wuz8@nih.gov

The privacy policy of the service you're connecting to details things like why the service requires this information, how long the information will be retained, who the information will be shared with, etc. In general, the information is needed to facilitate your access, it will not be shared, and it will be retained for a limited time. You can review the service's privacy policy here: https://www.cilogon.org/privacy

Do you agree to release the information listed above to this service?

Please select your sharing preference from the options below and click on the I Agree button.

● **Ask me again at next login**

- I agree to send my information this time, but I want to be prompted again the next time I access a service that requests this information.

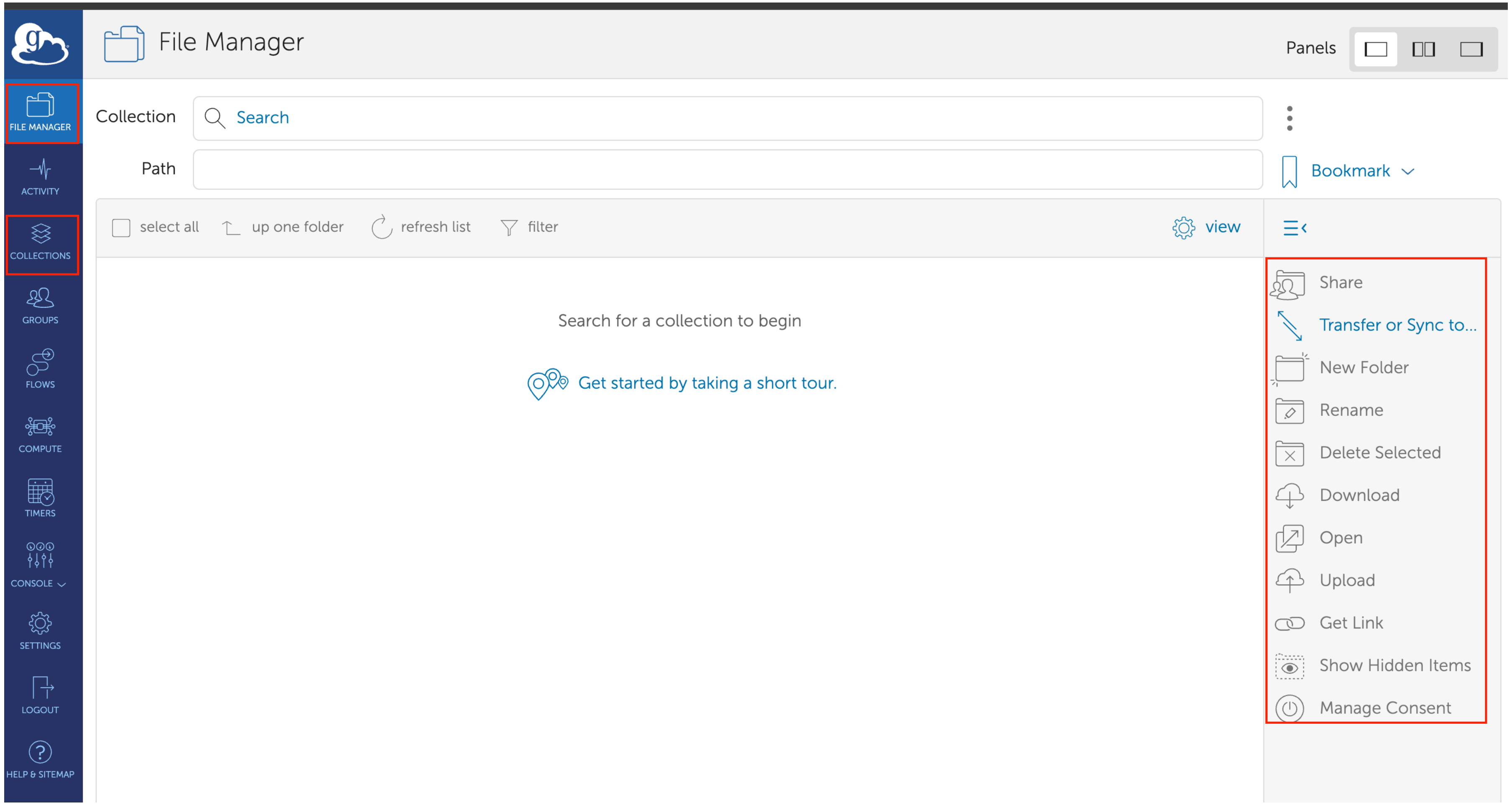
○ **Ask me again if information to be provided to this service changes**

- I agree to send the information listed above to this service now and in the future, but I wanted to prompted if the information the service requires changes.

○ **Send now and in the future - Do not ask me again**

- I agree to release the information listed above to this service and to any service that asks for the same information. I want to be prompted if a service asks for different information.

Globus Landing Page



Search for a collection to begin

Get started by taking a short tour.

Globus File Endpoint

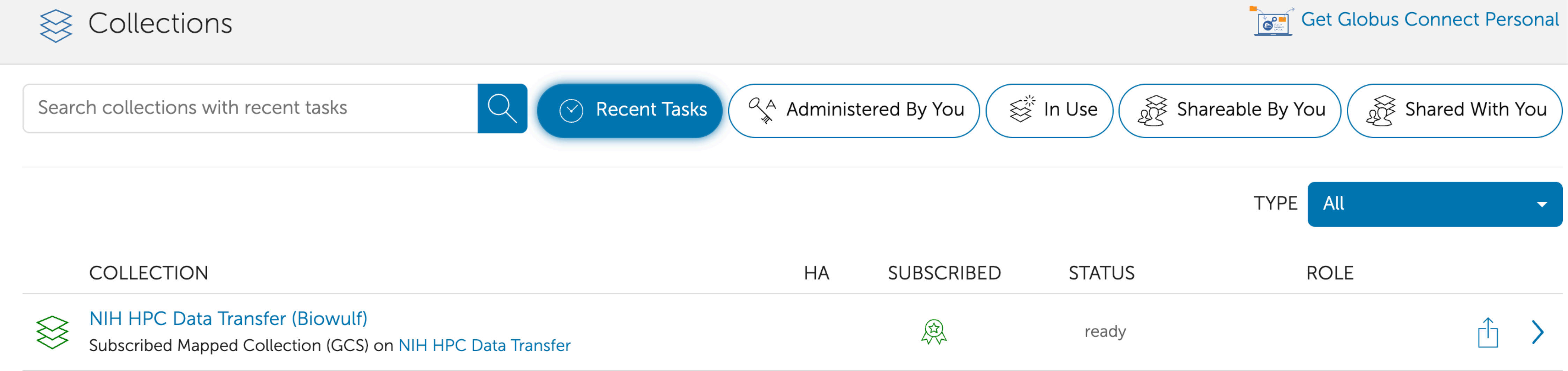
> "An "endpoint" is one of the two file transfer locations – either the source or the destination – between which files can move. Once a resource (server, cluster, storage system, laptop, or other system) is defined as an endpoint, it will be available to authorized users who can transfer files to or from this endpoint." -- Globus

Globus Collection

The Collections tab provides a table with metadata regarding the data transfer endpoints that a user has setup.

Globus Collection Table

Clicking on the "Collections" tab, the following table is shown.



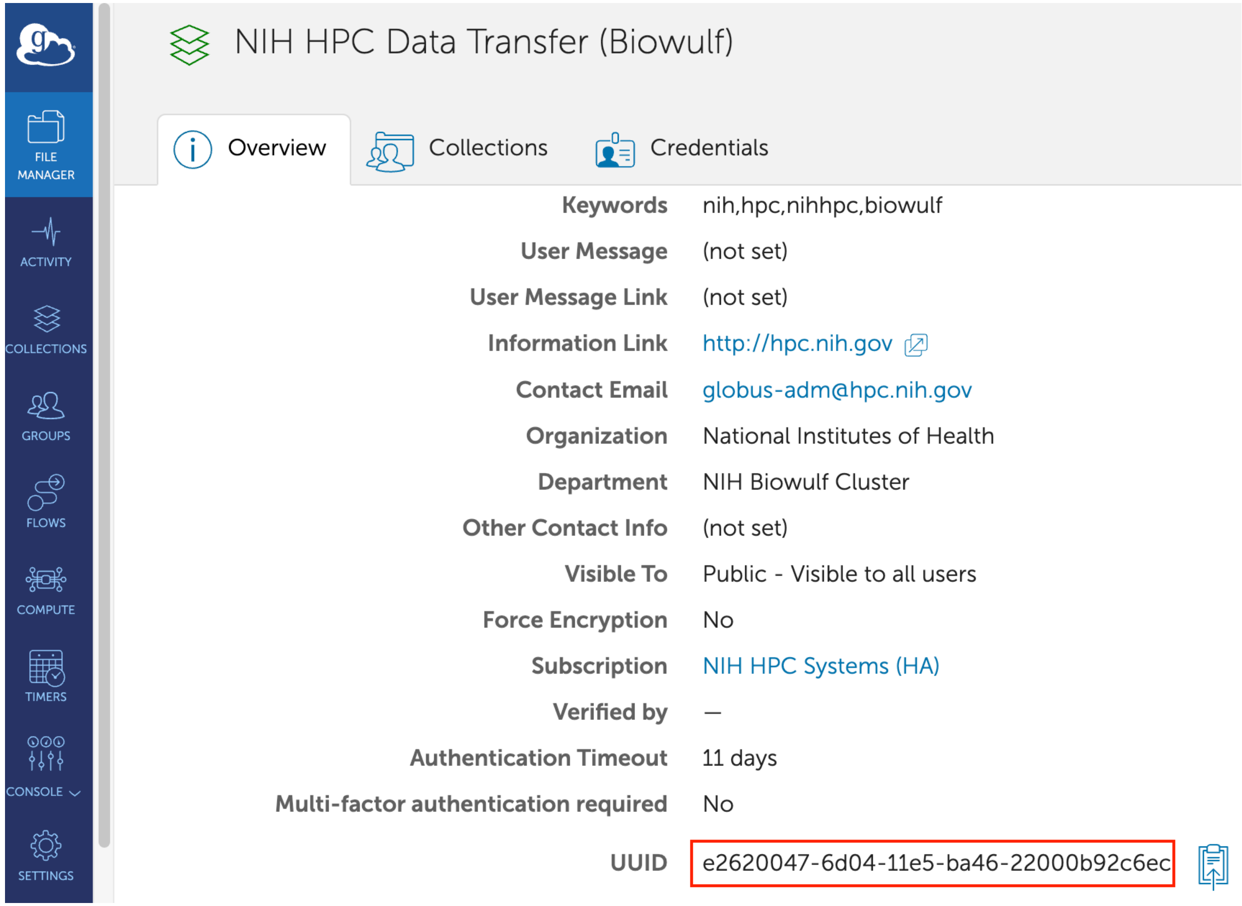
COLLECTION	HA	SUBSCRIBED	STATUS	ROLE
NIH HPC Data Transfer (Biowulf) Subscribed Mapped Collection (GCS) on NIH HPC Data Transfer			ready	

Globus Collection Table: Columns Explanation

- `COLLECTION` : contains name of the endpoint. The example below shows the NIH HPC data transfer endpoint.
- `SUBSCRIBED` : This column when checked indicates that the endpoint belongs to a organization that has a Globus subscription.
- `HA` : This column refers to high assurance collections and when checked indicates that the endpoint is suitable for things like personal health data (PHI). Biowulf cannot be used for PHI so this column is not checked.
- `STATUS` : This indicates whether the endpoint is ready to use.
- `ROLE` : Informs of whether the user has a things like administrative rights to the endpoint.
- Click on ⍐ in the far right to link to the file transfer manager.

Globus Endpoint Overview

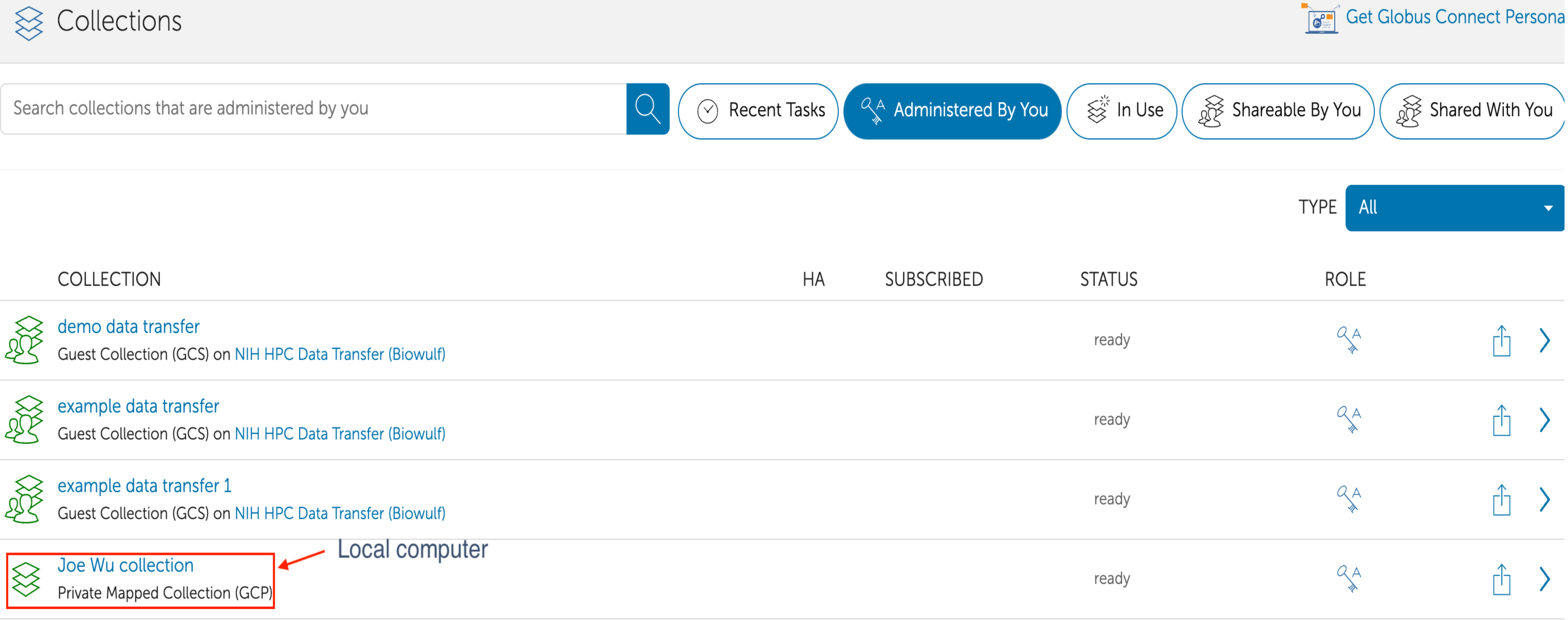
Click on ">" on the far right of the Collection table to see more detailed information regarding an endpoint. The UUID is important and needed for data transfer.



NIH HPC Data Transfer (Biowulf)

Keywords	nih,hpc,nihhpc,biowulf
User Message	(not set)
User Message Link	(not set)
Information Link	http://hpc.nih.gov
Contact Email	globus-adm@hpc.nih.gov
Organization	National Institutes of Health
Department	NIH Biowulf Cluster
Other Contact Info	(not set)
Visible To	Public - Visible to all users
Force Encryption	No
Subscription	NIH HPC Systems (HA)
Verified by	—
Authentication Timeout	11 days
Multi-factor authentication required	No
UUID	e2620047-6d04-11e5-ba46-22000b92c6ec

Globus Endpoint to Local Computer

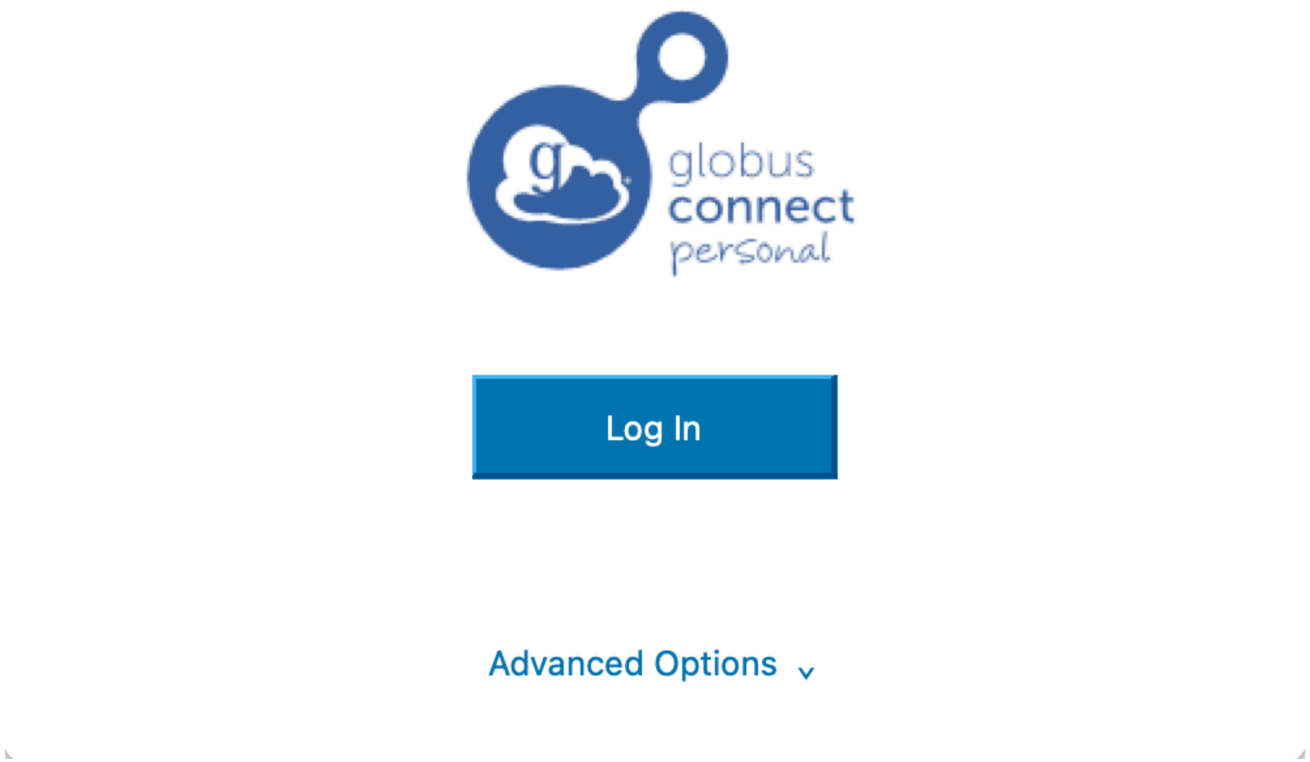


COLLECTION	HA	SUBSCRIBED	STATUS	ROLE
demo data transfer Guest Collection (GCS) on NIH HPC Data Transfer (Biowulf)			ready	
example data transfer Guest Collection (GCS) on NIH HPC Data Transfer (Biowulf)			ready	
example data transfer 1 Guest Collection (GCS) on NIH HPC Data Transfer (Biowulf)			ready	
Joe Wu collection Private Mapped Collection (GCP)			ready	

Local computer

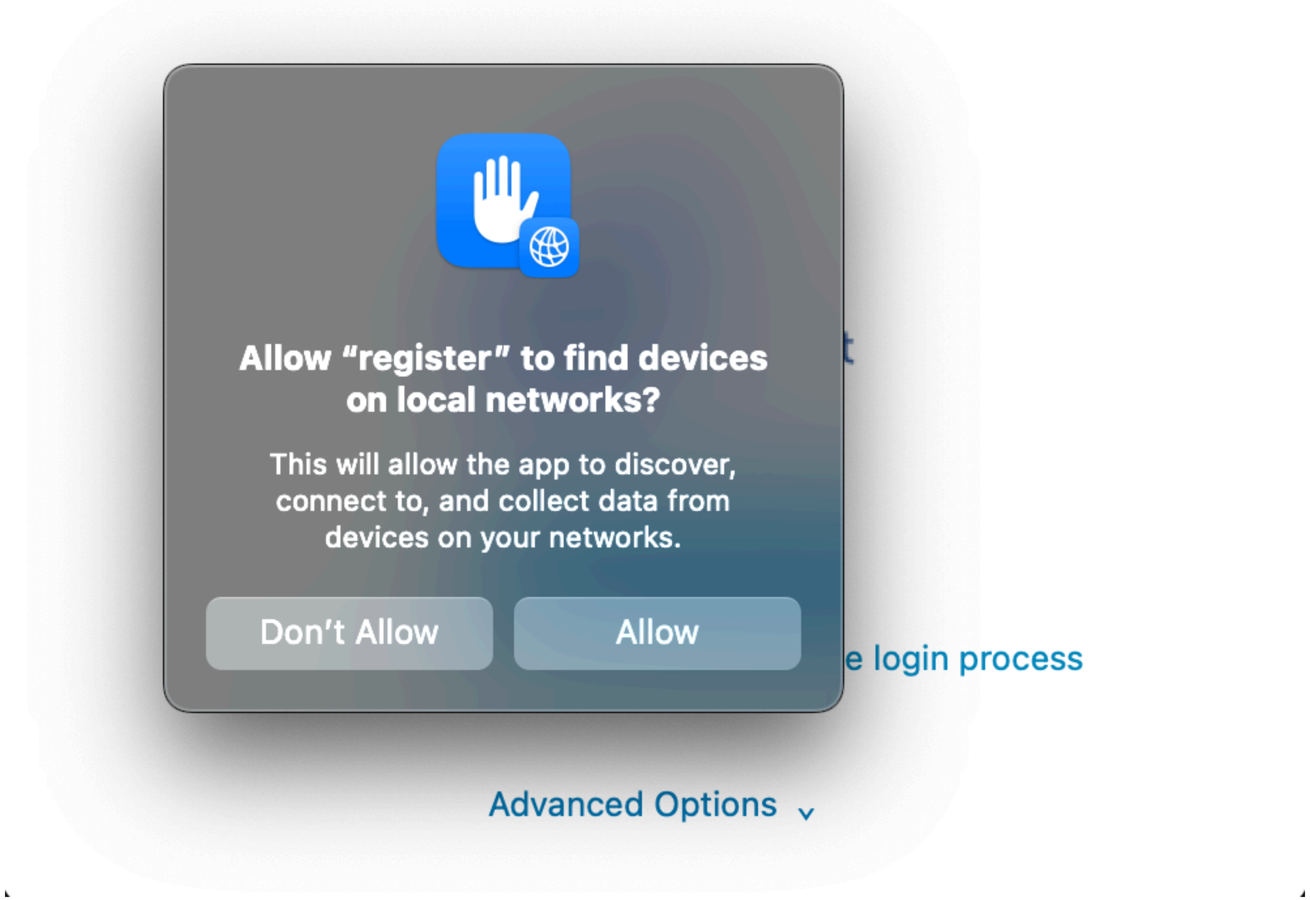
Setting up Globus Local Endpoint (step 1)

Launch Globus desktop client and choose "Log In".



Setting up Globus Local Endpoint (step 2)

Users may need to allow Globus desktop client to find local folders. Select "Allow" and sign onto Globus as shown in earlier slides.

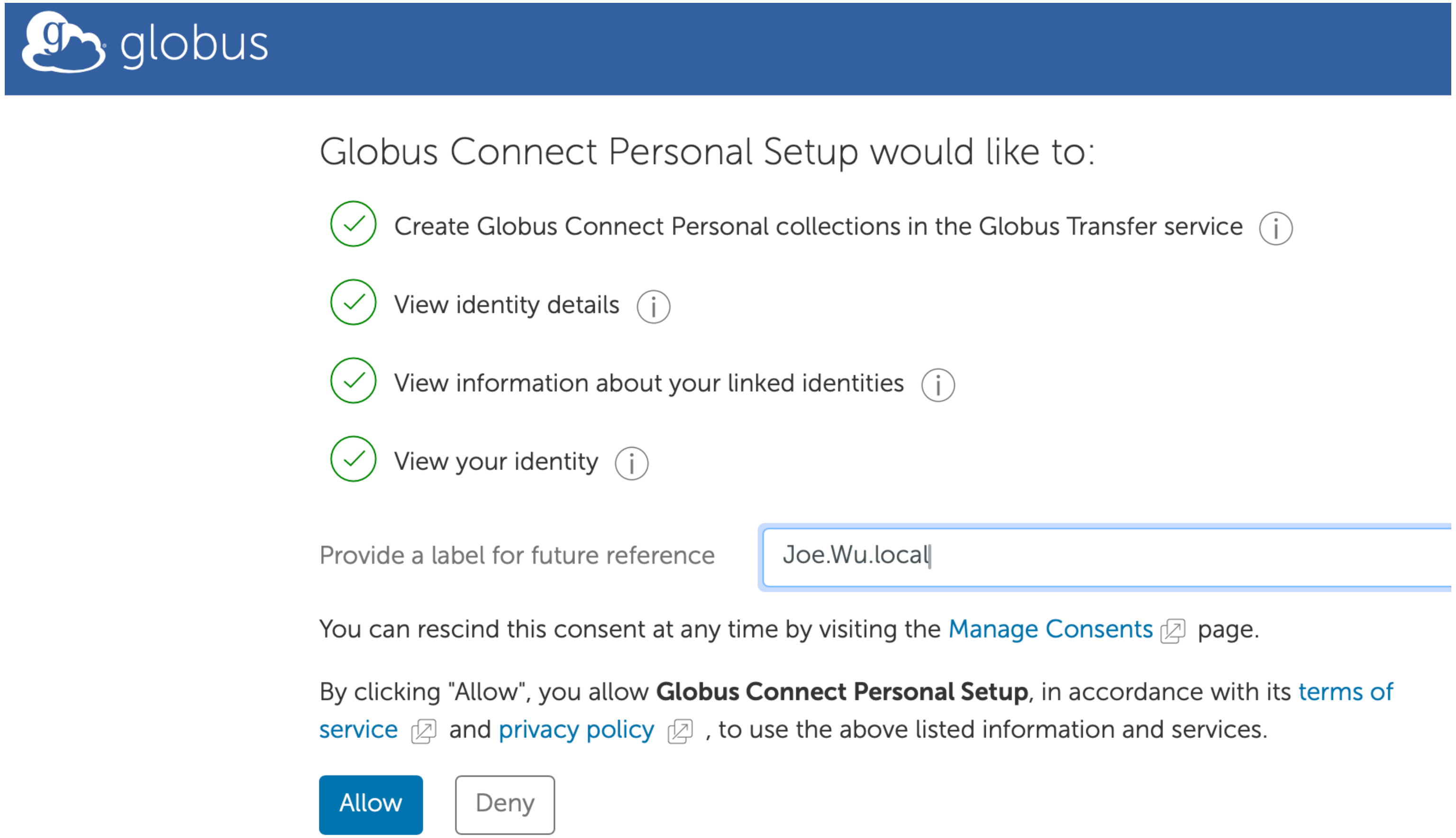


Allow “register” to find devices on local networks?

This will allow the app to discover, connect to, and collect data from devices on your networks.

Setting up Globus Local Endpoint (step 3)

After re-authenticating, fill out a label for the endpoint and click "Allow".



Globus Connect Personal Setup would like to:

- ✓ Create Globus Connect Personal collections in the Globus Transfer service ⓘ
- ✓ View identity details ⓘ
- ✓ View information about your linked identities ⓘ
- ✓ View your identity ⓘ

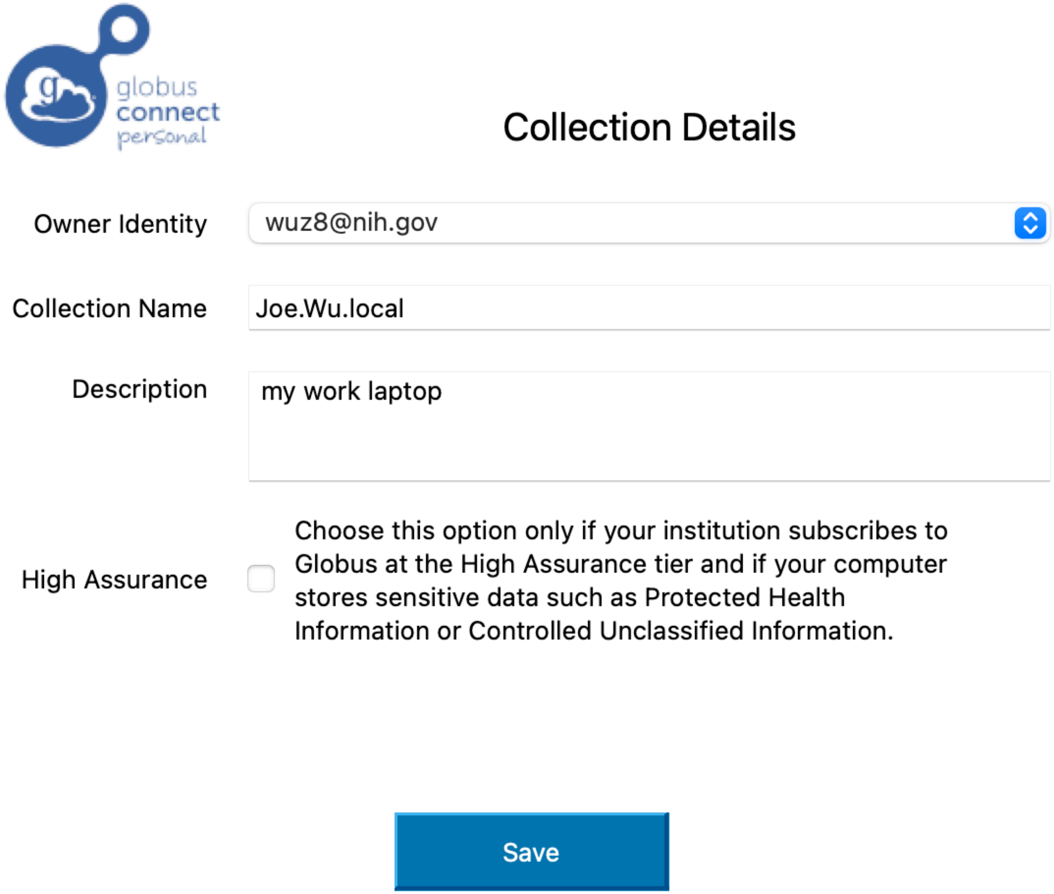
Provide a label for future reference: Joe.Wu.local

You can rescind this consent at any time by visiting the Manage Consents page.

By clicking "Allow", you allow **Globus Connect Personal Setup**, in accordance with its terms of service and privacy policy, to use the above listed information and services.

Setting up Globus Local Endpoint (step 4)

Subsequently, provide the name of the collection as well as a description. Hit "Save" when ready.



Collection Details

Owner Identity: wuz8@nih.gov

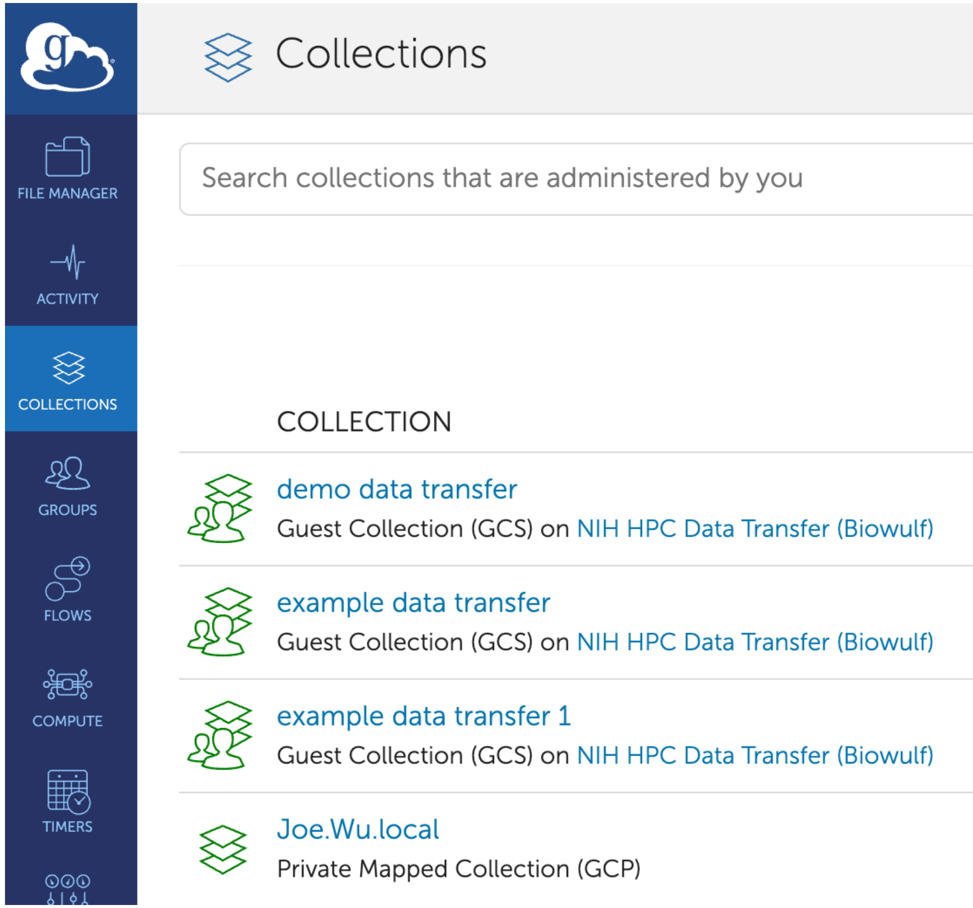
Collection Name: Joe.Wu.local

Description: my work laptop

High Assurance ☐ Choose this option only if your institution subscribes to Globus at the High Assurance tier and if your computer stores sensitive data such as Protected Health Information or Controlled Unclassified Information.

Setting up Globus Local Endpoint (step 5)

The endpoint appears under the "Administered By You" tab in the collections table.

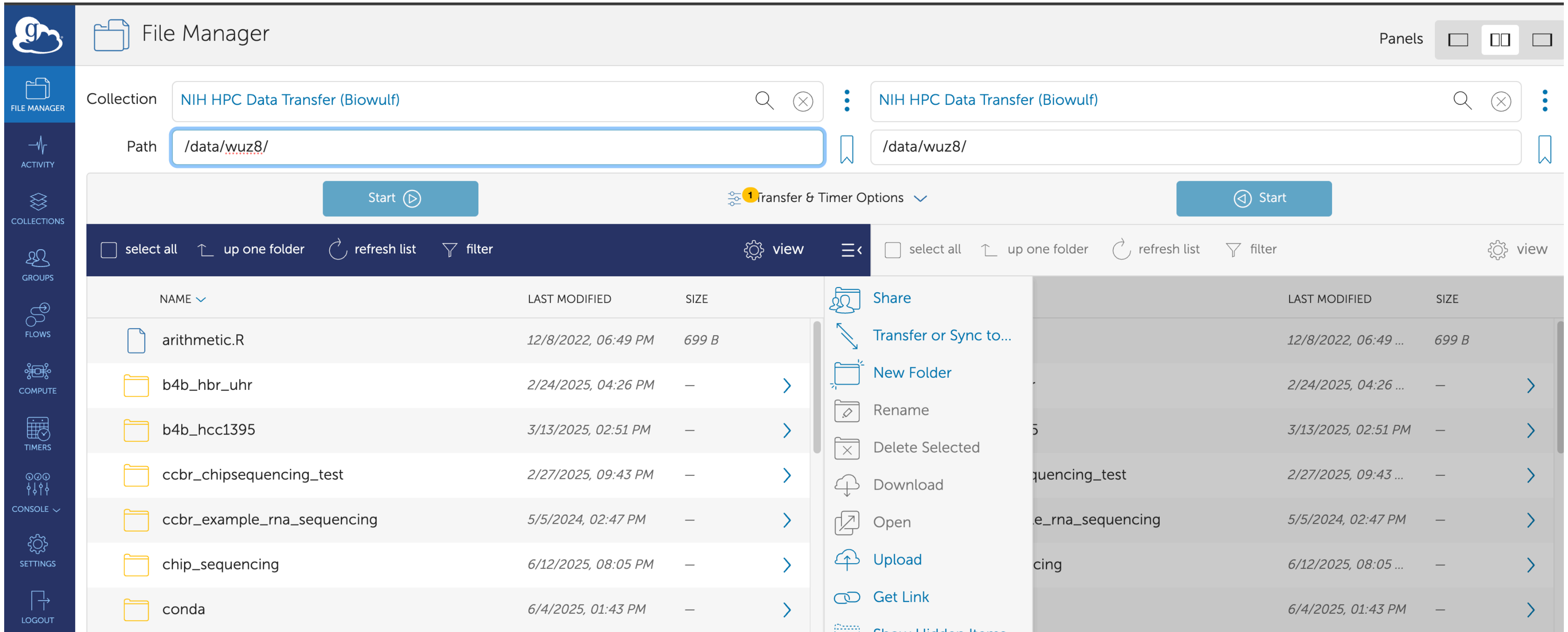


Search collections that are administered by you

COLLECTION
demo data transfer Guest Collection (GCS) on NIH HPC Data Transfer (Biowulf)
example data transfer Guest Collection (GCS) on NIH HPC Data Transfer (Biowulf)
example data transfer 1 Guest Collection (GCS) on NIH HPC Data Transfer (Biowulf)
Joe.Wu.local Private Mapped Collection (GCP)

Data Transfer from Local to Biowulf (step 1)

Click on the "COLLECTIONS" tab and select the "NIH HPC Data Transfer (Biowulf)" endpoint. Select "Transfer or Sync" to start a data transfer. A second file manager window opens. Here, click the magnifying glass to search for the endpoint to transfer to or from. This example will use the local computer endpoint that was setup.



Collection: NIH HPC Data Transfer (Biowulf)

Path: /data/wuz8/

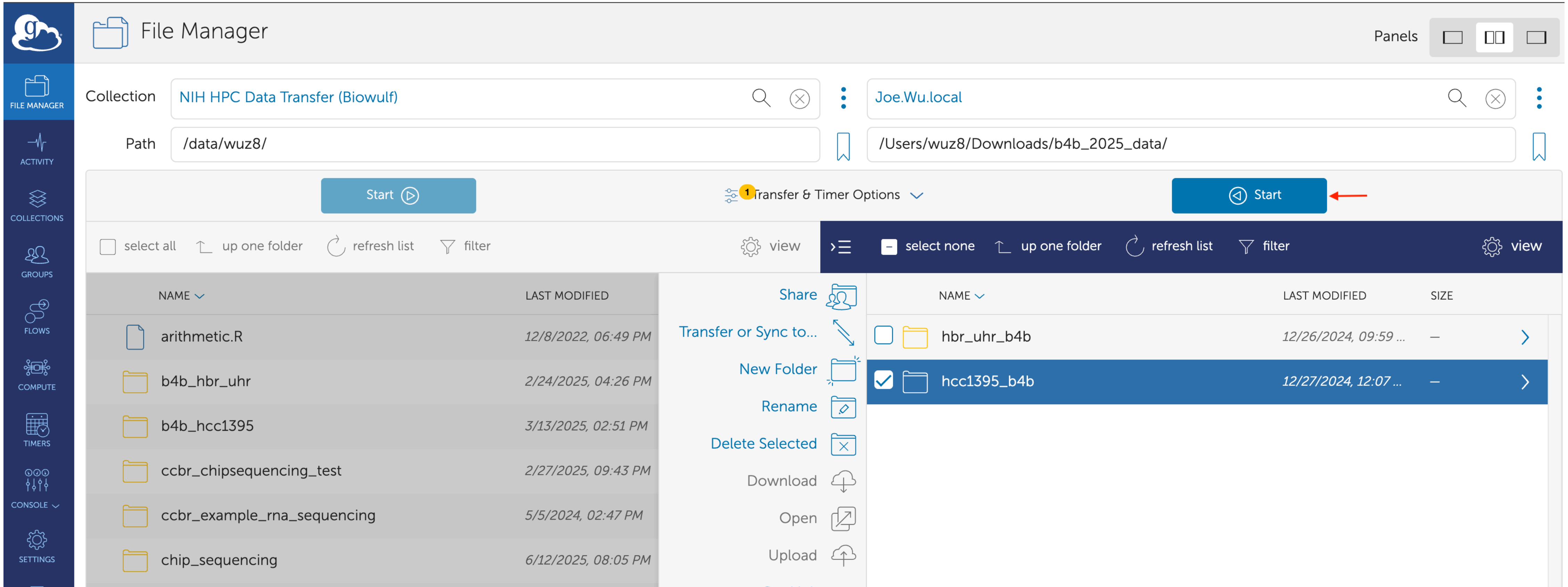
Collection: NIH HPC Data Transfer (Biowulf)

Path: /data/wuz8/

NAME	LAST MODIFIED	SIZE
arithmetic.R	12/8/2022, 06:49 PM	699 B
b4b_hbr_uhr	2/24/2025, 04:26 PM	–
b4b_hcc1395	3/13/2025, 02:51 PM	–
ccbr_chipsequencing_test	2/27/2025, 09:43 PM	–
ccbr_example_rna_sequencing	5/5/2024, 02:47 PM	–
chip_sequencing	6/12/2025, 08:05 PM	–
conda	6/4/2025, 01:43 PM	–

Data Transfer from Local to Biowulf (step 2)

Once the local endpoint ("Joe.Wu.local") is selected type the path to the file or folder that needs to be transferred. Click "Start" when ready. To transfer from Biowulf, just click on "Start" on the "NIH HPC Data Transfer (Biowulf)" endpoint panel.



Collection: NIH HPC Data Transfer (Biowulf)

Path: /data/wuz8/

NAME	LAST MODIFIED
arithmetic.R	12/8/2022, 06:49 PM
b4b_hbr_uhr	2/24/2025, 04:26 PM
b4b_hcc1395	3/13/2025, 02:51 PM
ccbr_chipsequencing_test	2/27/2025, 09:43 PM
ccbr_example_rna_sequencing	5/5/2024, 02:47 PM
chip_sequencing	6/12/2025, 08:05 PM

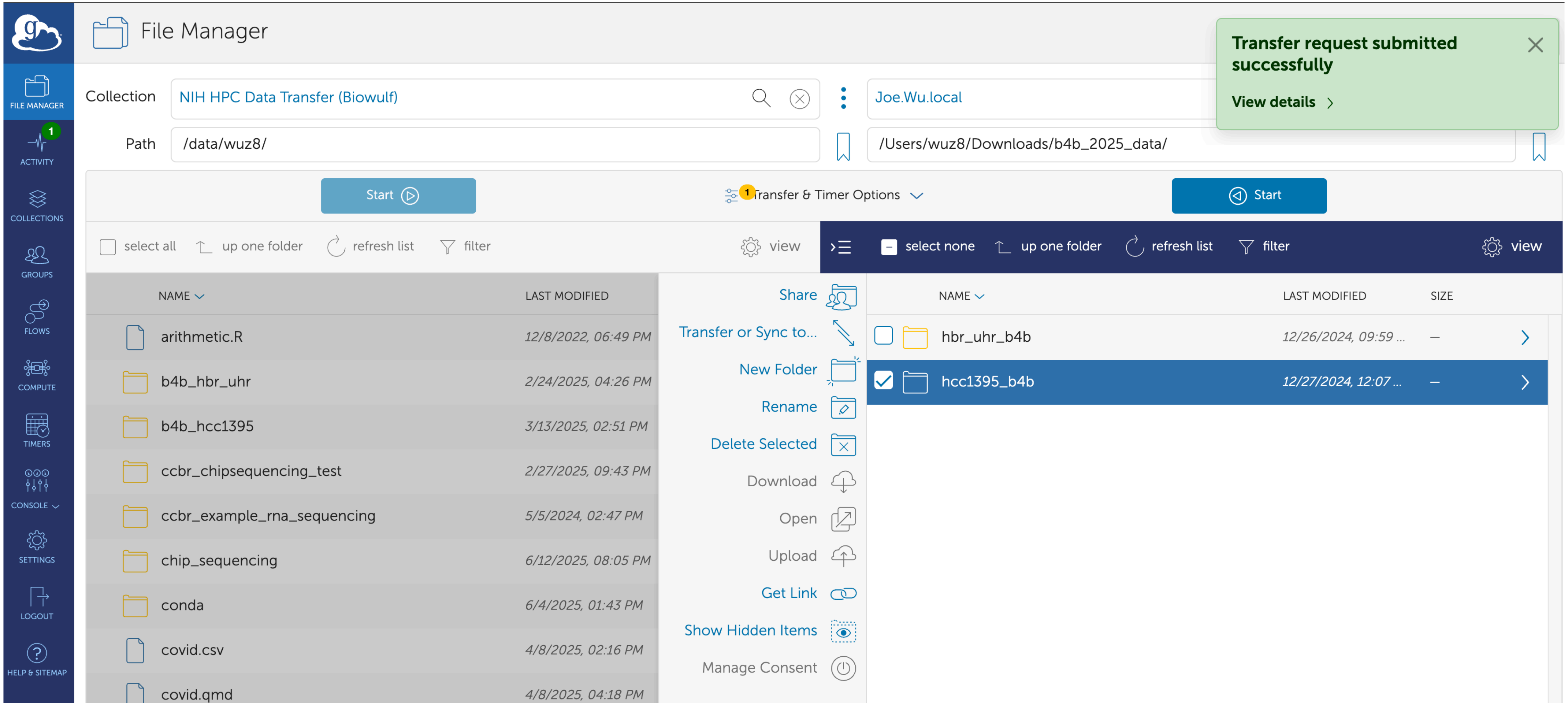
Collection: Joe.Wu.local

Path: /Users/wuz8/Downloads/b4b_2025_data/

	NAME	LAST MODIFIED	SIZE
☐	hbr_uhr_b4b	12/26/2024, 09:59 ...	–
☑	hcc1395_b4b	12/27/2024, 12:07 ...	–

Data Transfer from Local to Biowulf (step 3)

A message will appear if the transfer request was successfully submitted.



Transfer request submitted successfully

View details

Collection: NIH HPC Data Transfer (Biowulf)

Path: /data/wuz8/

Joe.Wu.local

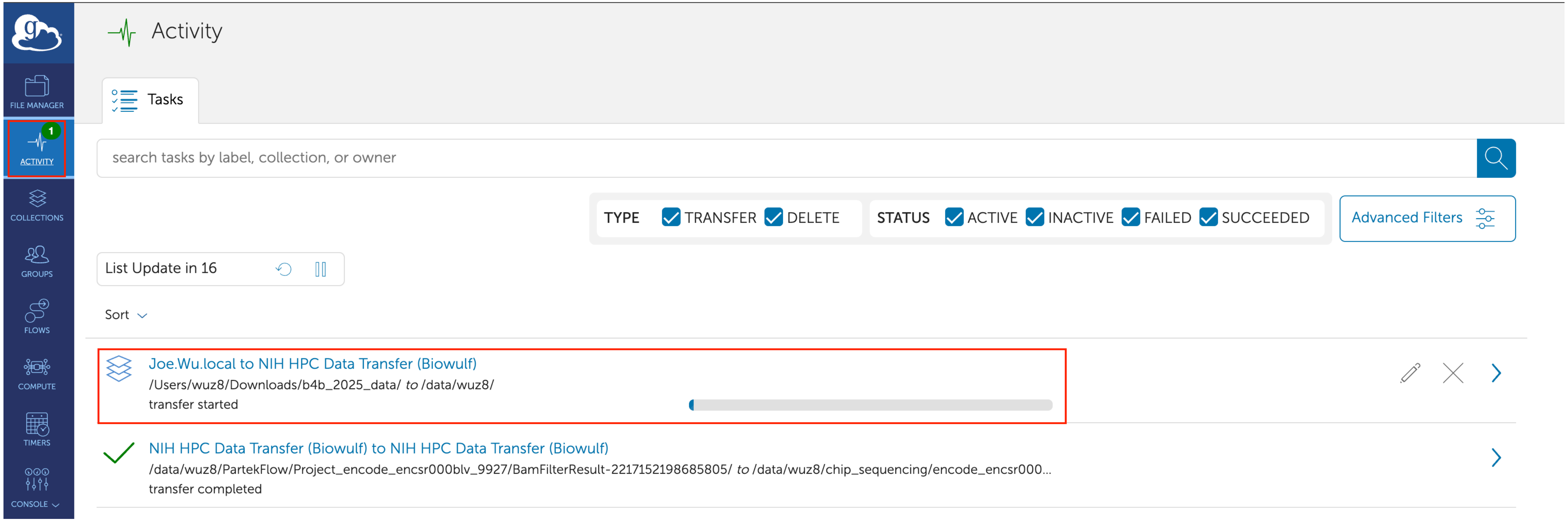
/Users/wuz8/Downloads/b4b_2025_data/

NAME	LAST MODIFIED
arithmetic.R	12/8/2022, 06:49 PM
b4b_hbr_uhr	2/24/2025, 04:26 PM
b4b_hcc1395	3/13/2025, 02:51 PM
ccbr_chipsequencing_test	2/27/2025, 09:43 PM
ccbr_example_rna_sequencing	5/5/2024, 02:47 PM
chip_sequencing	6/12/2025, 08:05 PM
conda	6/4/2025, 01:43 PM
covid.csv	4/8/2025, 02:16 PM
covid.qmd	4/8/2025, 04:18 PM

	NAME	LAST MODIFIED	SIZE
☐	hbr_uhr_b4b	12/26/2024, 09:59 ...	–
☑	hcc1395_b4b	12/27/2024, 12:07 ...	–

Data Transfer from Local to Biowulf (step 4)

Click on "ACTIVITY" to view details such as transfer progress. Once the transfer is complete, users will get an email.



TYPE ☑ TRANSFER ☑ DELETE

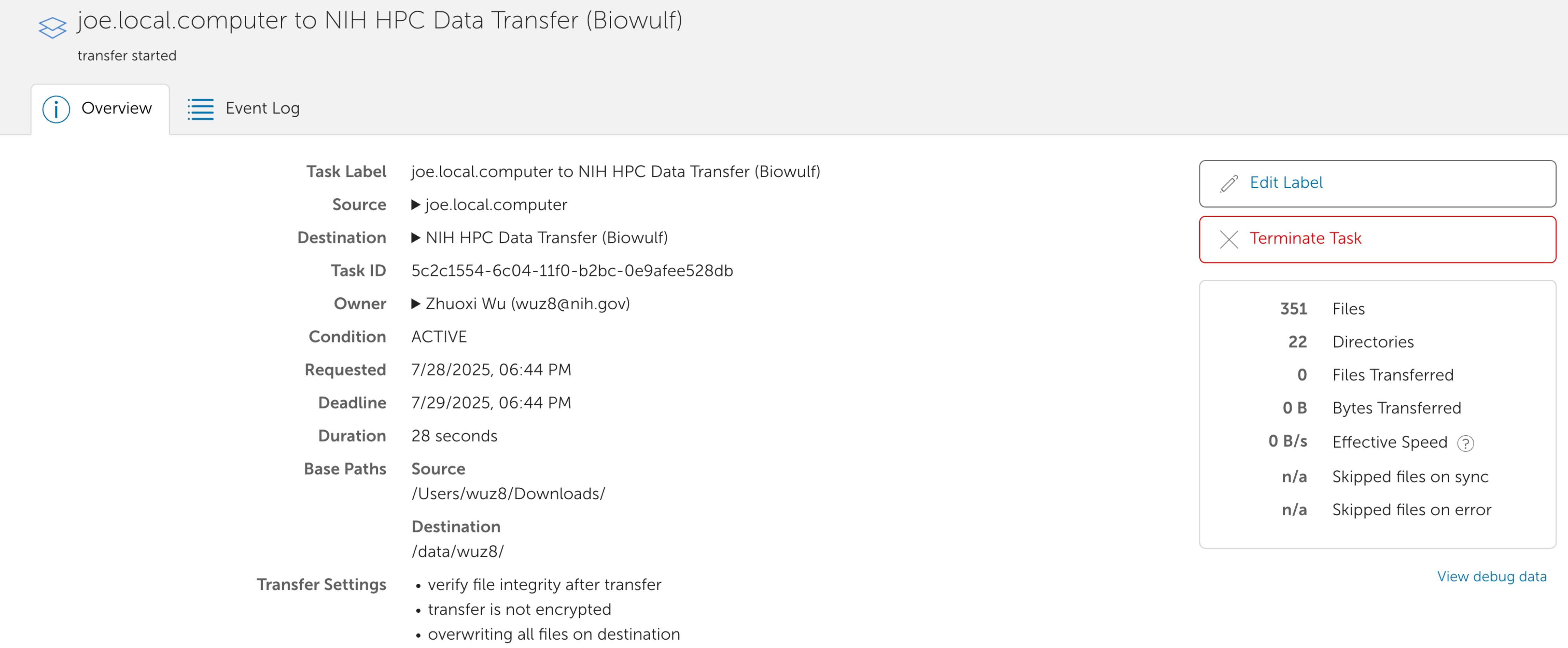
STATUS ☑ ACTIVE ☑ INACTIVE ☑ FAILED ☑ SUCCEEDED

Joe.Wu.local to NIH HPC Data Transfer (Biowulf)
/Users/wuz8/Downloads/b4b_2025_data/ *to* /data/wuz8/
transfer started

✓ NIH HPC Data Transfer (Biowulf) to NIH HPC Data Transfer (Biowulf)
/data/wuz8/PartekFlow/Project_encode_encsr000blv_9927/BamFilterResult-2217152198685805/ *to* /data/wuz8/chip_sequencing/encode_encsr000...
transfer completed

Terminating a Transfer

In the activity monitor, users can select a transfer task and subsequently terminate it if needed.



joe.local.computer to NIH HPC Data Transfer (Biowulf)

transfer started

Task Label joe.local.computer to NIH HPC Data Transfer (Biowulf)

Source ▶ joe.local.computer

Destination ▶ NIH HPC Data Transfer (Biowulf)

Task ID 5c2c1554-6c04-11f0-b2bc-0e9afee528db

Owner ▶ Zhuoxi Wu (wuz8@nih.gov)

Condition ACTIVE

Requested 7/28/2025, 06:44 PM

Deadline 7/29/2025, 06:44 PM

Duration 28 seconds

Base Paths **Source**
/Users/wuz8/Downloads/

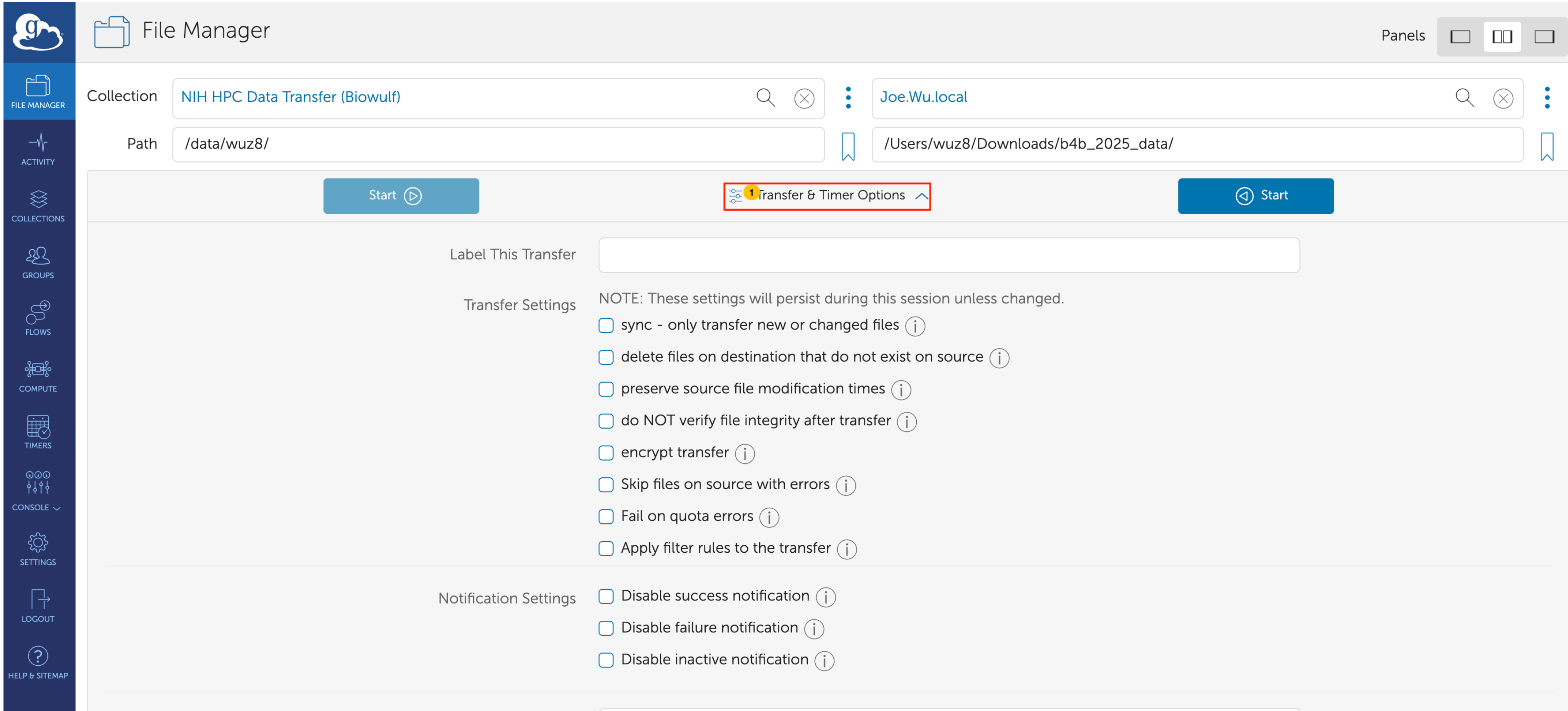
Destination
/data/wuz8/

Transfer Settings
- verify file integrity after transfer
- transfer is not encrypted
- overwriting all files on destination

351	Files
22	Directories
0	Files Transferred
0 B	Bytes Transferred
0 B/s	Effective Speed ?
n/a	Skipped files on sync
n/a	Skipped files on error

Schedule Data Transfer

Users can schedule data transfer.



Collection: NIH HPC Data Transfer (Biowulf)

Path: /data/wuz8/

Collection: Joe.Wu.local

Path: /Users/wuz8/Downloads/b4b_2025_data/

Label This Transfer

Transfer Settings

NOTE: These settings will persist during this session unless changed.

☐ sync - only transfer new or changed files

☐ delete files on destination that do not exist on source

☐ preserve source file modification times

☐ do NOT verify file integrity after transfer

☐ encrypt transfer

☐ Skip files on source with errors

☐ Fail on quota errors

☐ Apply filter rules to the transfer

Notification Settings

☐ Disable success notification

☐ Disable failure notification

☐ Disable inactive notification

Transfer from NCI CCR Sequencing Facility Data Management Environment: Overview

This example applies to those researchers who utilize the NCI CCR Sequencing Facility for sequencing experiments. The sequencing facility will:

- Provide a link to their Data Management Environment (DME) for researchers to access their data.
- Sequencing facility will also do many of the analysis steps for the researchers including QC.

Please check with the specific core for data management and transfer issues if not using NCI CCR Sequencing Facility.

Make a New Folder in Globus

Open the "NIH HPC Data Transfer (Biowulf)" endpoint and create a folder for Globus transfers in the `data`. In the example, below, the folder is named "globus_transfers".

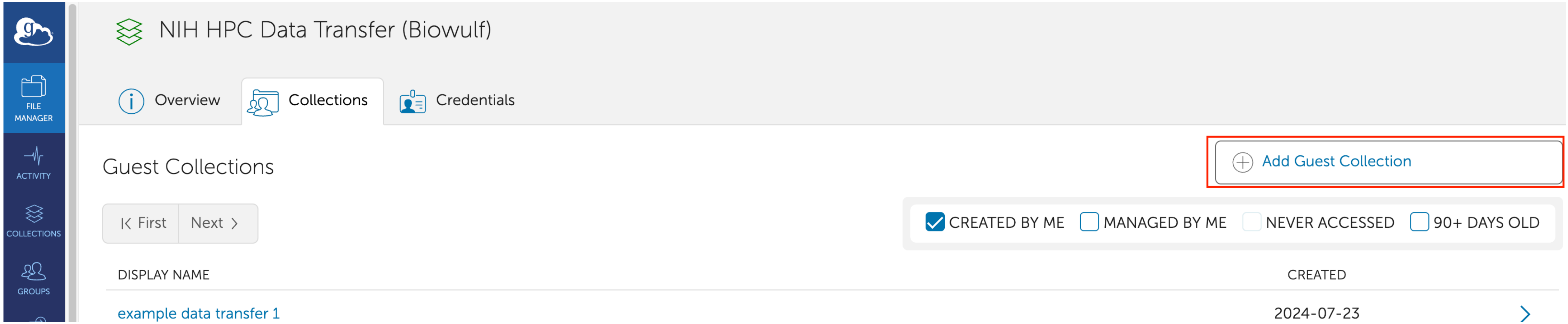
Collection: NIH HPC Data Transfer (Biowulf)

Path: /data/wuz8/

NAME	LAST MODIFIED	SIZE
globus_transfers	6/25/2025, 02:31 PM	–
hcc1395_b4b	6/25/2025, 10:33 AM	–

Add Guest Collection

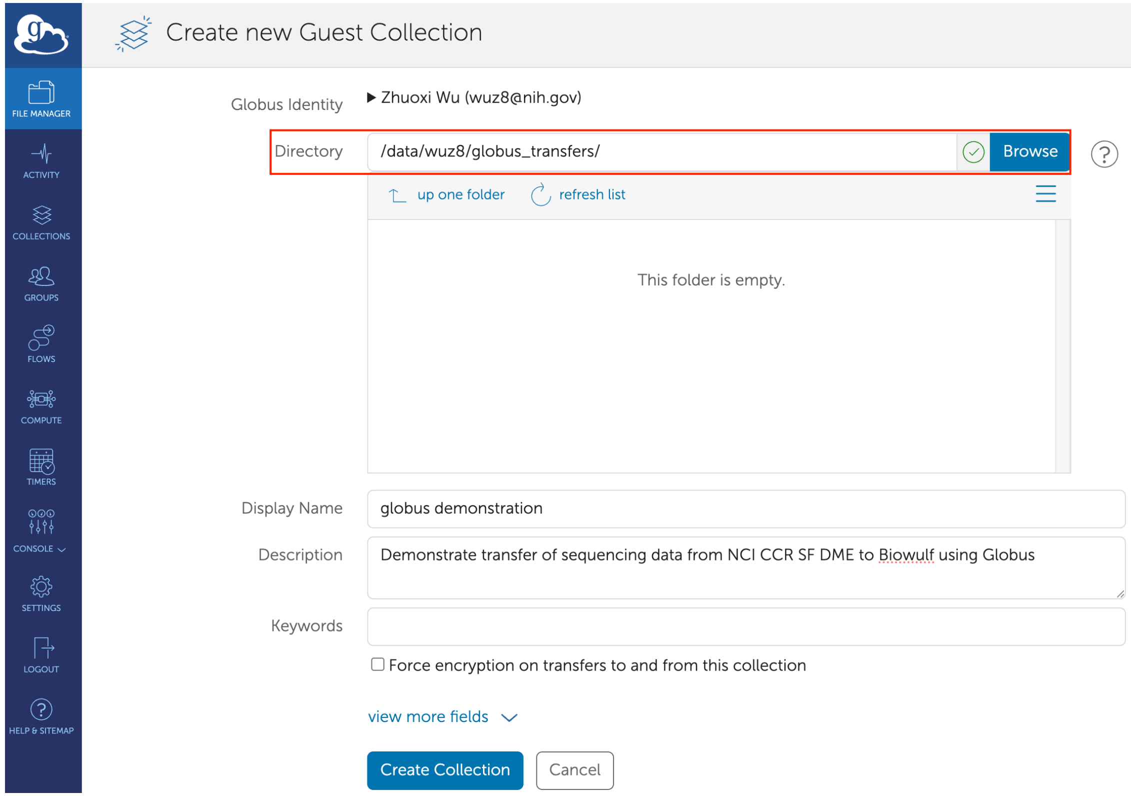
Next, go back to the "NIH HPC Data Transfer (Biowulf)" endpoint and click on "Add Guest Collection".



DISPLAY NAME	CREATED
example data transfer 1	2024-07-23

Provide Guest Collection Information

In the subsequent page, supply the directory in which to link the guest collection to. As an example, `globus_transfer` under Biowulf `data` folder. Provide a display name and description for the guest collection and hit "Create Collection" when done.



Globus Identity ▶ Zhuoxi Wu (wuz8@nih.gov)

Directory /data/wuz8/globus_transfers/

Display Name globus demonstration

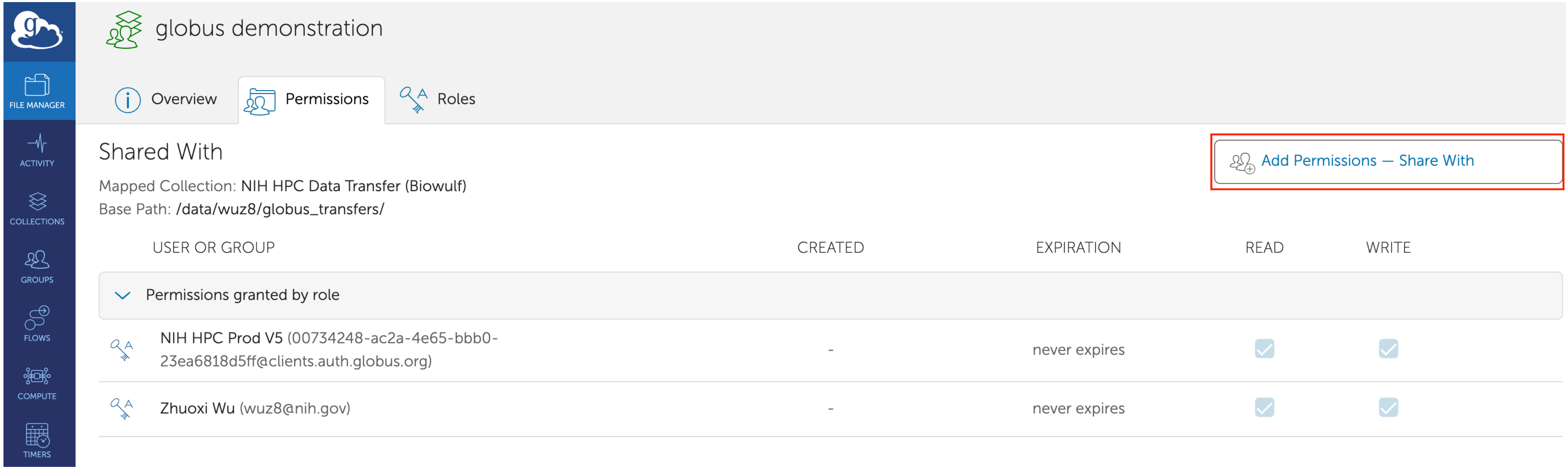
Description Demonstrate transfer of sequencing data from NCI CCR SF DME to Biowulf using Globus

Keywords

☐ Force encryption on transfers to and from this collection

Granting Permission for NCI CCR SF DME to Transfer to Guest Collection (step 1)

Next, grant permission for the NCI CCR Sequencing Facility Data Management Environment to share data with the globus demonstration collection.



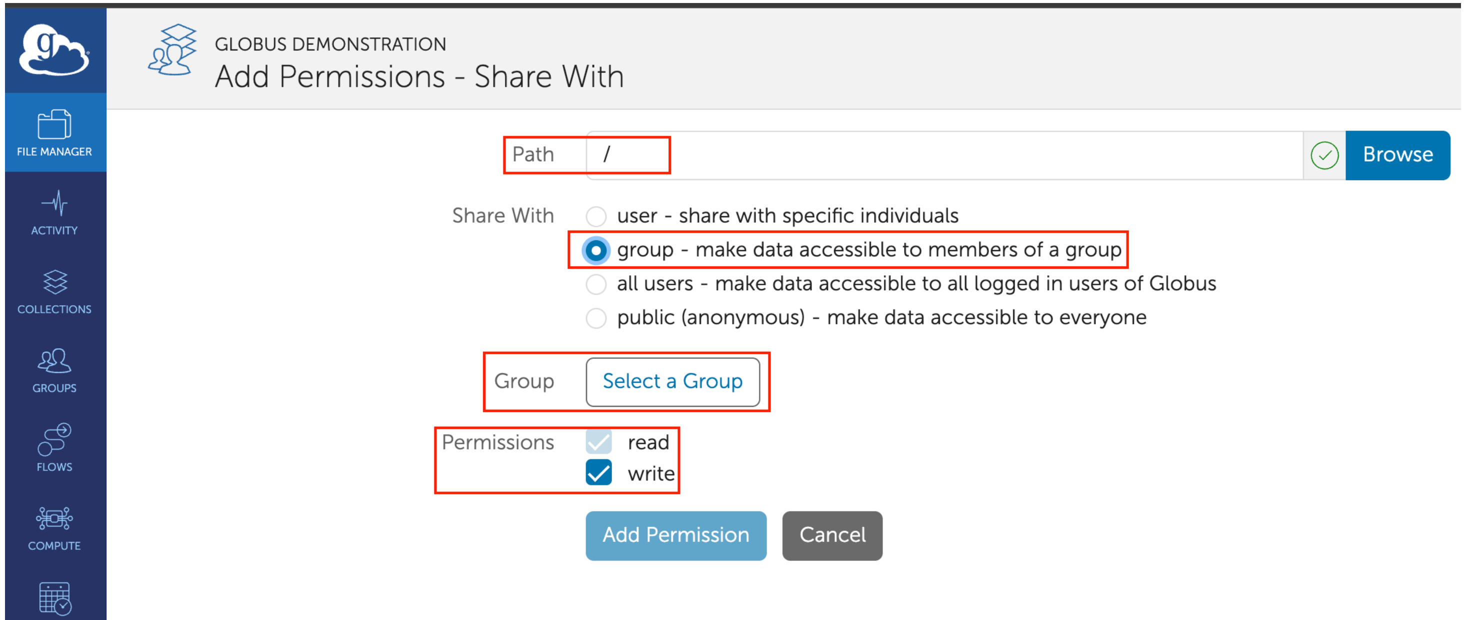
Shared With

Mapped Collection: NIH HPC Data Transfer (Biowulf)
Base Path: /data/wuz8/globus_transfers/

USER OR GROUP	CREATED	EXPIRATION	READ	WRITE
Permissions granted by role				
NIH HPC Prod V5 (00734248-ac2a-4e65-bbb0-23ea6818d5ff@clients.auth.globus.org)	-	never expires	☑	☑
Zhuoxi Wu (wuz8@nih.gov)	-	never expires	☑	☑

Granting Permission for NCI CCR SF DME to Transfer to Guest Collection (step 2)

- Keep `/` in the Path box.
- Be sure to select share with group.
- Make sure that permissions are set to read and write.
- Click "Select a Group" to find the group to grant permission to this endpoint to.



GLOBUS DEMONSTRATION

Add Permissions - Share With

Path: /

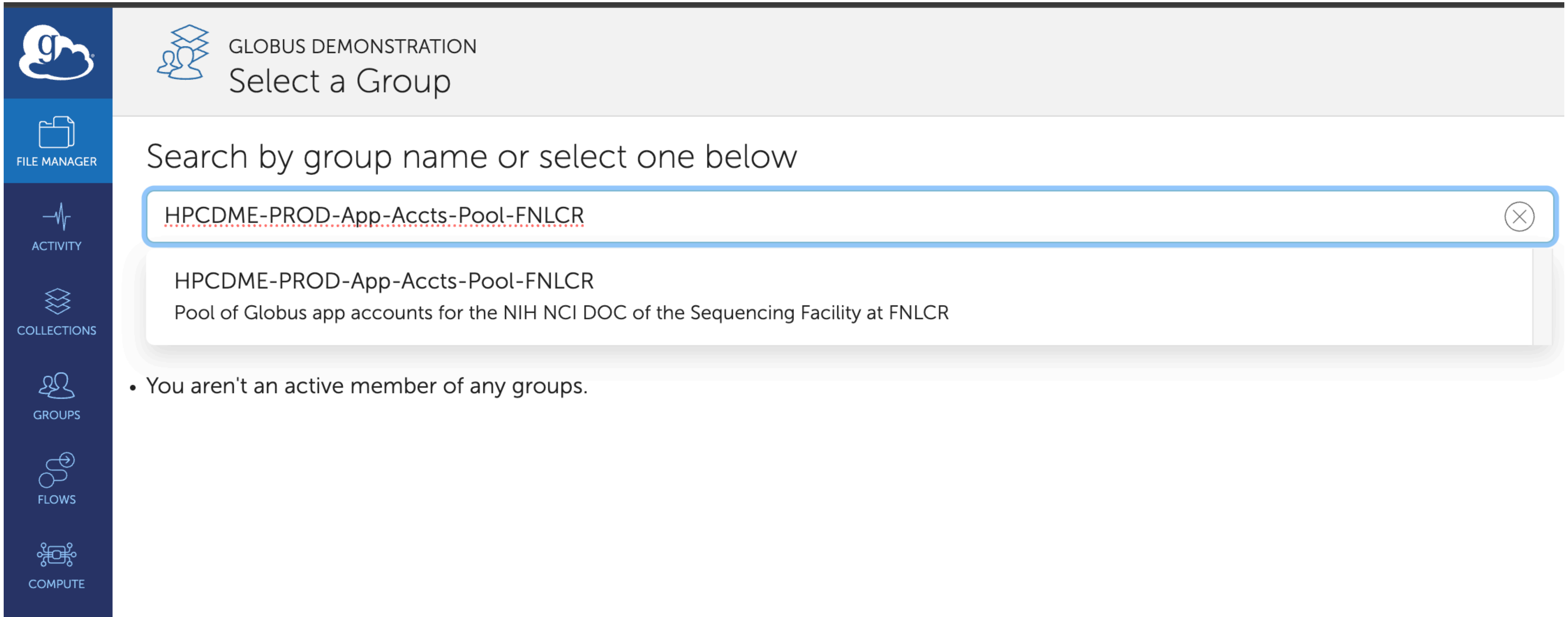
Share With
- ○ user - share with specific individuals
- ● group - make data accessible to members of a group
- ○ all users - make data accessible to all logged in users of Globus
- ○ public (anonymous) - make data accessible to everyone

Group: Select a Group

Permissions
- ☑ read
- ☑ write

Granting Permission for NCI CCR SF DME to Transfer to Guest Collection (step 3)

Enter the name of the group in which to grant permission for the collection to or start typing and a list of options will appear for users to choose from.



GLOBUS DEMONSTRATION

Select a Group

Search by group name or select one below

HPCDME-PROD-App-Accts-Pool-FNLCR

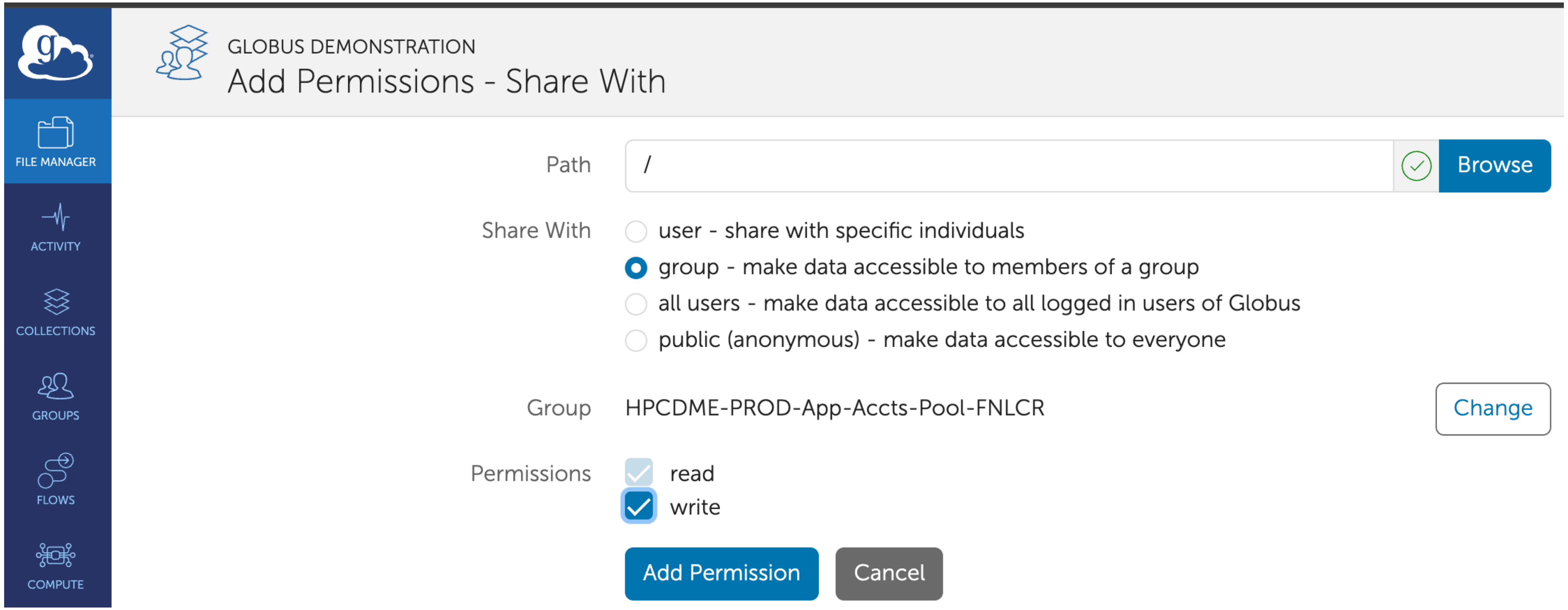
HPCDME-PROD-App-Accts-Pool-FNLCR

Pool of Globus app accounts for the NIH NCI DOC of the Sequencing Facility at FNLCR

- You aren't an active member of any groups.

Granting Permission for NCI CCR SF DME to Transfer to Guest Collection (step 4)

The group that is getting permission granted is now listed next to the "Group" column. Click "Add Permission" when ready.



GLOBUS DEMONSTRATION

Add Permissions - Share With

Path: /

Share With:

○ user - share with specific individuals

● group - make data accessible to members of a group

○ all users - make data accessible to all logged in users of Globus

○ public (anonymous) - make data accessible to everyone

Group: HPCDME-PROD-App-Accts-Pool-FNLCR

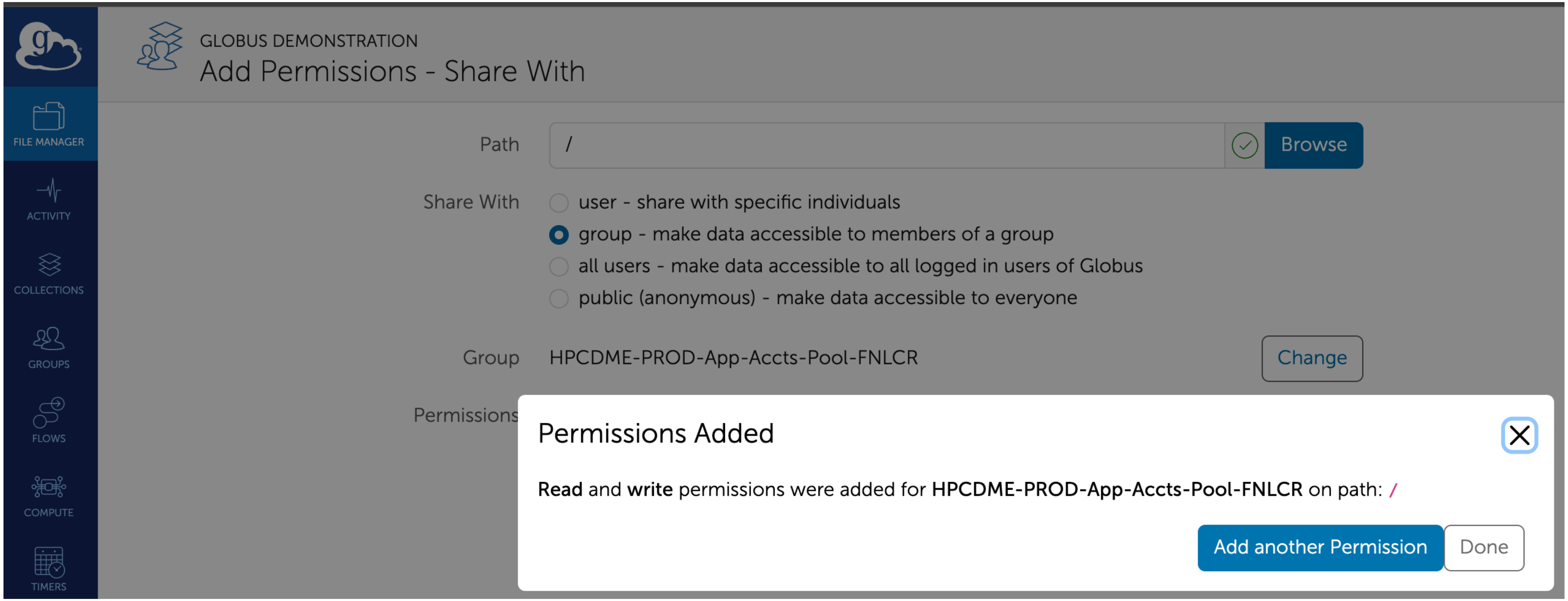
Permissions:

☑ read

☑ write

Granting Permission for NCI CCR SF DME to Transfer to Guest Collection (step 5)

When prompted, click "Done" to complete the permission granting process.



GLOBUS DEMONSTRATION

Add Permissions - Share With

Path: /

Share With:
○ user - share with specific individuals
● group - make data accessible to members of a group
○ all users - make data accessible to all logged in users of Globus
○ public (anonymous) - make data accessible to everyone

Group: HPCDME-PROD-App-Accts-Pool-FNLCR

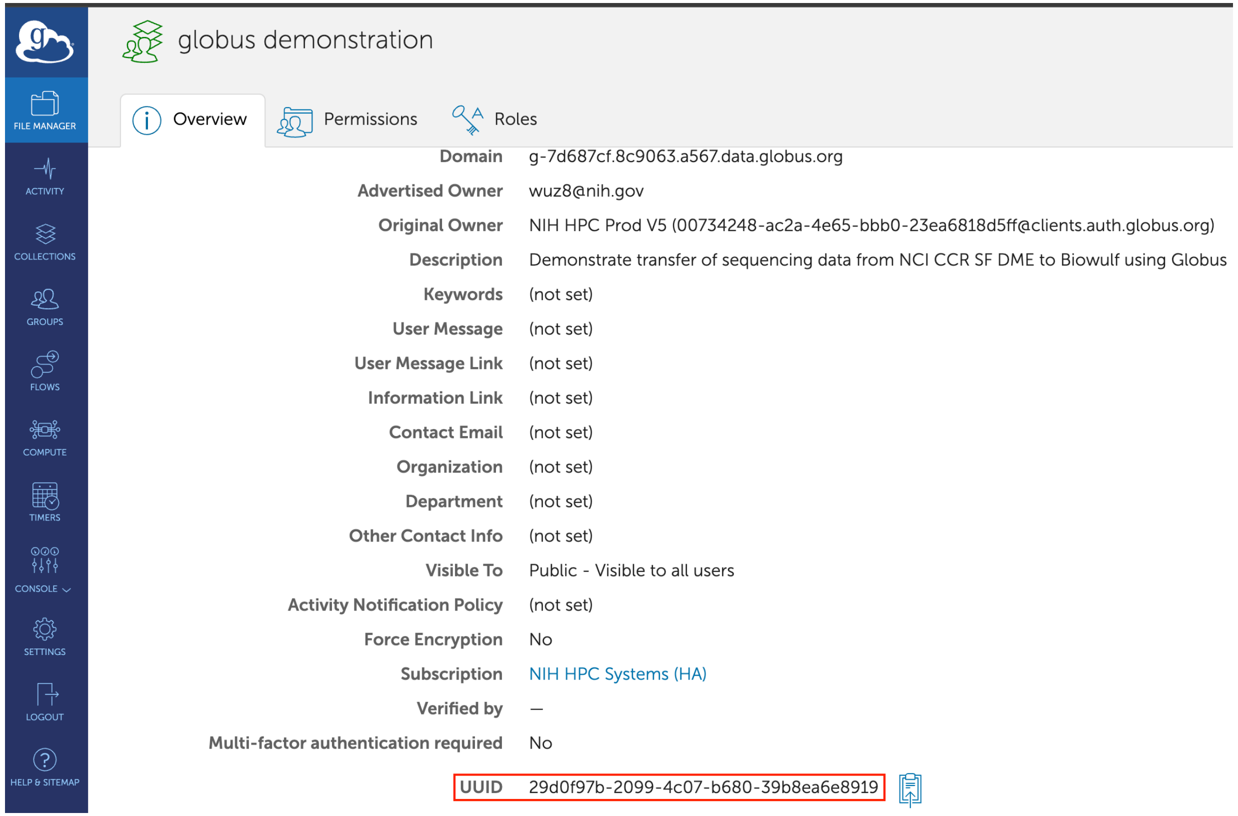
Permissions

Permissions Added

Read and **write** permissions were added for **HPCDME-PROD-App-Accts-Pool-FNLCR** on path: /

Granting Permission for NCI CCR SF DME to Transfer to Guest Collection (step 6)

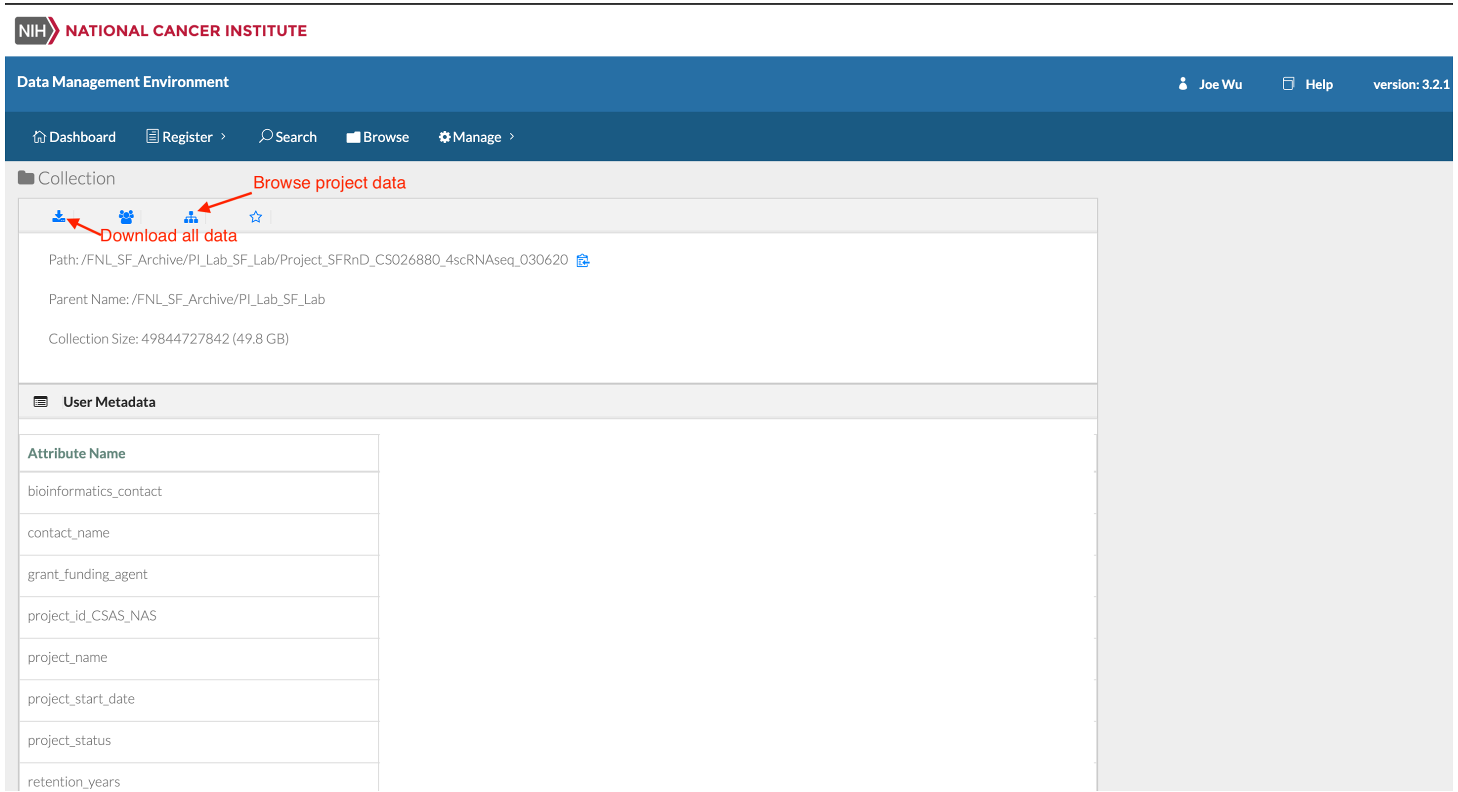
Look at the "Overview" for the "globus demonstration" collection. Note the UUID. This will be needed when transferring file from the sequencing facility DME to Biowulf.



Domain	g-7d687cf.8c9063.a567.data.globus.org
Advertised Owner	wuz8@nih.gov
Original Owner	NIH HPC Prod V5 (00734248-ac2a-4e65-bbb0-23ea6818d5ff@clients.auth.globus.org)
Description	Demonstrate transfer of sequencing data from NCI CCR SF DME to Biowulf using Globus
Keywords	(not set)
User Message	(not set)
User Message Link	(not set)
Information Link	(not set)
Contact Email	(not set)
Organization	(not set)
Department	(not set)
Other Contact Info	(not set)
Visible To	Public - Visible to all users
Activity Notification Policy	(not set)
Force Encryption	No
Subscription	NIH HPC Systems (HA)
Verified by	—
Multi-factor authentication required	No
UUID	29d0f97b-2099-4c07-b680-39b8ea6e8919

Getting Data from Sequencing Facility DME (step 1)

In the sequencing facility DME page where data is stored, click on "Browse project data" to peruse specific data or download the all of the data.



Browse project data

Download all data

Path: /FNL_SF_Archive/PI_Lab_SF_Lab/Project_SFRnD_CS026880_4scRNAseq_030620

Parent Name: /FNL_SF_Archive/PI_Lab_SF_Lab

Collection Size: 49844727842 (49.8 GB)

Attribute Name
bioinformatics_contact
contact_name
grant_funding_agent
project_id_CSAS_NAS
project_name
project_start_date
project_status
retention_years

Getting Data from Sequencing Facility DME (step 2)

This example will browse for specific data to download. Just click on the "Download" button to the far right of the file content table when ready.

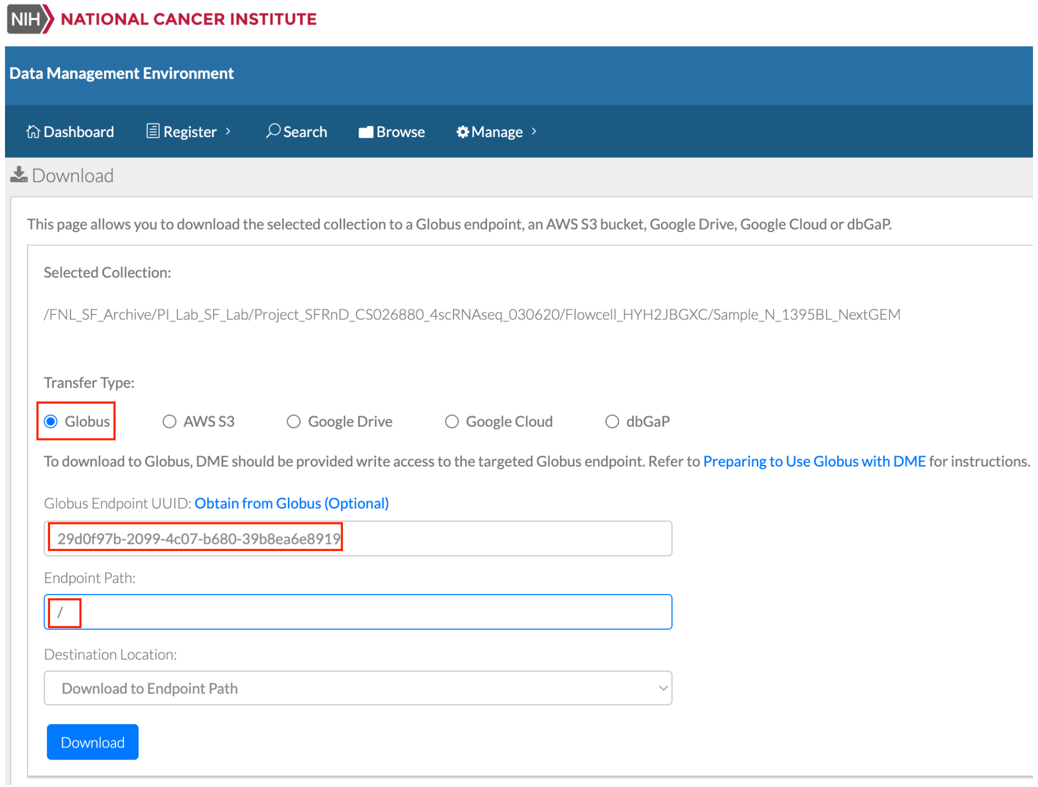
- FNL_SF_Archive
  - PI_Lab_SF_Lab
    - Project_SFRnD_CS026880_4scRNAseq_030620
      - Flowcell_HYH2JBGXC
        - AggregatedDatasets_aggregate.tar
        - N_1395BL_NextGEM_count.tar
        - T_1395_NextGEM_count.tar
        - T_N_NextGEM_Hash_GEX_A10_count.tar
- User Bookmarks

FNL_SF_Archive / PI_Lab_SF_Lab / Project_SFRnD_CS026880_4scRNAseq_030620 / Flowcell_HYH2JBGXC Total Size: 26798532642 (26.8 GB) Number of Entries: 5/5

File Name	File Size	Last Updated	Download
Sample_N_1395BL_NextGEM		2020-09-03 0...	⤓
Sample_T_1395_NextGEM		2020-09-03 0...	⤓
Sample_T_N_NextGEM_Hash_GEX_A10		2020-09-03 0...	⤓
Undetermined		2020-09-03 0...	⤓
200710_NB501156_0506_AHYH2JBGXC_supplem...	27525120	2020-09-03 1...	⤓

Getting Data from Sequencing Facility DME (step 3)

In the next page, select the Globus radial button under "Transfer Type" and supply the Globus endpoint UUID.



This page allows you to download the selected collection to a Globus endpoint, an AWS S3 bucket, Google Drive, Google Cloud or dbGaP.

Selected Collection:

/FNL_SF_Archive/PI_Lab_SF_Lab/Project_SFRnD_CS026880_4scRNAseq_030620/Flowcell_HYH2JBGXC/Sample_N_1395BL_NextGEM

Transfer Type:

● Globus ○ AWS S3 ○ Google Drive ○ Google Cloud ○ dbGaP

To download to Globus, DME should be provided write access to the targeted Globus endpoint. Refer to Preparing to Use Globus with DME for instructions.

Globus Endpoint UUID: Obtain from Globus (Optional)

29d0f97b-2099-4c07-b680-39b8ea6e8919

Endpoint Path:

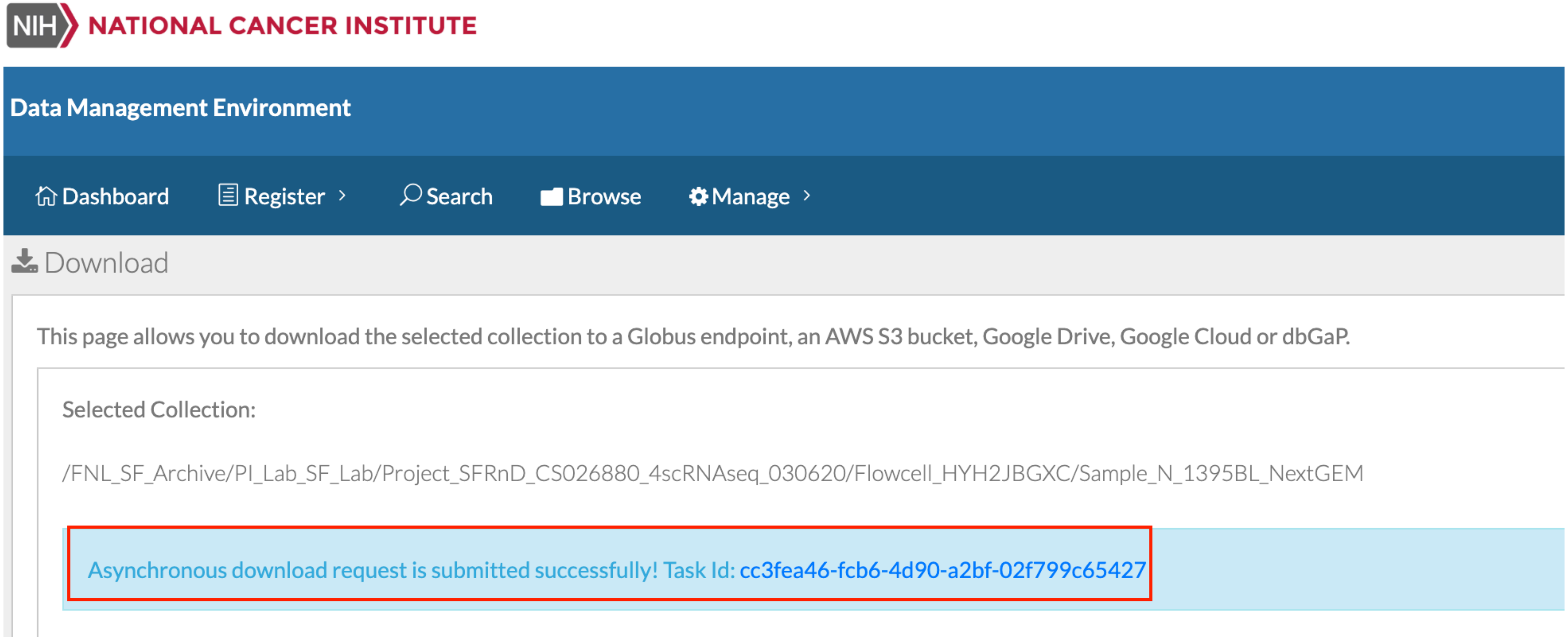
/

Destination Location:

Download to Endpoint Path

Getting Data from Sequencing Facility DME (step 4)

Users will see the message highlighted in blue in the image below when the transfer request has been successfully submitted.



This page allows you to download the selected collection to a Globus endpoint, an AWS S3 bucket, Google Drive, Google Cloud or dbGaP.

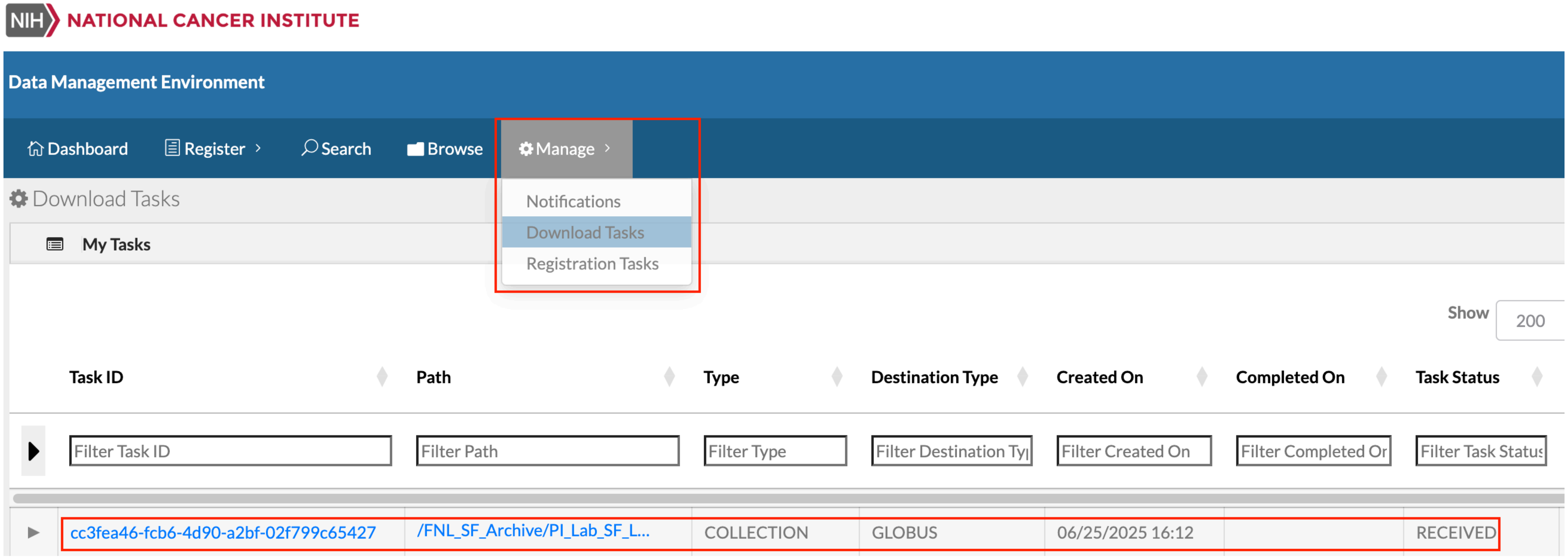
Selected Collection:

/FNL_SF_Archive/PI_Lab_SF_Lab/Project_SFRnD_CS026880_4scRNAseq_030620/Flowcell_HYH2JBGXC/Sample_N_1395BL_NextGEM

Asynchronous download request is submitted successfully! Task Id: cc3fea46-fcb6-4d90-a2bf-02f799c65427

Getting Data from Sequencing Facility DME (step 5)

Clicking on "Manage" and then "Download Tasks", the status for the data transfer will change to complete when done.



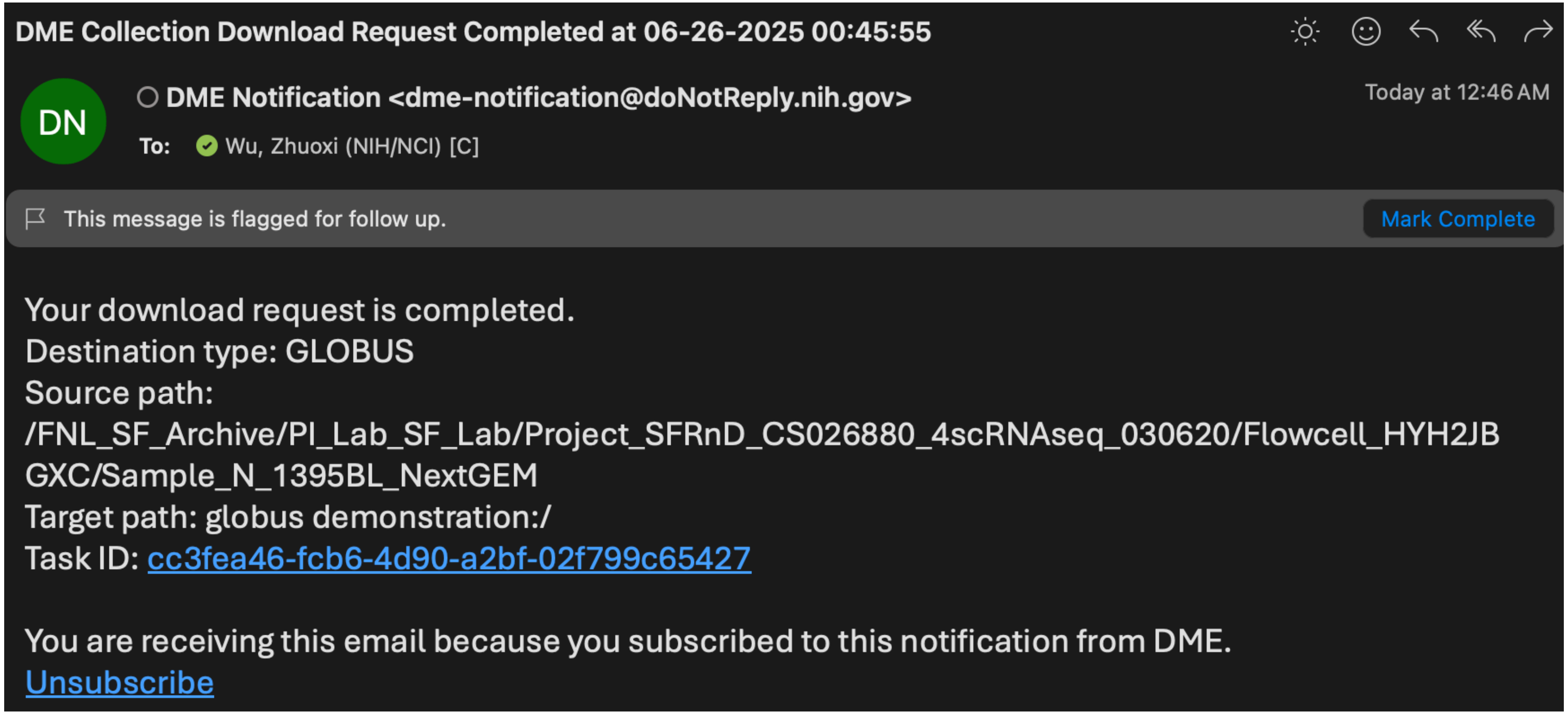
Task ID	Path	Type	Destination Type	Created On	Completed On	Task Status
cc3fea46-fcb6-4d90-a2bf-02f799c65427	/FNL_SF_Archive/PI_Lab_SF_L...	COLLECTION	GLOBUS	06/25/2025 16:12		RECEIVED

Getting Data from Sequencing Facility DME - task status

If a task status of "RESTORE_REQUESTED" appears, then it is likely that the dataset has been placed in archive and DME has to download it from a cloud service prior to transferring to Biowulf. This process may take approximately 12 hours.

Getting Data from Sequencing Facility DME (step 6)

Users will also receive an email from sequencing facility DME that transfer was completed.



DME Collection Download Request Completed at 06-26-2025 00:45:55

○ **DME Notification <dme-notification@doNotReply.nih.gov>**

To: Wu, Zhuoxi (NIH/NCI) [C]

Today at 12:46 AM

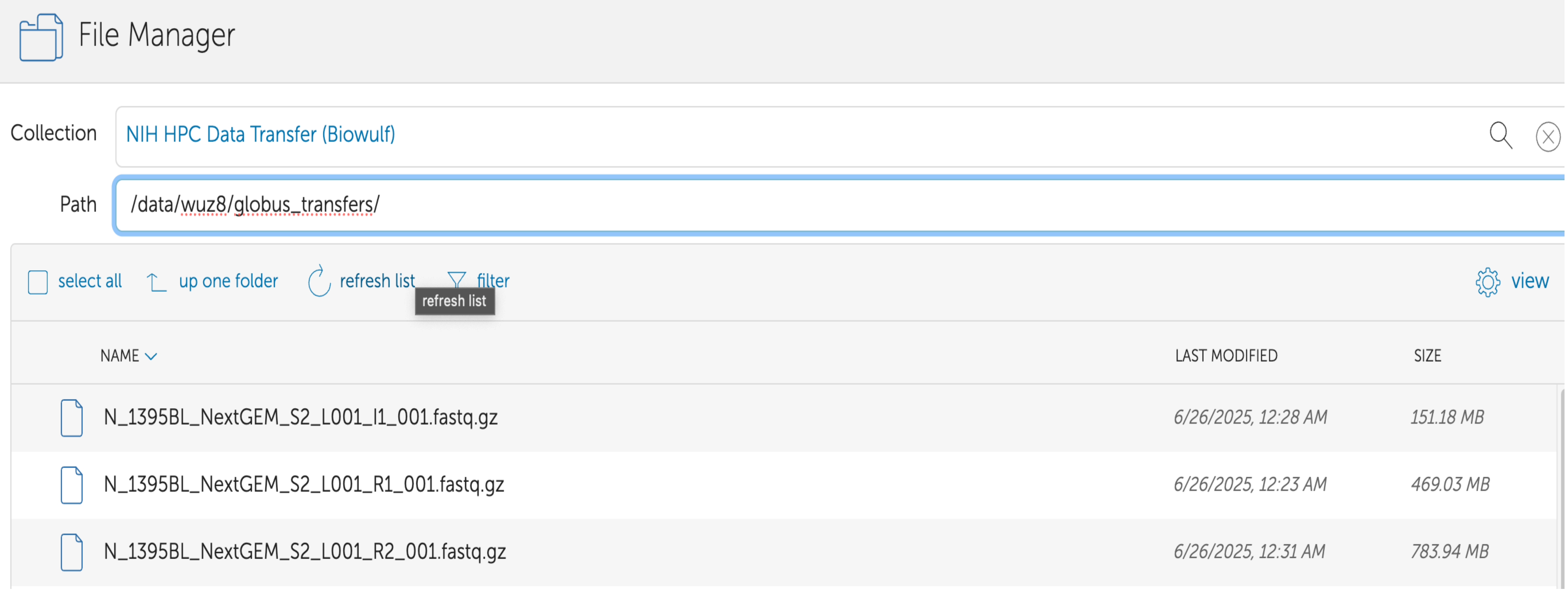
This message is flagged for follow up.

Your download request is completed.
Destination type: GLOBUS
Source path: /FNL_SF_Archive/PI_Lab_SF_Lab/Project_SFRnD_CS026880_4scRNAseq_030620/Flowcell_HYH2JBGXC/Sample_N_1395BL_NextGEM
Target path: globus demonstration:/
Task ID: cc3fea46-fcb6-4d90-a2bf-02f799c65427

You are receiving this email because you subscribed to this notification from DME.
Unsubscribe

Getting Data from Sequencing Facility DME (step 7)

The folder in which the data was transferred to is now populated with content.



Collection: NIH HPC Data Transfer (Biowulf)

Path: /data/wuz8/globus_transfers/

NAME	LAST MODIFIED	SIZE
N_1395BL_NextGEM_S2_L001_I1_001.fastq.gz	6/26/2025, 12:28 AM	151.18 MB
N_1395BL_NextGEM_S2_L001_R1_001.fastq.gz	6/26/2025, 12:23 AM	469.03 MB
N_1395BL_NextGEM_S2_L001_R2_001.fastq.gz	6/26/2025, 12:31 AM	783.94 MB