

Introduction to Data Transfer using Globus

Joe Wu, PhD

NCI CCR Bioinformatics Training and Education Program

ncibtep@nih.gov

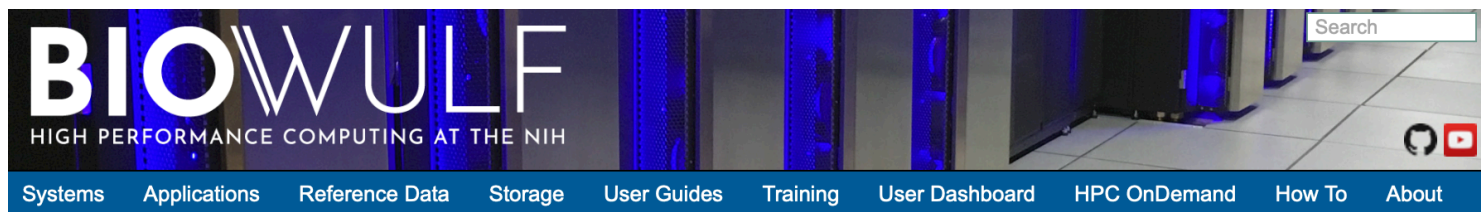
Globus

"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)

Quick Links

[NIH HPC Globus Endpoints](#)

[Globus Setup ▶](#)

[Globus Data Transfers ▶](#)

[Globus Sharing ▶](#)

[Extras ▶](#)

[Biowulf to/from desktop](#)

[desktop to/from desktop](#)

[via the command line](#)

[Recurring/Scheduled transfers](#)

[to/from AWS](#)

[to/from Google Cloud Storage](#)

[to/from Google Drive](#)

[shared Google Drive data](#)

[to/from NIH OneDrive](#)

[to/from NIH Box](#)


[to/from NIH NCI DME](#)

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 major research facilities. [\[Globus website\]](#)

When moving data in and out of our systems, be aware that PII and PHI data cannot be moved to or from NIH HPC systems.

For more information, see the menu at left for details.



Sharing data with collaborators

If your data is on the NIH HPC (free) Globus account. The advantage is that you do not need to transfer your data anywhere. This prevents data duplication, wastage of storage space, and you can have full control over which files your collaborator can access, and whether they have read-only or write access.

Sharing data from NIH HPC to collaborators

If your data is on the NIH HPC (free) Globus account. The advantage is that you do not need to transfer your data anywhere. This prevents data duplication, wastage of storage space, and you can have full control over which files your collaborator can access, and whether they have read-only or write access.

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 cloudevironments)
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 <i>may</i> 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.



The screenshot shows the Globus website homepage. The top navigation bar includes the Globus logo (a cloud with a 'g') and the text 'globus a ucjicagu non-profit service'. To the right of the logo are two buttons: 'GET STARTED' with a rocket icon and 'LOG IN' with a login icon, which is highlighted with a red square. Below the navigation bar is a dark blue banner with the text 'Go beyond data' and 'Globus Compute' in large white font. Underneath this is the tagline 'Reliable, distributed Function-as-a-Service' and 'COMPUTE ANYWHERE: EDGE TO SUPERCOMPUTER'. To the left of the text is a diagram showing a central figure with a laptop, connected to various computing environments: leadership class computing, commercial computing, research computing, institutional computing, and personal computing. Below the banner is a light blue section with two columns. The left column has an icon of a server rack and the text 'Transfer your data' followed by 'Gigabytes, terabytes, petabytes—research data is large and distributed. Globus lets you'. The right column has an icon of a folder and the text 'Share your data' followed by '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.

The screenshot shows the Biowulf HPC OnDemand web interface. The header includes the Biowulf logo and navigation links: HPC OnDemand, Files (highlighted with a red box), Interactive Apps, and My Interactive Sessions. The 'Files' dropdown menu is open, showing options: Data /data/wuz8, Data (GAU) /data/GAU, Data (gau) /data/gau, and Home Directory. The main navigation bar includes links to HPC Dashboard, Help, and a user login status (Logged in as wuz8). The main content area shows a directory listing for /data/wuz8/. The 'Globus' button is highlighted with a red box. The directory listing shows two folders: ccb_r_chipsequencing_test and ccb_r_example_rna_sequencing.

Biowulf
HIGH PERFORMANCE COMPUTING AT THE NIH

HPCC OnDemand Files Interactive Apps My Interactive Sessions

HPCC Dashboard Help Logged in as wuz8 Log Out

Data /data/wuz8
Data (GAU) /data/GAU
Data (gau) /data/gau
Home Directory

Refresh New File New Directory Upload Download **Globus** Copy/Move Delete

↑ / data / wuz8 / Change directory Copy path

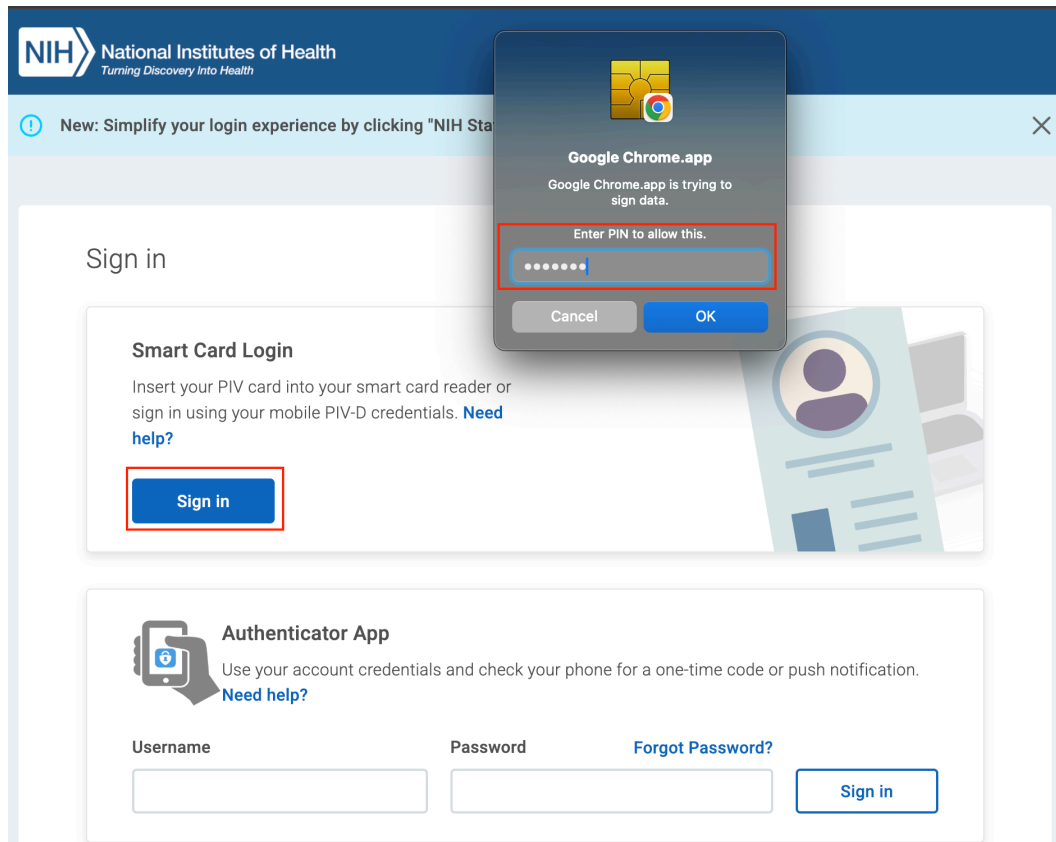
☐ Show Owner/Mode ☐ Show Dotfiles Filter:

Showing 41 of 47 rows - 0 rows selected

<input type="checkbox"/>	Type ▲	Name	Size	Modified at
<input type="checkbox"/>	Folder	ccbr_chipsequencing_test	-	2/27/2025 9:43:37 PM
<input type="checkbox"/>	Folder	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.



The screenshot shows the NIH Sign in page. At the top, the NIH logo and "National Institutes of Health" are displayed. Below this, a notification bar states: "New: Simplify your login experience by clicking 'NIH Sta...". The main sign-in area has a "Sign in" heading. Underneath, there are two primary login methods:

- Smart Card Login:** Includes the instruction "Insert your PIV card into your smart card reader or sign in using your mobile PIV-D credentials." and a blue "Sign in" button, which is highlighted with a red rectangle.
- Authenticator App:** Includes the instruction "Use your account credentials and check your phone for a one-time code or push notification." and a "Need help?" link.

At the bottom, there are input fields for "Username" and "Password", a "Forgot Password?" link, and a "Sign in" button.

Overlaid on the page is a "Google Chrome.app" dialog box. It contains the text "Google Chrome.app is trying to sign data." and "Enter PIN to allow this." followed by a PIN input field (represented by dots) and "Cancel" and "OK" buttons. The PIN input field is highlighted with a red rectangle.

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 be prompted if the information the service requires changes.

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

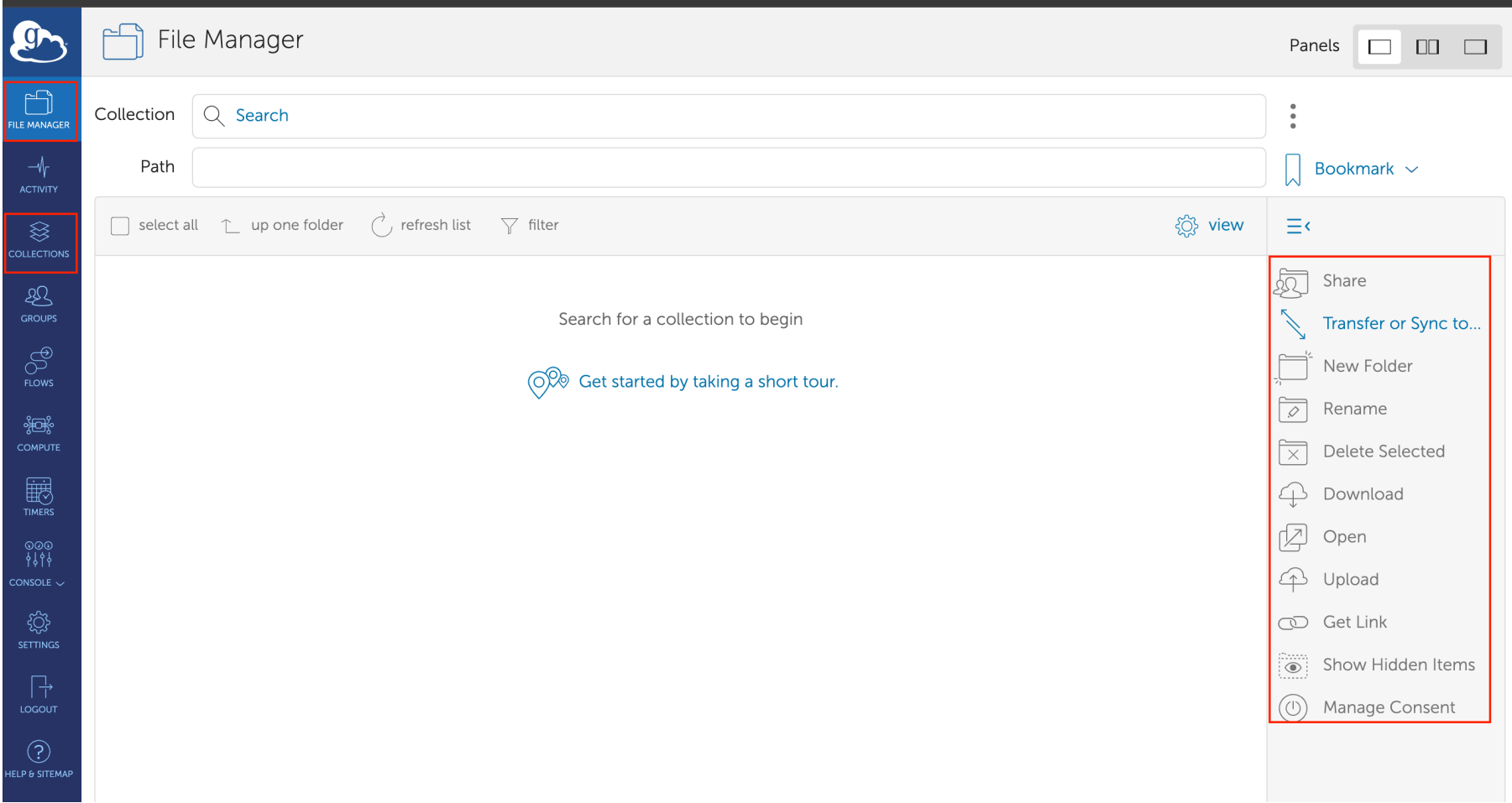
- I agree to release the information listed above to this service and to any service that asks for the same information. I want to be prompted if a service asks for different information.

I Do Not Agree

I Agree



Globus Landing Page



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.

Collections

Get Globus Connect Personal

Search collections with recent tasks

Recent Tasks

Administered By You

In Use

Shareable By You

Shared With You

TYPE

All

COLLECTION

HA

SUBSCRIBED

STATUS


ROLE

NIH HPC Data Transfer (Biowulf)

Subscribed Mapped Collection (GCS) on NIH HPC Data Transfer

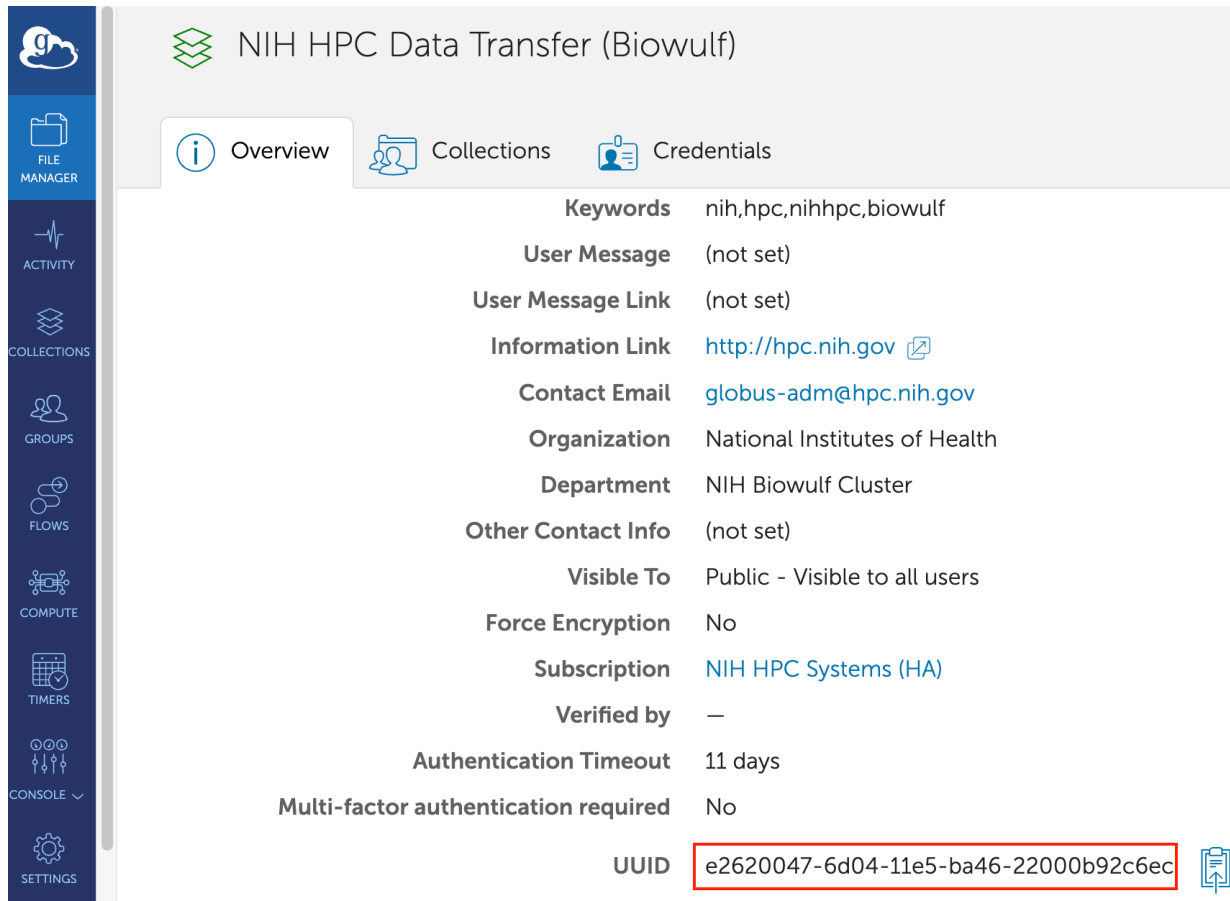
ready

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)

Overview Collections Credentials

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

Collections

Get Globus Connect Persona

Search collections that are administered by you

Recent Tasks

Administered By You

















In Use

Shareable By You

Shared With You

TYPE

All

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

Local computer

Setting up Globus Local Endpoint (step 1)

Launch Globus desktop client and choose "Log In".

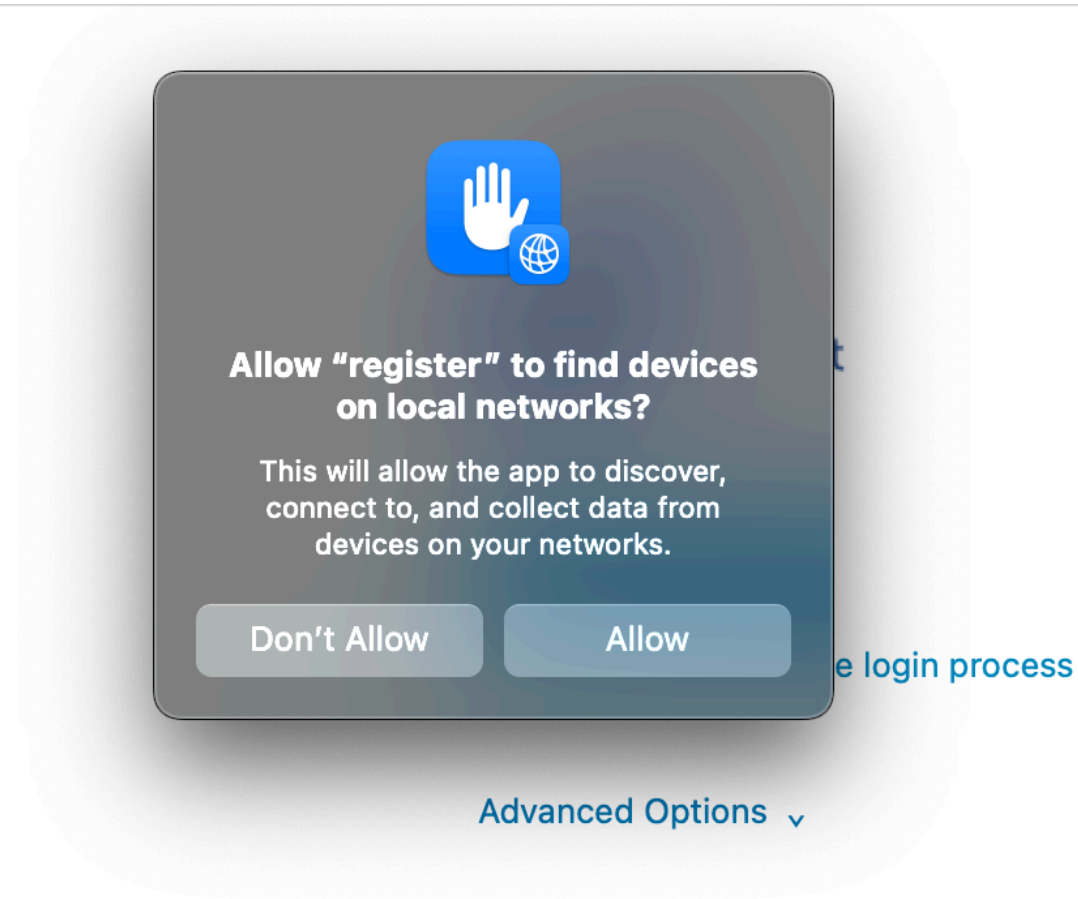


Log In

Advanced Options ▾


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.



Setting up Globus Local Endpoint (step 3)

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

 globus

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

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.

Save

Setting up Globus Local Endpoint (step 5)

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



Collections

FILE MANAGER

ACTIVITY

COLLECTIONS

GROUPS


FLows


COMPUTE


TIMERS


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.

The screenshot displays two instances of the 'File Manager' application. The left instance shows a collection named 'NIH HPC Data Transfer (Biowulf)' with a path of '/data/wuz8/'. The right instance is a second file manager window, also titled 'File Manager', showing the same collection and path. A context menu is open over the file list in the right window, showing options like 'Share', 'Transfer or Sync to...', 'New Folder', 'Rename', 'Delete Selected', 'Download', 'Open', 'Upload', 'Get Link', and 'Show Hidden Items'.

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.

g

FILE MANAGER

ACTIVITY

COLLECTIONS

GROUPS

Flows

COMPUTE

TIMERS

CONSOLE

SETTINGS

File Manager

Panels

Collection NIH HPC Data Transfer (Biowulf)

Path /data/wuz8/

Start

Transfer & Timer Options

Start

select all up one folder refresh list filter view

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

Share

Transfer or Sync to...

New Folder

Rename

Delete Selected

Download

Open

Upload

Joe.Wu.local

/Users/wuz8/Downloads/b4b_2025_data/


Start

select none up one folder refresh list filter view

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.

File Manager

FILE MANAGER

1
ACTIVITY

COLLECTIONS

GROUPS

FLows

COMPUTE

TIMERS

CONSOLE

SETTINGS

LOGOUT

HELP & SITEMAP

CollectionNIH HPC Data Transfer (Biowulf)

Path/data/wuz8/

Start

Transfer & Timer Options

Start

select allup one folderrefresh listfilterview

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

Share

Transfer or Sync to...

New Folder

Rename

Delete Selected

Download

Open

Upload

Get Link

Show Hidden Items

Manage Consent

Joe.Wu.local

/Users/wuz8/Downloads/b4b_2025_data/

Start

select noneup one folderrefresh listfilterview

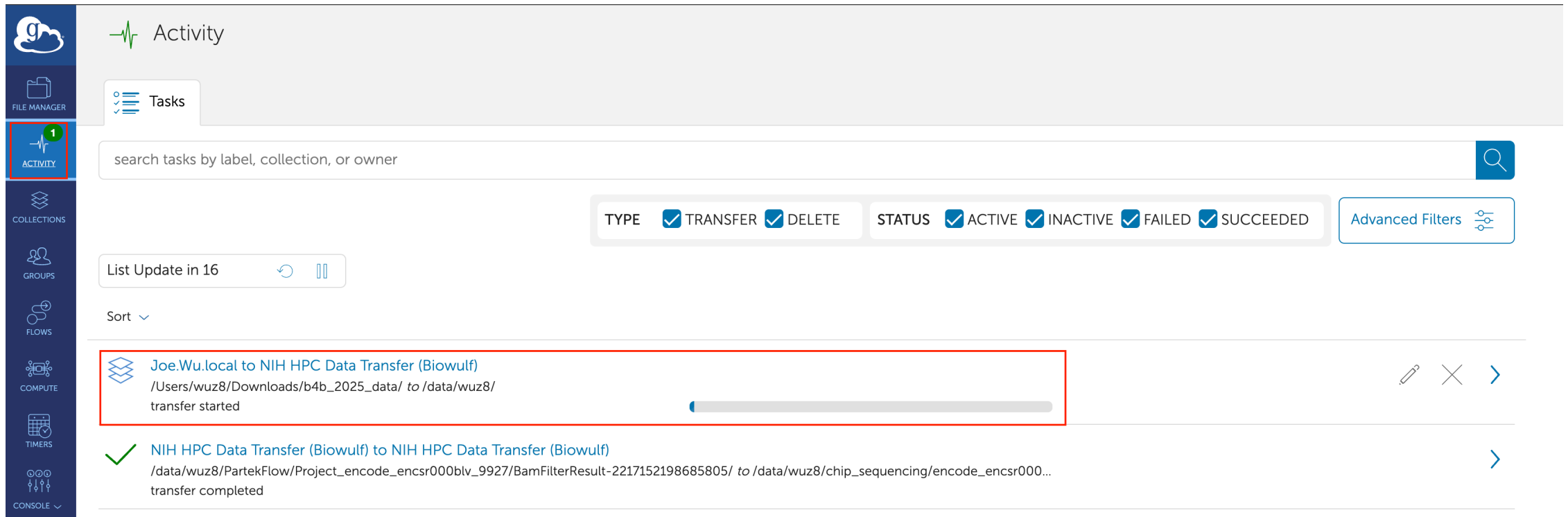
NAME	LAST MODIFIED	SIZE
hbr_uhr_b4b	12/26/2024, 09:59 ...	—
hcc1395_b4b	12/27/2024, 12:07 ...	—

Transfer request submitted successfully

View details

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.



The screenshot displays the 'Activity' page in a web application. On the left is a dark blue sidebar with icons for 'FILE MANAGER', 'ACTIVITY' (highlighted with a red box and a '1'), 'COLLECTIONS', 'GROUPS', 'FLOWS', 'COMPUTE', 'TIMERS', and 'CONSOLE'. The main content area has a light gray header with a green pulse icon and the word 'Activity'. Below this is a 'Tasks' tab and a search bar labeled 'search tasks by label, collection, or owner'. A filter bar shows 'TYPE' with checkboxes for 'TRANSFER' and 'DELETE', and 'STATUS' with checkboxes for 'ACTIVE', 'INACTIVE', 'FAILED', and 'SUCCEEDED'. There is also an 'Advanced Filters' button. Below the filters, a 'List Update in 16' button with a refresh icon is visible. The task list shows two entries: 1. 'Joe.Wu.local to NIH HPC Data Transfer (Biowulf)' with a folder icon, a path '/Users/wuz8/Downloads/b4b_2025_data/ to /data/wuz8/', the status 'transfer started', and a progress bar. 2. 'NIH HPC Data Transfer (Biowulf) to NIH HPC Data Transfer (Biowulf)' with a green checkmark icon, a path '/data/wuz8/PartekFlow/Project_encode_encsr000blv_9927/BamFilterResult-2217152198685805/ to /data/wuz8/chip_sequencing/encode_encsr000...', and the status 'transfer completed'. Each task entry has edit, delete, and expand icons on the right.

Activity

Tasks

search tasks by label, collection, or owner

TYPE ☒ TRANSFER ☒ DELETE STATUS ☒ ACTIVE ☒ INACTIVE ☒ FAILED ☒ SUCCEEDED Advanced Filters

List Update in 16


Sort


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

 Overview

 Event Log

Task Label

Source

Destination

Task ID

Owner

Condition

Requested

Deadline

Duration

Base Paths

Transfer Settings

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

► joe.local.computer

► NIH HPC Data Transfer (Biowulf)

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

► Zhuoxi Wu (wuz8@nih.gov)

ACTIVE

7/28/2025, 06:44 PM

7/29/2025, 06:44 PM

28 seconds


Source


Destination

• verify file integrity after transfer

• transfer is not encrypted

• overwriting all files on destination

 Edit Label


 Terminate Task

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


[View debug data](#)


Schedule Data Transfer


Users can schedule data transfer.


FILE MANAGER


ACTIVITY


COLLECTIONS


GROUPS


FLOWS


COMPUTE

TIMERS

CONSOLE

SETTINGS

LOGOUT

HELP & SITEMAP

File Manager

Panels

CollectionNIH HPC Data Transfer (Biowulf)

Joe.Wu.local

Path/data/wuz8/

/Users/wuz8/Downloads/b4b_2025_data/

Start

Transfer & Timer Options

Start

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".



The screenshot displays the Globus File Manager interface. On the left, a vertical sidebar contains navigation icons for File Manager, Activity, Collections, Groups, and Flows. The main area is titled 'File Manager' and shows the 'NIH HPC Data Transfer (Biowulf)' collection. The path is set to '/data/wuz8/'. Below the path, there are controls for 'select all', 'up one folder', 'refresh list' (highlighted with a blue box), and 'filter'. A table lists the files and folders in the current directory:

NAME	LAST MODIFIED	SIZE
<input type="checkbox"/> globus_transfers	6/25/2025, 02:31 PM	—
<input type="checkbox"/> hcc1395_b4b	6/25/2025, 10:33 AM	—

The 'globus_transfers' folder is highlighted with a red box. On the right side of the interface, there is a 'Panels' section with three icons. Below it, a 'Bookmark' dropdown is visible. At the bottom right, a menu is open showing options: 'Share', 'Transfer or Sync to...', and 'New Folder' (highlighted with a red box).

Add Guest Collection

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

FILE MANAGER

ACTIVITY

COLLECTIONS

GROUPS

NIH HPC Data Transfer (Biowulf)

Overview

Collections

Credentials

Guest Collections

< First

Next >

☒ CREATED BY ME

☐ MANAGED BY ME

☐ NEVER ACCESSED

☐ 90+ DAYS OLD

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

Add Guest Collection

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.

Create new Guest Collection

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

Directory ?

[up one folder](#) [refresh list](#)

This folder is empty.

Display Name

Description


Keywords

☐ Force encryption on transfers to and from this collection

[view more fields](#) ▼

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.

 globus demonstration

FILE MANAGER

ACTIVITY

COLLECTIONS

GROUPS

FLows

COMPUTE

TIMERS

Overview

Permissions



Roles

Shared With

Mapped Collection: NIH HPC Data Transfer (Biowulf)

Base Path: /data/wuz8/globus_transfers/

Add Permissions — Share With

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	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
 Zhuoxi Wu (wuz8@nih.gov)	-	never expires	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

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

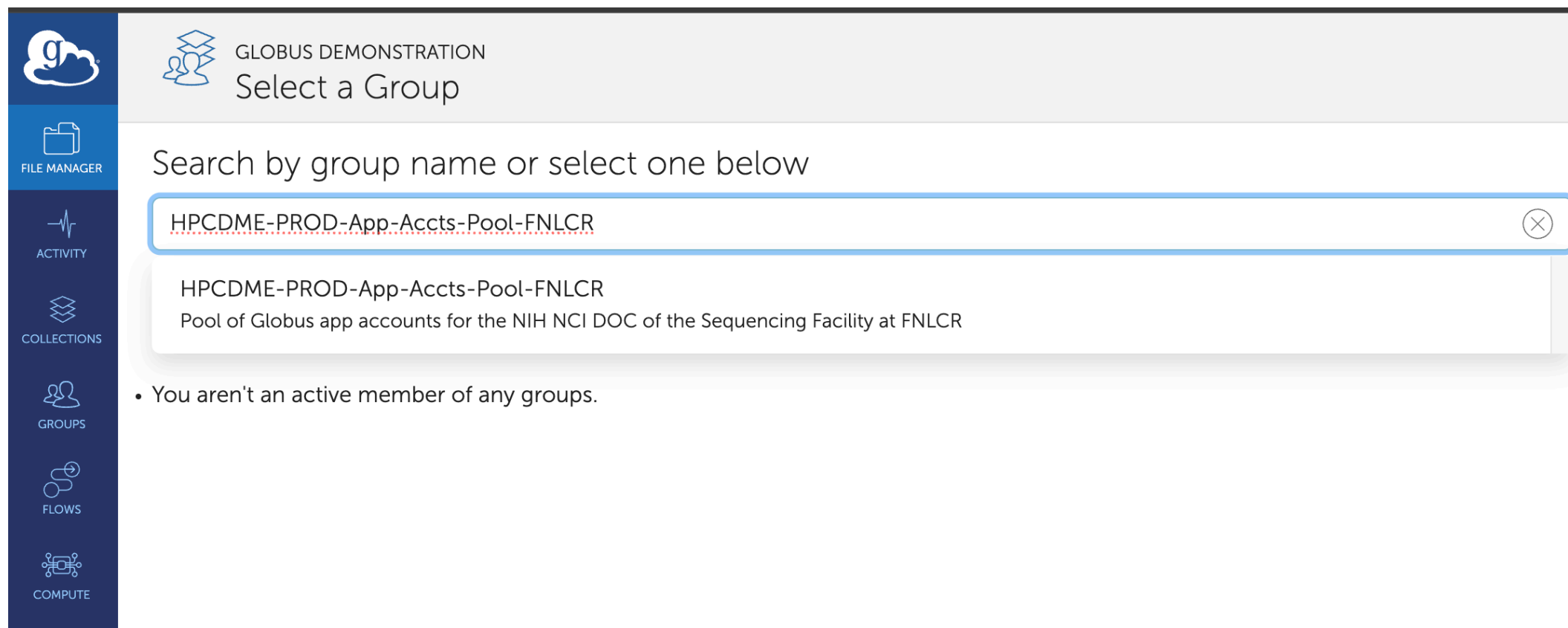
Group

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.



The screenshot shows the 'Select a Group' interface in the Globus DEMONSTRATION. On the left is a dark blue sidebar with icons and labels for 'FILE MANAGER', 'ACTIVITY', 'COLLECTIONS', 'GROUPS', 'FLOWS', and 'COMPUTE'. The main area has a light gray header with the 'GLOBUS DEMONSTRATION' logo and the title 'Select a Group'. Below the header, the text 'Search by group name or select one below' is displayed. A search input field contains the text 'HPCDME-PROD-App-Accts-Pool-FNLCR' and has a close button (X) on the right. Below the input field, a dropdown menu is open, showing a single result: 'HPCDME-PROD-App-Accts-Pool-FNLCR' with the description 'Pool of Globus app accounts for the NIH NCI DOC of the Sequencing Facility at FNLCR'. At the bottom of the main area, a message states: '• You aren't an active member of any groups.'

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.

The screenshot shows the 'Add Permissions - Share With' interface in the Globus DEMONSTRATION application. On the left is a dark blue sidebar with icons and labels for 'FILE MANAGER', 'ACTIVITY', 'COLLECTIONS', 'GROUPS', 'FLOWS', and 'COMPUTE'. The main content area has a light gray header with the Globus logo and the text 'GLOBUS DEMONSTRATION Add Permissions - Share With'. Below the header, the 'Path' is set to '/'. The 'Share With' section has four radio button options: 'user - share with specific individuals', 'group - make data accessible to members of a group' (which is selected), 'all users - make data accessible to all logged in users of Globus', and 'public (anonymous) - make data accessible to everyone'. The 'Group' field displays 'HPCDME-PROD-App-Accts-Pool-FNLCR' with a 'Change' button to its right. The 'Permissions' section has two checked checkboxes: 'read' and 'write'. At the bottom are 'Add Permission' and 'Cancel' buttons.

GLOBUS DEMONSTRATION
Add Permissions - Share With

Path: / Browse

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 Change

Permissions:

- ☒ read
- ☒ write

Add Permission Cancel

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

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

FILE MANAGER

ACTIVITY

COLLECTIONS

GROUPS

FLOWS

COMPUTE

TIMERS

GLOBAL DEMONSTRATION

Add Permissions - Share With

Path /

Browse

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

Change

Permissions

Permissions Added

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

Add another Permission Done

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.

globus demonstration

Overview Permissions Roles

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.

NIH NATIONAL CANCER INSTITUTE

Data Management Environment

Joe Wu Help version: 3.2.1

Dashboard Register Search Browse Manage

Collection

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)

User Metadata

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.



Data Management Environment

Joe Wu

Help

version: 3.2.1

Dashboard

Register >

Search

Browse

Manage >

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.

 **NATIONAL CANCER INSTITUTE**

Data Management Environment

[Dashboard](#) [Register >](#) [Search](#) [Browse](#) [Manage >](#)

 Download

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/PL_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

Download

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.



Data Management Environment

[Dashboard](#) [Register >](#) [Search](#) [Browse](#) [Manage >](#)

Download

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.



Data Management Environment

Dashboard Register > Search Browse **Manage >**

Download Tasks

My Tasks

- Notifications
- Download Tasks**
- Registration Tasks

Show 200

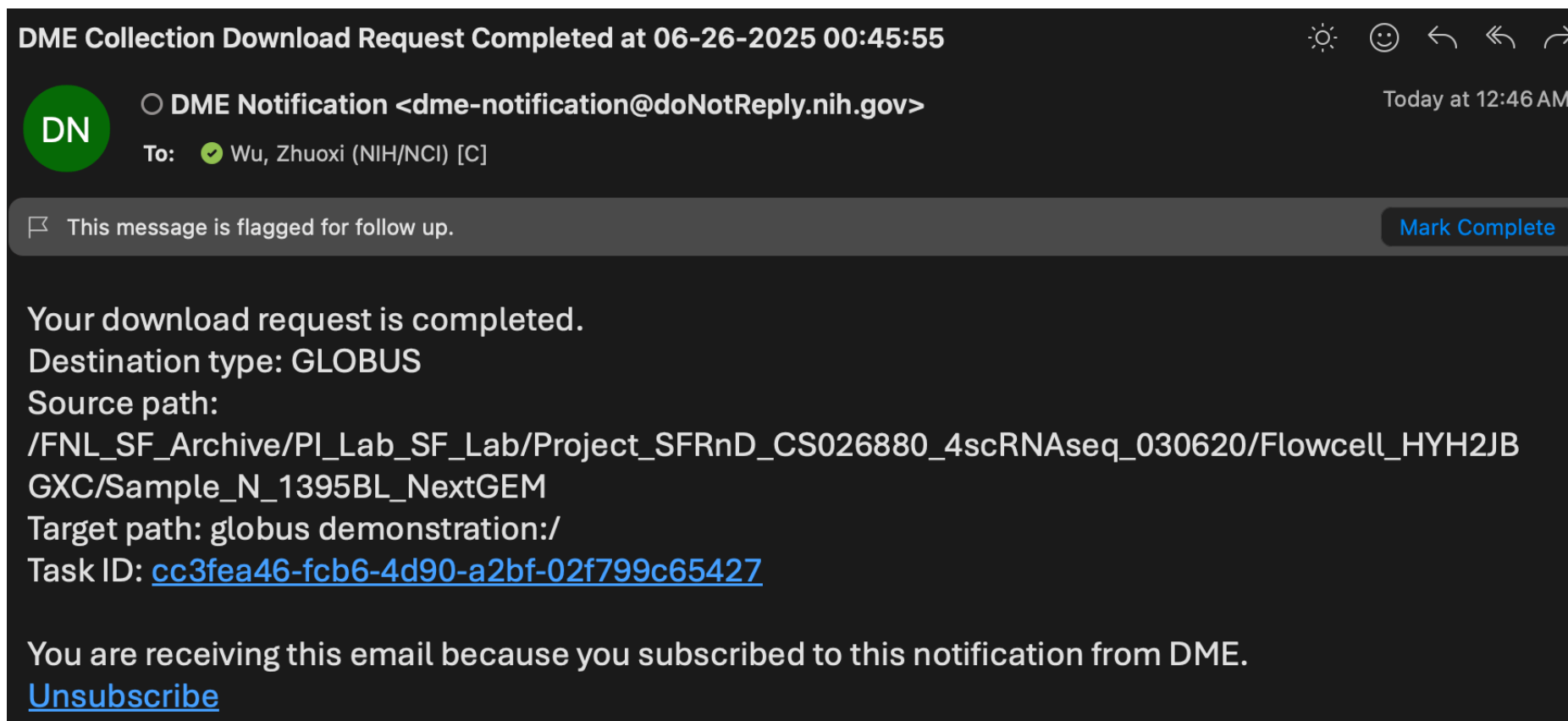
Task ID	Path	Type	Destination Type	Created On	Completed On	Task Status
<input type="text" value="Filter Task ID"/>	<input type="text" value="Filter Path"/>	<input type="text" value="Filter Type"/>	<input type="text" value="Filter Destination Ty"/>	<input type="text" value="Filter Created On"/>	<input type="text" value="Filter Completed Or"/>	<input type="text" value="Filter 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.



Getting Data from Sequencing Facility DME (step 7)


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


 File Manager


Collection NIH HPC Data Transfer (Biowulf)  


Path /data/wuz8/globus_transfers/





☐ select all

 up one folder

 refresh list

 filter

 view

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