

GLOBUS TOOLS

Framework Reference Card

Data transfer, remote compute, and workflow automation for Genesis Mission • globus.org; docs: docs.globus.org

Globus provides secure, reliable high-performance data movement (Transfer), remote function execution (Compute), and workflow orchestration (Flows). Globus services are available at thousands of institutions, including most DOE laboratories. This card covers connecting to AmSC infrastructure for Genesis Mission teams.

1 • Access Methods

Every Globus service is available through **three interfaces**:

Interface	Best for	URL / Install
Web App	Interactive use, browsing, one-off transfers	app.globus.org
CLI	Scripting, automation, batch operations	<code>pip install globus-cli</code>
Python SDK / API	Application integration, programmatic workflows	<code>pip install globus-sdk</code>

Choose the interface that fits your workflow — they all use the same underlying services and authentication.

2 • Installation (CLI and SDK)

Skip this section if you only need the Web interface.

Globus CLI

Platform	Command	Updates
pip (recommended)	<code>pip install globus-cli</code>	<code>pip install --upgrade globus-cli</code>
pipx (isolated)	<code>pipx install globus-cli</code>	<code>pipx upgrade globus-cli</code>
Conda	<code>conda install -c conda-forge globus-cli</code>	<code>conda update globus-cli</code>

Globus SDK (Python)

```
pip install globus-sdk globus-compute-sdk
```

Globus Connect Personal (optional)

To make your laptop or workstation a Globus endpoint:

Platform	Installer
macOS	Download from globus.org/globus-connect-personal
Linux	<code>wget https://downloads.globus.org/globus-connect-personal/linux/stable/globusconnectpersonal-latest.tgz</code>
Windows	Download from globus.org/globus-connect-personal

Globus Compute Endpoint

To run your own Compute endpoint on a cluster or server:

```
pip install globus-compute-endpoint
globus-compute-endpoint configure <endpoint-name>
```

This creates a configuration directory at `~/ .globus_compute/<endpoint-name>/`. Edit `config.yaml` to customize for your cluster (many examples are provided in the Globus Compute documentation).

For HPC systems (Slurm, PBS, etc.), configure the appropriate provider:

```
engine:
  type: HighThroughputEngine
  provider:
    type: SlurmProvider
    partition: "compute"
    account: "<your-allocation>"
    nodes_per_block: 1
    max_blocks: 10
    walltime: "01:00:00"
```

Start and manage your endpoint:

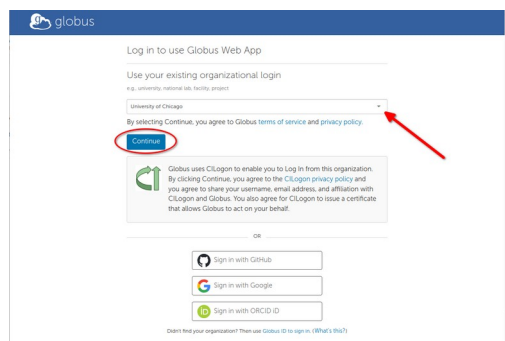
```
globus-compute-endpoint start <endpoint-name>    # start (prints endpoint UUID)
globus-compute-endpoint status <endpoint-name>    # check status
globus-compute-endpoint stop <endpoint-name>      # stop
globus-compute-endpoint list                     # list all endpoints
```

3 · Authentication

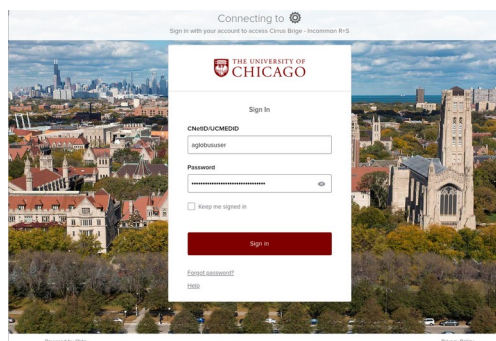
Genesis Mission teams authenticate via their institutional login and in the future via the AmSC identity provider. Authentication works the same way regardless of which interface you use.

Web App

1. Go to **app.globus.org**
2. Click **Log In**
3. Search for and select your institution
4. Complete login with your identity provider



Select your institution (e.g., UChicago)



Authenticate with selected identity

CLI

```
globus login
```

A browser window opens — select your institutional identity when prompted.

Verify your identities

To simplify use of diverse resources, you can link identities into an identity set.

Web: Click your name (top right) → **Account** → **Identities**

CLI:

```
globus whoami --linked-identities
```

Session management

Web: Account → Consents (manage app permissions), or Log Out from menu

CLI:

```
globus session show      # list active sessions
globus logout            # revoke all tokens
globus login --force      # re-authenticate
```

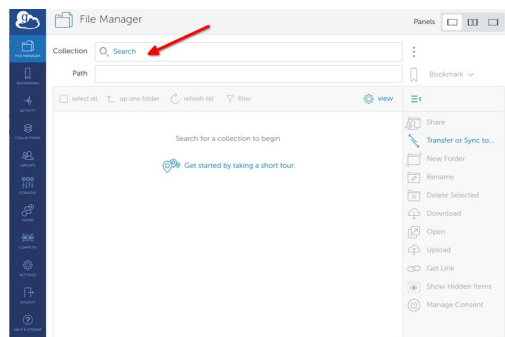
Tokens: CLI stores tokens at `~/ .globus .cfg`. For automation, use client credentials or refresh tokens — never embed tokens in scripts committed to version control.

4 · Globus Transfer (Data Access, Transfer, and Sharing)

Genesis Mission teams have access to tens of thousands of Globus Collections [a file system or directory accessible via Globus], including on most AmSC storage systems.

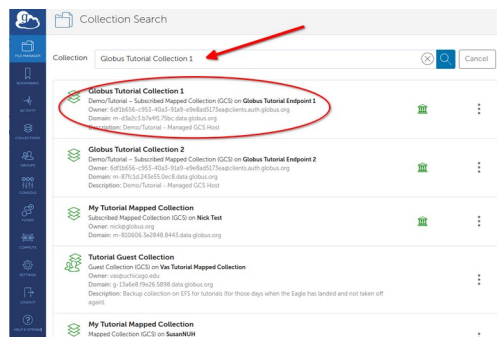
Find collections

Web: Go to **File Manager** → type collection name in the search box



Search

for collections



from results

Select

CLI:

```
globus endpoint search "ALCF"      # search by name
globus bookmark list              # list saved bookmarks
```

Bookmark frequently-used collections

Web: Open a collection → click the **Bookmark** icon (star)

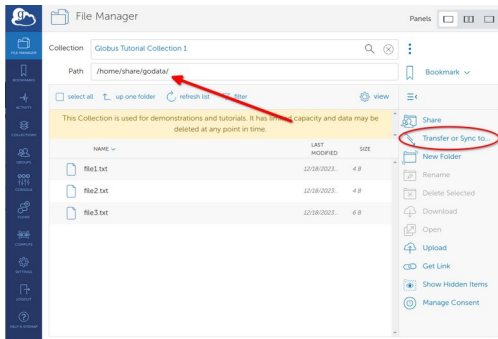
CLI:

```
globus bookmark create <shared-collection-uuid> "AmSC Shared"
globus bookmark create <scratch-collection-uuid> "AmSC Scratch"
```

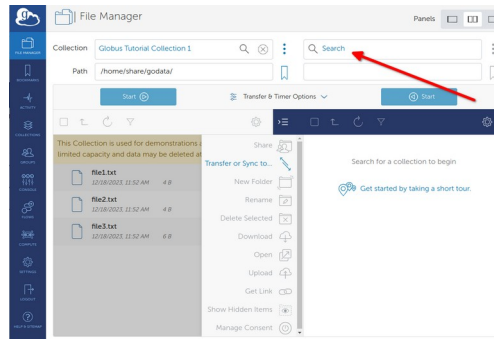
Web App — Interactive Transfer

1. Go to **app.globus.org** → **File Manager**
2. In the left panel, select your **source** collection and navigate to files
3. Click the **two-panel icon** to enable Transfer view

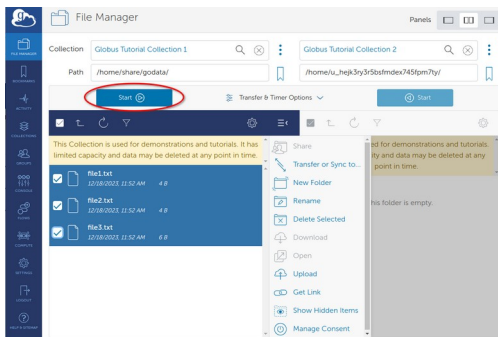
- In the right panel, select your **destination** collection
- Select files/folders, then click **Start** (or **Sync** for incremental)



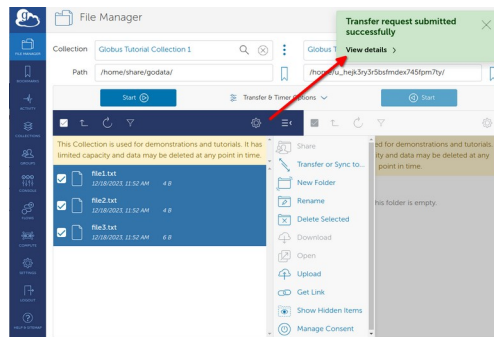
Browse source collection



Two-panel transfer view



Select files, click Start



submitted

Transfer

CLI — Transfer Commands

Single file

```
globus transfer <source-uuid>:/path/to/file.dat <dest-uuid>:/path/to/file.dat
```

Directory (recursive)

```
globus transfer <source-uuid>:/data/experiment/ <dest-uuid>:/data/experiment/ --recursive
```

With Label for tracking

```
globus transfer <source-uuid>:/data/ <dest-uuid>:/data/ --recursive --label "Genesis run 42 results"
```

Sync directories (incremental)

Web: Select files → click **Sync** instead of Start, choose sync level

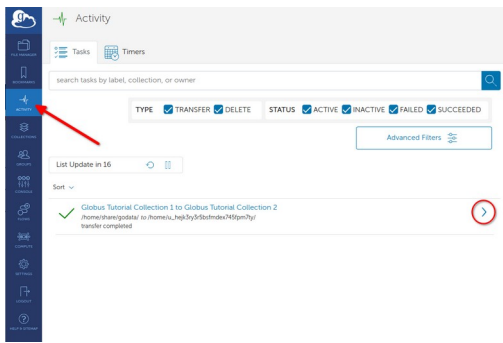
CLI:

```
globus transfer <source-uuid>:/data/ <dest-uuid>:/data/ --recursive --sync-level checksum
```

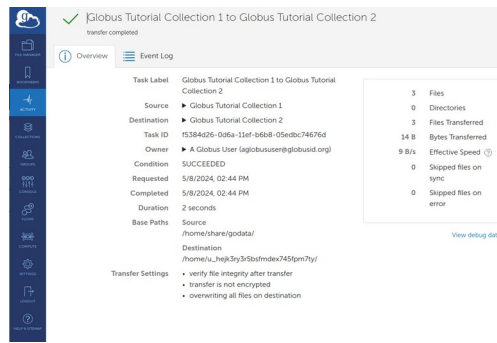
Sync levels: exists, size, mtime, checksum

Monitor transfers

Web: Go to **Activity** in the left sidebar to see all tasks



Activity page with task list



Task details and event log

CLI:

```
globus task list
globus task show <task-id>
globus task wait <task-id>
globus task event-list <task-id>
```

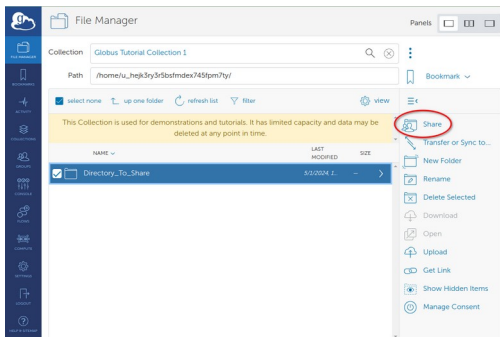
```
# list recent tasks
# detailed status
# block until complete
# event log
```

File Sharing

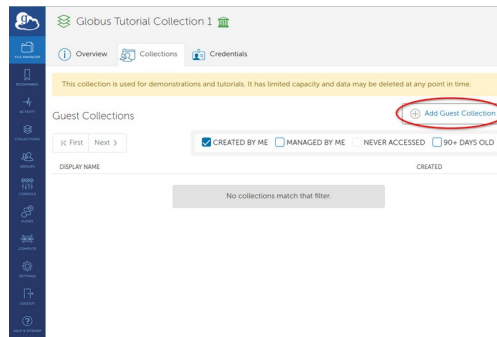
Share data with collaborators by creating a Guest Collection. *This is possible only if the underlying collection (e.g., ALCF Eagle) allows sharing; also, administrators may restrict the directories that can be shared.*

Web App — Share Files

1. In File Manager, navigate to the folder you want to share
2. Select the folder and click **Share** in the right panel
3. Click **Add Guest Collection**
4. Name your guest collection and click **Create**
5. Use the **Permissions** tab to add users or groups

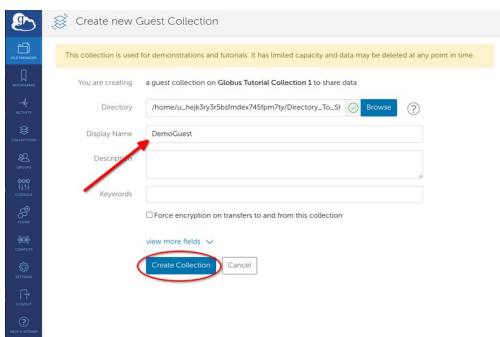


Select folder, click Share

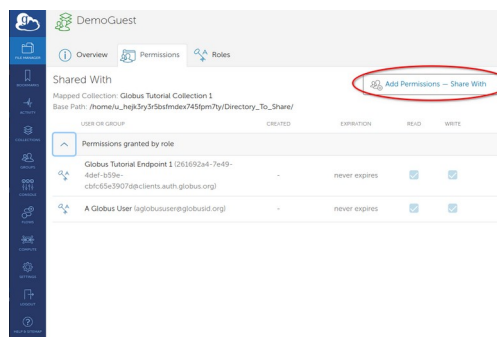


Add

Guest Collection



Name your collection



Manage

permissions

5 • Globus Compute (Remote Function Execution)

Globus Compute runs Python functions on remote endpoints without manual SSH or job submission.

Web App — Compute Dashboard

1. Go to **app.globus.org** → **Compute** (in left sidebar)
2. Discover accessible endpoints, view your registered endpoints, function history, and task status

Python SDK — Run Remote Functions

from globus_compute_sdk **import** Executor

AMSC_COMPUTE_ENDPOINT = "<compute-endpoint-uuid>"

```
def my_analysis(data_path):  
    import pandas as pd  
    df = pd.read_parquet(data_path)  
    return df.describe().to_dict()
```

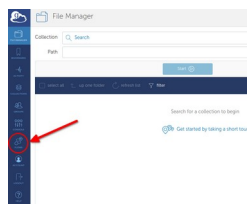
```
with Executor(endpoint_id=AMSC_COMPUTE_ENDPOINT) as ex:  
    future = ex.submit(my_analysis, "/path/to/data.parquet")  
    result = future.result()  
    print(result)
```

6 • Globus Flows (Workflow Automation)

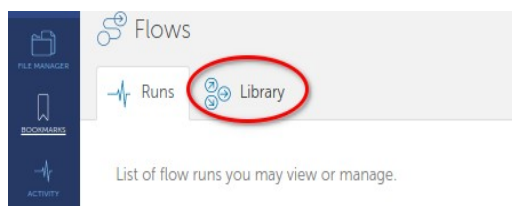
Flows orchestrate multi-step operations: e.g., transfer data, compute, transfer results.

Web App — Flows Dashboard

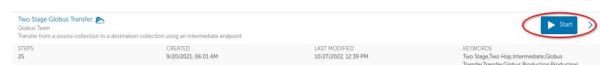
1. Go to **app.globus.org** → **Flows** (in left sidebar)
2. Browse available flows in the **Library** tab
3. Click **Start** to run a flow
4. Monitor runs in the **Runs** tab



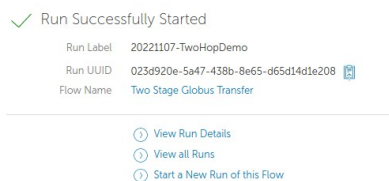
Flows menu



Runs and Library tabs



Click Start to run a flow



Run confirmation

CLI — Flow Commands

```
globus flows login  
globus flows list
```

```
# authenticate for Flows  
# list available flows
```

```
globus flows run <flow-id> --input input.json
globus flows run-log <run-id> # check run status
```

Python SDK — Define and Run Flows

Flows can be authored via the web app or using the Python SDK. Flow definitions are finite state machines written in JSON that declares a network of states with transitions between them. Comprehensive documentation is provided on the Globus web app to support authoring.

7 · Quick Reference

Web App URLs

Service	URL
File Manager (Transfer)	app.globus.org/file-manager
Activity (Task Monitor)	app.globus.org/activity
Collections	app.globus.org/file-manager/collections
Compute	app.globus.org/compute
Flows	app.globus.org/flows
Account & Settings	app.globus.org/account

CLI Commands

Command	What it does
globus login	Authenticate with Globus (opens browser)
globus whoami	Show your Globus identity
globus ls <uuid>:/path/	List files on a collection
globus transfer <src> <dst>	Transfer files between collections
globus task list	List recent transfer tasks
globus task show <id>	Show task status
globus task wait <id>	Wait for task to complete
globus bookmark list	List saved collection bookmarks
globus endpoint search <query>	Search for endpoints by name
globus-compute whoami	Show Compute identity
globus flows list	List available flows
globus --help	Full CLI help

8 · Troubleshooting

Authentication issues

Web: Log out, clear browser cache, log in again selecting correct institution

CLI:

```
globus logout
globus login --force
globus session show
```

Permission denied on collection

- **Web:** Check collection permissions under **Permissions** tab; request access from owner

- **CLI:** Verify your identity is linked: `globus whoami --linked-identities`
- Contact the collection administrator for access

Transfer stuck or slow

Web: Go to **Activity** → click the task → view **Event Log** for errors

CLI:

```
globus task show <task-id>           # check status
globus task event-list <task-id>      # look for errors
globus task cancel <task-id>         # cancel if needed
```

Compute function fails

- Ensure required packages are installed on the compute endpoint
- View logs in Web App under **Compute** → **Tasks**
- CLI: `globus-compute-endpoint status` to verify endpoint health

9 • Resources

Resource	URL
Globus Web App	app.globus.org
Globus Documentation	docs.globus.org
File Transfer Tutorial	docs.globus.org/guides/tutorials/manage-files/transfer-files/
File Sharing Tutorial	docs.globus.org/guides/tutorials/manage-files/share-files/
Globus CLI Reference	docs.globus.org/cli
Globus SDK (Python)	globus-sdk-python.readthedocs.io
Globus Compute Docs	globus-compute.readthedocs.io
Globus Flows Docs	docs.globus.org/api/flows
AmSC Support	<i>contact your Genesis Mission liaison</i>
