SHARING DATA USING ALCF'S COMMUNITY FILESYSTEM (EAGLE)

AVANTHI MADDURI

Argonne NATIONAL LABORATORY

75 1946–2021

WHY EAGLE?

Primary use of Eagle is data sharing with the research community, using <u>Globus</u>.

Data sharing has been a long anticipated service that ALCF will soon be providing to allow PIs a storage space for collaboration.

Argonne ▲ | 75
NATIONAL LABORATORY   1946–2021

EAGLE SPECIFICATIONS

- **HPE ClusterStorE1000**
- **100 Petabytes of usable capacity**
- **Far exceeds the capacity of the theta filesystems**
- **8480 disk drives**
- **Network: HDR Infiniband**
- **Filesystem: Lustrev2.12**
- **160 Object Storage Targets**
- **40 Metadata Targets**
- **Sustained transfer rate: 650GB/s**
- **Far exceeds that of Theta filesystems at 240GB/s**
- **Primary use: Data sharing with research community using Globus**

GLOBUS NOMENCLATURE RELEVANT TO EAGLE

Guest and Mapped Collections

- https://docs.globus.org/globus-connect-server/v5.4/reference/collection/

Groups

- https://docs.globus.org/how-to/managing-groups/

Collaborators (ALCF account holders, external)

Access Manager (Proxy)

- https://docs.globus.org/api/transfer/endpoint/#managed_endpoints

More information : https://docs.globus.org/faq/

GETTING STARTED WITH EAGLE

- Eagle storage allocations will be granted to projects focused on data sharing.

- Interested users can submit an allocation request using the ALCF's Directors Discretionary Allocation Request form: **https://accounts.alcf.anl.gov/allocationRequests**

U.S. DEPARTMENT OF ENERGY Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.

Argonne NATIONAL LABORATORY | 75 1946–2021

OVERALL WORKFLOW FOR ONBOARDING ON EAGLE

1. PI requests an Eagle allocation project

2. Allocations Committee reviews and approves requests

3. A project with a unixgroup, project directory and quota is setup for the approved Eagle allocation project

4. A Globus sharing policy is created for the project with appropriate access controls

5. PI creates a **guest collection** for the project, using the Globus **mapped collection** for Eagle.
   Note: PI needs to have an active ALCF Account and will need to log in to Globus using their ALCF credentials.
   - If PI already has a Globus account, it needs to be linked to their ALCF account

6. PI adds collaborators to the guest collection. Collaborators can be ALCF users and external collaborators (who need to have Globus accounts)
   - Added with Read only or Read-Write permissions

U.S. DEPARTMENT OF **ENERGY** Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.

Argonne ▲ | 75 NATIONAL LABORATORY | 1946–2021

ACTIONS A PI CAN PERFORM

- Create, delete and share data with ALCF users and external collaborators (who have Globus accounts)

- Specify someone as a Proxy (Access Manager) for the guest collections

- Transfer data between the guest collection on Eagle and other Globus endpoints/collections

Argonne **75** NATIONAL LABORATORY 1946–2021

ACTIONS AN EXTERNAL COLLABORATOR CAN PERFORM

- Collaborator can read files from a collection *

- Collaborator can write to a collection **

- Collaborator can delete files in a collection **

* if the PI has read permissions for those files on the POSIX side

** if the PI has write permissions for those files on the POSIX side AND the collaborator is given write permissions in Globus
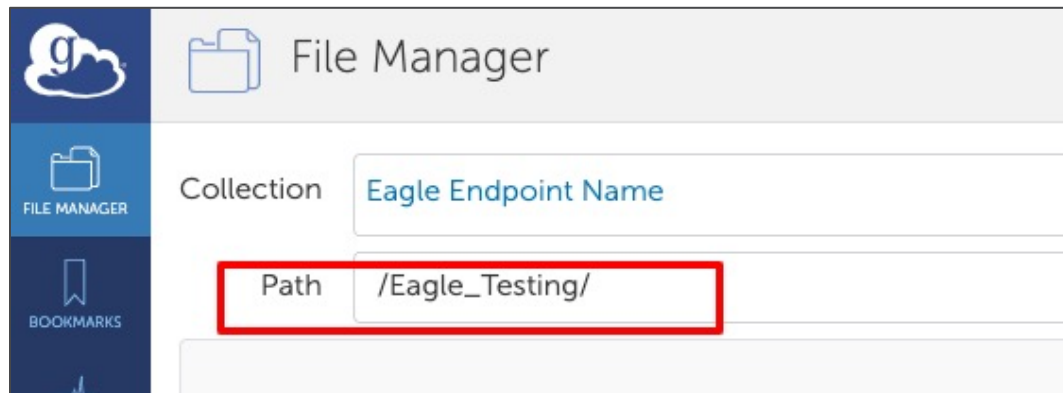
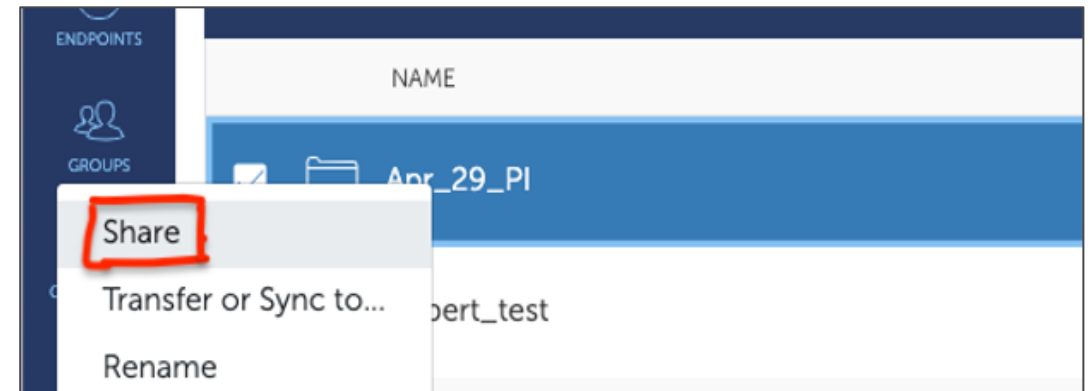# SETTING UP AND SHARING A GUEST COLLECTION
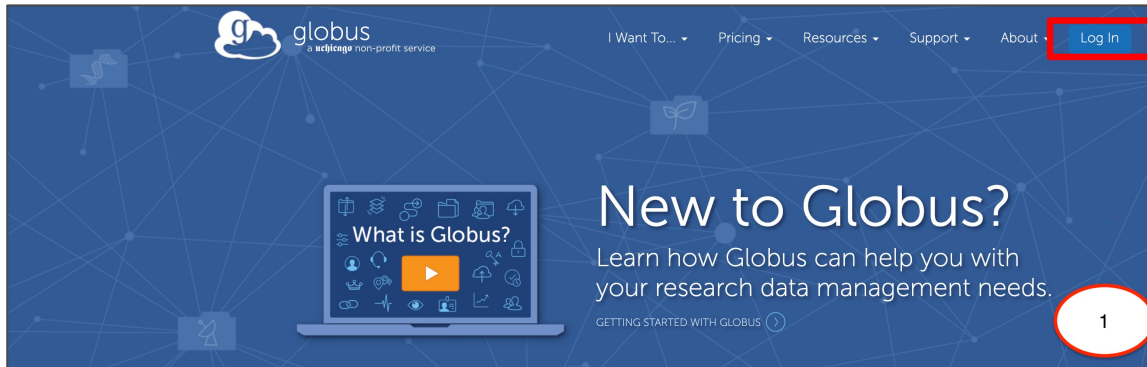
1. Logging in to Globus
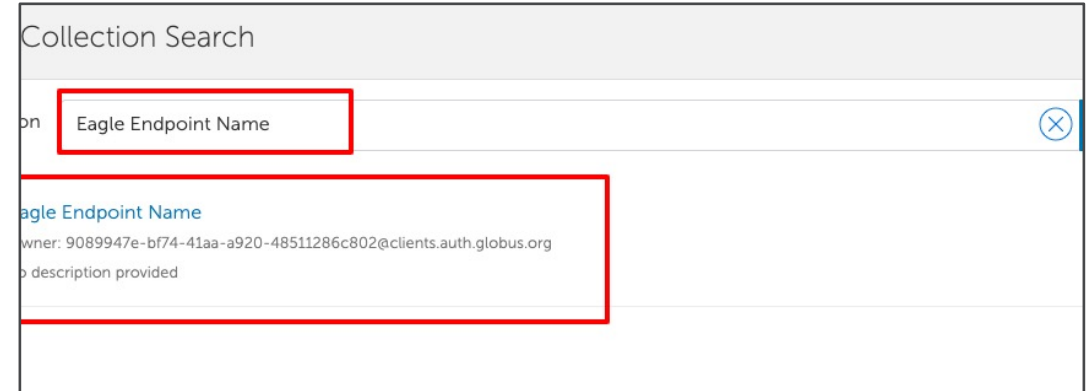
2. Searching for Eagle endpoint

3. Navigating to the specific project
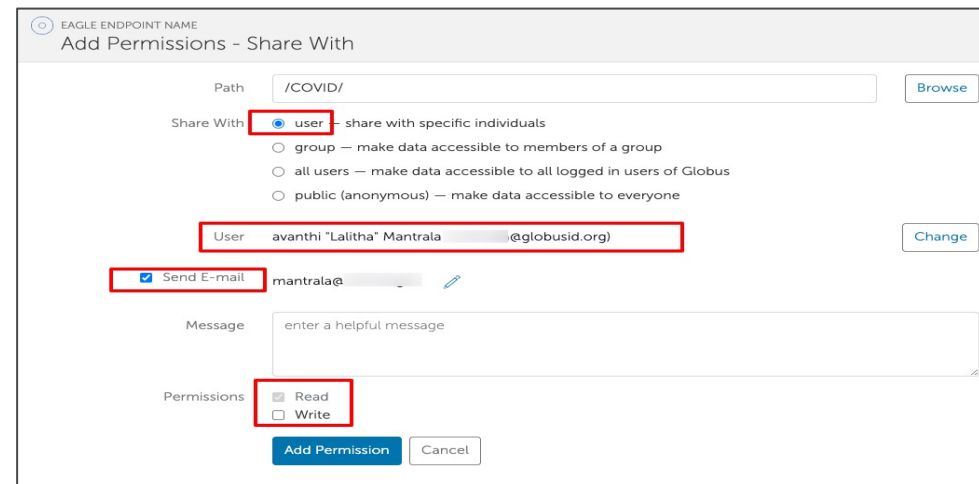

4. Selecting Transfer/Share

SETTING UP AND SHARING A GUEST COLLECTION

5. Add a Guest Collection

6. Share the collection

7. Share it with collaborators with required permissions

AND FINALLY…

8. Email notification sent to collaborators

THANK YOU FOR YOUR TIME AND ATTENTION

Argonne
NATIONAL LABORATORY | 75
1946–2021