Using Globus to transfer files to and from NYU HPC storage

The Globus project aims at providing powerful tools for scientific data management, to help researchers to focus on their domain subjects and solve data intensive research problems. Globus has been grown maturely to enable grid computing by connecting computing resources distributed globally across organizational boundary. Universities, national laboratories and computing facilities are using services of Globus.

Transferring data between endpoints

**Endpoint**
A globus **Endpoint** is a data transfer location, a location where data can be moved to or from using Globus transfer, sync and sharing service. An endpoint can either be a **personal endpoint** (on a user’s personal computer) or a **server endpoint** (located on a server, for use by multiple users). For more info see: [https://www.globus.org/how-it-works](https://www.globus.org/how-it-works)

**Data Sharing**
How to share data using Globus is described: [https://docs.globus.org/how-to/share-files/](https://docs.globus.org/how-to/share-files/)

The first step in transferring data is to get a **Globus account** at [https://www.globus.org](https://www.globus.org). Click on "Log in" at upper right corner. Select "New York University" (Step 1) from the pull-down menu and click on "Continue" (Step 2).
Enter your NYU username (NetId) and password in the familiar screen and hit "LOGIN"

The "Transfer Files" panel should come up as the following image. In order to be able to transfer files, you will need to specify the two Endpoints. One of the endpoints can be nyu#prince, the Server Endpoint on the NYU HPC storage.
Server and Personal Endpoints

The NYU HPC Server Endpoint: nyu#prince
**Globus Connect Server** is already installed on the NYU HPC cluster creating a **Server Endpoint** named **nyu#prince**, that is available to authorized users (users with a valid HPC account) using Globus. If you want to move data to or from your computer and the NYU HPC cluster, you need to install **Globus Connect Personal** on your computer, thus creating an **Personal Endpoint** on your computer.

---

**Moving data between Server Endpoints**
If you plan to transfer data between **Server Endpoints**, such as between the NYU server endpoint **nyu#prince** and a server endpoint at another institution, you do not need to install **Globus Connect Personal** on your computer.

---

**Creating a Personal Endpoint on your computer**
To create a Personal Endpoint, visit Manage Endpoints. Click the link on the upper right: add Globus Connect Personal endpoint.

More information about **Globus Connect Personal** and download links for Linux, Mac and Windows can be found at: [Globus Connect Personal](https://www.globus.org/globus-connect-personal)

---

Transfer files between your Personal Endpoint and NYU **nyu#prince**
To transfer files you need to specify two endpoints. Specify one of the endpoints as nyu\#prince. Once the endpoint is specified, you will be prompted for your username and password on Prince. Enter you NetId and password and click on Authenticate.

Your home directory content will show up, which is the same as seen on the Prince login nodes.

Set Path to /scratch/NetId
When using the nyu\#prince server endpoint, the default Path is set to the user home directory. User home directories on Prince are limited in size by enforcing disk quota (~20GB). If you plan to transfer data to the Prince cluster, we recommended that you change the Path in the Transfer Files window to your scratch directory: /scratch/NetID (see below).
Specify the second Endpoint as the personal endpoint e.g. named “My desktop at NYU”. Your desktop computer home directory content will show up. Now select directory and files (you may select multiple files when clicking on file names while pressing down “shift” key), click one of the two blue arrow buttons to indicate the transfer direction. After clicking the blue arrow button, you should see a message indicating a transfer request has been submitted successfully, and a transfer ID is generated. Globus file transfer service takes care of the actual copying.

When the transfer is done, you should receive an email notification. Click “RECENT ACTIVITY” on the Globus portal, select the transfer you want to check, a finished transfer should look like the following:
Activity

<table>
<thead>
<tr>
<th>Task ID</th>
<th>3e6bf3a-371a-11e7-bcda-22000b9a448b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owner</td>
<td>Wensheng Deng (<a href="mailto:wd35@nyu.edu">wd35@nyu.edu</a>)</td>
</tr>
<tr>
<td>Source</td>
<td>nyu@prince</td>
</tr>
<tr>
<td>Destination</td>
<td>My desktop at NYU</td>
</tr>
<tr>
<td>Condition</td>
<td>SUCCEEDED</td>
</tr>
<tr>
<td>Requested</td>
<td>2017-05-12 09:52 am</td>
</tr>
<tr>
<td>Completed</td>
<td>2017-05-12 09:53 am</td>
</tr>
<tr>
<td>Transfer Settings</td>
<td>- verify file integrity after transfer</td>
</tr>
<tr>
<td></td>
<td>- transfer is not encrypted</td>
</tr>
<tr>
<td></td>
<td>- overwriting all files on destination</td>
</tr>
</tbody>
</table>

Files: 1
Directories: 0
Bytes Transferred: 5.36 GB
Effective Speed: 82.45 MB/s
Pending: 0
Succeeded: 1
Cancelled: 0
Expired: 0
Failed: 0
Retrying: 0
Skipped: 0

Additional info can be found at this page [https://docs.globus.org/how-to/get-started/](https://docs.globus.org/how-to/get-started/). Feel free to send any question. Good luck!