Data Migration from SuperMUC to SuperMUC-NG

This document provides a step-by-step guide on how to migrate data from SuperMUC to SuperMUC-NG using Globus. Although the system also allows transferring files via scp or rsync, Globus should be the preferred solution for migrating large volumes of data (several TByte), while the aforementioned two other options should be only used for transferring a few small files.

- Transfer via Globus Online
- Hints
- Transfer via scp or rsync
- References

Transfer via Globus Online

(Note: the file permissions may change when copying files into SuperMUC-NG (where files seem to be created with the default umask). If files permissions need to be preserved, the recommendation is to create a tarball containing these files, transfer that tarball instead, and un-tar once copied.)

The provided solution is based on Globus Online, which is accessible online via app.globus.org. Notice that migrating files using other Globus tools as the globus-url-copy provided by the Globus Toolkit is not available at the moment.

Therefore, the first step is to open your favourite internet browser and navigate to app.globus.org. If your internet browser is old, it may not be supported. In this case, a message indicating this issue will be showed to you. At the time of writing, we have not experienced problems with Google Chrome, Safari or Firefox. We have experienced some troubles with Microsoft Internet Explorer, Firefox should work. We have not tested other Internet browsers.

The login interface of the Globus Online looks like the following picture.

Choose “Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften” as organisation from the list and click on “Continue”.

The previous step should redirect you to the login system of LRZ based on Shibboleth. Without going into details, this system should authorise you using your LRZ username and password. Insert your username for SuperMUC-NG (it is very important you use the username for SuperMUC-NG and not for SuperMUC in case they are different) and password and click on "Anmelden". The example depicts the login for the user di57sal.
If the login credentials are correct, the previous step redirects you to the file manager. By default, it looks like the following picture.

As you are going use the system to copy files between two locations, choose the view with two panels (maybe that view is the default for you). Changing from one to two panels can be done with the "Panels" button.
The two panels interface looks somewhat like this:

For each of the panels, you need to choose a collection using the "Collection" field of the interface (see picture below). Each collection represents a storage system from/to where you want to transfer files.
As you want to copy files from SuperMUC to SuperMUC-NG, the endpoints “LRZ SuperMUC Data Migration (RESTRICTED)” and “LRZ SuperMUC-NG Data Migration (RESTRICTED)” should be selected. When you start typing in any of the collection fields, the system would bring you to the interface for choosing collections (see the picture below). Typing LRZ would probably suffice for the system for offering you the aforementioned two collections as the system uses auto completion.

First, choose the “LRZ SuperMUC-NG Data Migration (RESTRICTED)” for the panel you like (in this tutorial we choose it for the right panel.)

Second, choose the “LRZ SuperMUC Data Migration (RESTRICTED)” for the other panel (in this tutorial we choose it for the left panel.) In this case, you should get a message in that panel indicating you need to authenticate again.

Click on "Continue", and you will be redirected to the interface for authenticating with your SuperMUC account.
In the above window login with your SuperMUC account. This way, if you have different usernames in SuperMUC, you can choose the account for which you want to migrate the data.

Once the collections are chosen you should see something like the following picture.

Some key elements in the previous interface are the *Path* fields, the *Start* buttons for triggering the transfers, and the *Transfer & Sync Options* button for accessing a more advanced options menu.
The system is pretty straightforward. On the left panel set the path from where you copy files (this path must exist in SuperMUC and your user must have access rights). Do the same on the right panel (i.e., choose the path to where you want to copy). Choose the files on the left side and click on "Start". Once clicked, the transfer will start asynchronously. You will be provided a link to check the status of the transfer. Once finished, you will receive a mail indicating its completion.

Although the previous step should suffice, we recommend several options accessible via the "Transfer & Sync Options button". In particular, the system offers the following ones.

We recommend at least to tick on "verify integrity after transfer". We do not see any need in choosing the encryption option as the data would not leave the LRZ network at any point.

**Hints**

1. Transfer of a lot of small files will be slow. For that use case, we recommend creating a tarball with all these files and transfer the tarball instead.
2. **Globus does not preserve file permissions** when performing a transfer. When you transfer files with Globus, their permissions are determined entirely by the destination endpoint’s configuration. If you want to avoid this, create a tarball including the file and transfer the tarball instead. Files within the tarball should not be affected. (For reasons see: Globus FAQ)
3. If you want to transfer files from different accounts on SuperMUC, use the button "Manage Activation" in the middle of the two transfer windows. Click on this button and indicate to deactivate the endpoint (the system will ask twice for deactivating it). Once deactivated, login again with a different SuperMUC account (as indicated before). IMPORTANT: deactivating an endpoint should never be done while there are active data transfers. Changing the account should be only done when all the requested data transfer are finished.
Transfer via scp or rsync

scp and rsync can be used for transferring small amounts of data (some GBs) from SuperMUC to SuperMUC-NG. However, the transfer must be initiated from the login nodes of SuperMUC-NG i.e., you can only pull data to SuperMUC-NG.

Examples:

SuperMUC-NG Login Node: scp -r <USER ID on SuperMUC>@hw.supermuc.lrz.de:<SOURCE PATH on SuperMUC> <DESTINATION PATH on SuperMUC-NG>
SuperMUC-NG Login Node: rsync -av <USER ID on SuperMUC>@hw.supermuc.lrz.de:<SOURCE PATH on SuperMUC> <DESTINATION PATH on SuperMUC-NG>

References

- https://docs.globus.org/
- Globus FAQs