User-Centered Usage Model

LRZ will introduce a new usage model for SuperMUC-NG. The new model was asked for in several user discussions. It will overcome well-known difficulties in using the system, such as the complexities of handling several user accounts, as well as sharing of data between project members and even between different projects. These changes will apply for SuperMUC-NG, while the usage model for the old SuperMUC will be kept unchanged for the rest of its lifetime.

The new model is user-centered, i.e. each user will have only one user account on the system. All assigned project resources can be accessed via this account. Particular attention was paid to data handling: The new model enables a user to seamlessly access his or her data in different projects for which he/she is validated. In addition to that, a user can share his or her data with other users within the same project, by giving the project group access to his or her user directory within the project.

To get this model up and running, all users will need to obtain a new single user account on SuperMUC-NG. With this new user account, the user will be able access his/her data in projects to which he/she is assigned to. All old user accounts will be phased out together with decommissioning of the SuperMUC system. The password for the new account is the one that the user has most recently changed before the target date of December 1st, 2018. The expiration date of the new password will be that of the old one.

Users are encouraged to test the assigned user account as soon as possible by logging in to https://idportal.lrz.de/r/entry.pl?Sprache=en.

Also the naming convention of the HOME, WORK and SCRATCH directories on SuperMUC-NG will change. The path to HOME will include a 2-character hash which is used to balance the I/O load.

All users will be informed by email and in the file

$HOME/README_LRZ_Migration_to_SuperMUC_NG

in all their HOME directories on SuperMUC.

The following example shows the current accounts of user “Erika Mustermann” on SuperMUC, the scheme for HOME and WORK as well as the date of the most recent password change.

<table>
<thead>
<tr>
<th>User Account</th>
<th>Project</th>
<th>HOME</th>
<th>WORK</th>
<th>Date of password change</th>
</tr>
</thead>
<tbody>
<tr>
<td>di12faq</td>
<td>pr12xf</td>
<td>/home/hpc/pr12xf/di12faq</td>
<td>/gpfs/work/pr12xf/di12faq</td>
<td>2018-05-01</td>
</tr>
<tr>
<td>di12faq2</td>
<td>pr23ys</td>
<td>/home/hpc/pr23ys/di12faq2</td>
<td>/gpfs/work/pr23ys/di12faq2</td>
<td>2018-07-02</td>
</tr>
<tr>
<td>di12faq4</td>
<td>pr12ab</td>
<td>/home/hpc/pr12ab/di12faq4</td>
<td>/gpfs/work/pr12ab/di12faq4</td>
<td>2018-02-01</td>
</tr>
</tbody>
</table>

The new user account on SuperMUC-NG for “Erika Mustermann” will be: di12faq5 and the password for SuperMUC-NG will be that of her account di12faq2, because it was most recently changed. She can inform herself about the projects which she can access on SuperMUC-NG by the command: groups

The Linux groups for WORK consist of the project-ID followed by “-c”. Other groups appear here may be used for administrative purposes, such as access to software

Example:

<table>
<thead>
<tr>
<th>User Account</th>
<th>Projects the user is involved</th>
<th>HOME and SCRATCH</th>
<th>WORK directories the user can access</th>
<th>Password expiration date</th>
</tr>
</thead>
</table>
Predefined Variables

LRZ provides predefined variables for accessing the directories

```bash
$WORK_<Project>
$SCRATCH
```

Users can select a specific WORK directory by applying the appropriate projectID e.g.,

```bash
export WORK=$WORK_pr12xf
```

in scripts or setting it in their .profile.

LRZ provides a command to assist users:

<table>
<thead>
<tr>
<th>command</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>workon info</code></td>
<td>displays information about groups, quota and budget about all projects the user is assigned to</td>
</tr>
<tr>
<td><code>source workon [-q] &lt;project&gt;</code></td>
<td>displays information about the specified project and sets the following environment variables which can be used in subsequent scripts. With -q only the variables are set. (example: <code>source workon pr12xf</code>)</td>
</tr>
</tbody>
</table>

- **WORK**: containing the path the WORK directory of the specified project (example: `/hppfs/work/pr12xf/di12faq5`)
- **DSM_CONFIG**: containing the path of the configuration for the tape archive (example: `/etc/adsm/dsm.opt_pr12xf`)
- **ACCOUNT**: to which project the cpu-hours and resource are to be accounted (example: `pr12xf`)

Batch processing and budget

Since there is no more a default project and budget which can be charged for the resources needed by a job, the user must specify a specific project for accounting during job submission, like:

```bash
sbatch --account=pr12ab $HOME/myjob
```

File systems and Tape Archive

HOME

The default size of HOME of each user is 100 GB.

We remind the user that (based on the storage systems’ characteristics) the intended use of HOME is for small files rather than large files. Large files should be stored in WORK.

As a measure of safety against technical defects, data in HOME are replicated to another system and backups to tape are regularly be performed. However this mechanism does not prevent the accidental deletion of files by users. LRZ will try to provide snapshots for HOME later-on.

Deletion of data

Data in SCRATCH are deleted when necessary, i.e. if certain capacity marks are exceeded (high water mark deletion). The oldest files in SCRATCH are then deleted first. Users who want to store data for longer time must use WORK or the Data Science Storage.

The data in the HOME directory of a user will be deleted one year after the latest involvement of the user in a SUPERMUC-NG project.

The data in WORK will be deleted one year after the end of its associated project. Project Managers will be informed before deletion is done.

One pitfall however may be that a user is not aware of the granting period of the project the data are stored in. LRZ will therefore provide tools to monitor the lifetime of the project a user is assigned to.

New Storage layer: Data Science Storage (DSS)

Additionally, LRZ introduces a new storage layer, called Data Science Storage (DSS). In contrast to user-centered $HOME, and project-centered $WORK and $SCRATCH, DSS implements a data-centric management approach. This means that the lifetime of data on DSS is not directly coupled to the lifetime of a particular compute project and that data can be shared with arbitrary LRZ-internal and external users as well using an invitation approach just as you are used to from Dropbox or Google Drive. Additionally, data on DSS can be accessed within the complete LRZ HPC ecosystem, like our HPC cloud, visualisation environment, and so on. Further details on DSS on SuperMUC-NG and its application process will be announced at a later time. For now, you can check out the generic DSS documentation at:

https://doku.lrz.de/display/PUBLIC/Data+Science+Storage

Tape Archive

Regarding the tape archive system, this will be operated in the status quo you are already familiar with SuperMUC until 2020. In 2020 we will install a successor system, and we also will change the usage model of the archive.
Data Transfer from SuperMUC to SuperMUC-NG

User must organize the transfer of data from SuperMUC to SuperMUC-NG themselves until the end of operation of SuperMUC Phase 2 at the end of June 2019. The migration process will require coordination with the corresponding project managers of the project they want to store the data in.

For transferring the data, we will shortly provide two high speed Globus Online end points on both SuperMUC and SuperMUC-NG that will allow you to copy or sync the data between the systems, using a convenient Web interface. A throughout description on how to accomplish this task will be made available shortly.

Terms and conditions

Support

If you have further questions:

- Servicedesk for SuperMUC-NG