Terms of use for the HPC systems at the ZIH
March 5, 2018

In addition to the general Terms of Use of the ZIH, and (only if necessary) diverging from the provisions of the IT-Regulations, the applicant henceforth acknowledges the following "Terms of Use for the HPC Systems at the ZIH", which focus on the handling of user data on high-performance computers.

The high-performance computers of the ZIH are highly specialized research devices that are operated by the ZIH for various user groups. In order to ensure trouble-free operation and efficient maintenance of the systems, the following conditions are recognized:

§1 - Storage of Personal Data in the HPC Project System

The HPC project system supports the ZIH staff in day-to-day operations (e.g. project management, user support, statistics). In addition to project-related data, personal contact data of HPC users (surname, first name, title, office e-mail address, office telephone number, and office postal address) are also stored. HPC support can use these communication channels for the following purposes:
- Postal sending of confidential information (e.g. access data),
- Notification of failures, maintenance or training via e-mail,
- Inquiry support for user-specific problems during operation (e.g. impairment of other users) by e-mail or telephone.

The period for storing this data in the HPC project system is governed by §4 (4) of the "Regulations for the Establishment and Operation of an Identity Management System at the Dresden University of Technology". At the end of this period, the personal data of the users will be made anonymous in this system.

Descriptions of the project, the proof of the resources used (accounting), as well as the contact data of the project manager are excluded from this. These are permanently stored and used for scientific purposes and to ensure the effective use of high-performance computers of the Gauß-Allianz.

§2 - Work in the Project

The HPC systems may only be used for research purposes and within the scope of the project applied for. Any other or deviating use is not permitted.

The project leader is usually a professor or head of a non-university research group. The **project leader** bears the overall responsibility for the project as well as for all activities within his project on the HPC systems of the ZIH. The project leader can appoint a **project administrator** who is responsible for the technical and administrative matters of the HPC project. The project leader and the project administrator are the ZIH's contacts for the entire project group.

In particular, they are authorized to:
- Add additional users (special legal regulations for foreign citizens may need to be taken into account!)
- Update members' contact details when necessary
- Block HPC access for employees who have left the working group
- Monitor the required resources of the project
- View a user's data in the project directory and save data worthy of backup before the user leaves the project

To ensure efficient operation of the HPC systems, and for general and statistical evaluation, data on the utilization of resources (CPU, RAM, file system, software, energy) of each job are automatically recorded (accounting). The following applies:

- Project managers and administrators have access to an overview of resource consumption in their respective project.
- In order to be able to coordinate accordingly, the resource consumption is specified by jobs of the individual users.
- The ZIH only uses a project-related accounting view vis-à-vis the scientific advisory board of the ZIH, the CIO of the TUD and its advisory board, or other comprehensive committees.
- Within a period according to §4 (4) of the "Regulations for the Establishment and Operation of an Identity Management System at the Dresden University of Technology", the associated personal accounting data will be made anonymous in the project system.

No new jobs can be started at the end of a project runtime. For the purpose of securing important research data, the access rights for members of this project can be used for another 30 days before they are deactivated. 100 days after the end of the project, the data assigned to the project is deleted from the HPC file systems.

§3 - Handling of Data

For the scientific work on the high-performance computers, different file systems with different properties are provided. Temporary, frequently changing files, and directories should not be located under /home/ or /projects. Scratch file systems are available especially for this purpose. Archives at the ZIH can be used for the long-term data storage. Because of the hierarchy of storage systems in the HPC realm, these specifics are followed:
- The HOME directories of the users under /home and the project directories under /projects are automatically backed up to a tape system. The retention period here is 180 days.
  Additional snapshots within the file system facilitate access to older files.
- The temporary data file systems (/tmp, /scratch, /lustre/ssd) are not backed up. Files stored there are automatically deleted after a reasonable period of time if necessary. This happens on local disks (/tmp) after 7 days (login node) or immediately after job end (compute node), in /scratch after 100 days, in global SSD scratch (/lustre/ssd) after 50 days.

The ZIH assumes that all data in the file systems of the HPC machines are project-related data. This results in the following rules:
- The project manager or the project administrator is granted access to read and write project directories of the project members. This also applies expressly to directories of already retired users. The HOME directories are excluded from this.
- Users and project managers/project administrators are obliged to review the project data and save data worthy of backup before a user leaves the project. For this purpose, the storage services of the ZIH (long-term archiving of research data, intermediate archive for unstructured data) can be used.
- 100 days after the expiration of the HPC access authorization, the user data on the HPC file systems are automatically cleaned up. Data in the HOME directory of the former user will be deleted. The data of the former user under /projects is transferred to the project administrator.

§4 - Problems and Malfunctions in Operation

HPC systems are susceptible to usage errors. HPC Support is authorized to cancel or stop jobs that interfere with normal operation. In this case, the user will be informed and asked for troubleshooting participation. The user is responsible to follow the instructions of HPC Support. In case of non-compliance, the user can be excluded from using the batch system or his login can be blocked.

All log data in the HPC context is recorded in a database and used for long-term analyses. After one month, messages with a login reference are deleted.

§5 - Recording Performance Data for User Jobs

For research purposes, job- and login-related performance data are stored and evaluated in a database. These include metrics from processors, networks, storage systems and the power supply. This data can be passed on to HPC research partners in an anonymous form.
HPC support has access to performance data to identify optimization potential and inform users about it.