Data Analysis Service (142-004)

Data analysis infrastructure for the multi-disciplinary users of PSI large scale facilities

PSI develops, constructs and operates complex large-scale facilities, and every year more than 2000 guest scientists from Switzerland, as well as from around the world, come to PSI. PSI's large scale facilities support fields as varied as physics, chemistry, biology, material sciences, energy technology, environmental science, medical technology and cultural heritage. Applications are numerous, for example, crystallography can reveal the structures of viruses and proteins important for the development of new drugs; neutron scattering can identify stresses within engineering components such as turbine blades, and X-ray phase-contrast tomography can image microscopic details of the 3D-structure of the brain. Commercial users include the pharmaceutical, petrochemical and microelectronic industries. Beam time is allocated based on a peer-reviewed process with scientific excellence being the selection criterion. Recent developments, particularly in the field of X-ray detectors, yield unprecedented data quality, speed and sample throughput, thus allowing for new type of experiments including e.g. in-situ time-resolved applications.
Today, data are taken at PSI and then transferred to the users' home institutions for further analysis. With the new generation of detectors, producing data volumes in the Petabyte range, this model has severe limitations.

Therefore PSI aims to provide a new service to the scientists in form of a *data analysis center*.

The online analysis, providing am immediate feedback to users while data are recorded, is provided by PSI and regularly expanded. The proposed new infrastructure covers the following analysis steps. It allows storage and analysis of the experimental data directly after data taking and will allow to pursue subsequent offline data analysis on site and remotely from anywhere within Switzerland. This will result in more efficiently pursued scientific projects and reduces thereby the time to the publication of the data or will in some cases even be the factor that enables analysis and publication at all.

This infrastructure can be combined with other national or international services, particularly in the area of identity management, meta-data handling and generic e-science services. It can therefore be a crystallization point for a network of similar services at other institutes.

In a second phase of the present project, a co-location of storage and compute resources in an IT infrastructure commonly used within Switzerland will be evaluated, provided such services emerge within this program. In any case PSI is committed to provide this service in a sustainable manner.