



CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

Lugano, 11 July 2018

CSCS will store petabyte data for the Paul Scherrer Institute

In the future, research data collected by the large-scale research facilities at the Paul Scherrer Institute (PSI) in Villigen will be archived at the Swiss National Supercomputing Centre (CSCS) in Lugano. Collaboration between PSI and CSCS enabled major improvements to the data transfer and storage process.

Storing and archiving data is part of everyday life at CSCS. However, transferring huge amounts of data — including ongoing research projects online across Switzerland as well as an entire data archive containing decades of major completed projects — via fibre-optic cable presents a logistical challenge. That was made clear by a recent project between the Paul Scherrer Institute (PSI) in Villigen and CSCS in Lugano. PSI and its researchers have been using supercomputers at CSCS, part of ETH Zurich, for around 15 years. The institute will now also store its data at CSCS — specifically data generated by the large-scale research facilities of the new SwissFEL X-ray laser and Swiss Lightsource (SLS). A total of 10 to 20 petabytes of data is produced every year. For comparison, if you wanted to save a petabyte of data on DVDs, you would need more than 200,000 of them.

Secure data transfer

Transferring the data generated at PSI, which comes directly from experiments, was enabled by the recent successful completion of a project between employees of CSCS and the PSI IT department. PSI and CSCS worked together to develop a process by which data could be securely transferred, archived, retrieved and deleted after the end of the archival period — a kind of guide for how data, much like parcels, must be packed for secure dispatch. In these “packaging regulations”, they integrated an intelligent workflow manager into the data transfer, storage and migration process to coordinate the procedure. They also incorporated a data catalogue into the archiving process that systematically lists the data and thus makes it easier to access. Using the specially-developed, dedicated redundant network connection between PSI and CSCS, 10 gigabytes of data can now be transferred per second. For CSCS and PSI, this new solution is a significant contribution towards fulfilling growing demands in the area of open research data.

Those who set up the project are very happy with its successful conclusion. “For us, this project is a logical development of the services that we offer our users,” says ETH professor Thomas Schulthess, Director of CSCS. “With our computing infrastructure at CSCS, we also possess the necessary resources to analyse this data if necessary. Keeping data storage, analysis and processing close together saves time and energy.”

PSI is also pleased with the successful collaboration. “Carrying out data-intensive research projects in the era of big data makes it logical for us to transfer our data to CSCS and reduce costs compared to a local solution,” says Dr Gerd Mann, Head of IT at PSI. “The collaboration with CSCS has allowed us to focus our IT resources at PSI on supporting scientific data management and processing.”

At CSCS, a tape library with a current storage capacity of 120 petabytes is available for users’ data archiving. By 2022, PSI will transfer around 85 petabytes of data to CSCS for archiving. Around 35 petabytes come from SwissFEL experiments, and 40 come from SLS. The storage



CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

capacity can be increased to 2,000 petabytes if needed over the course of the next five years, in order to meet the growing need for data storage and archiving in Switzerland.

Photos:

<https://www.cscs.ch/publications/press-releases/2018/589/>

Contact CSCS

Dr. Valentina Tamburello, Communications Officer

Phone: +41 78 964 36 40

Email: communication@cscs.ch

Dr. Michele De Lorenzi, Deputy Director

Phone: +41 91 610 82 08

Email: michele.delorenzi@cscs.ch

Contact PSI

Dr. Gerd Mann, Head IT PSI and Head IT Department

Phone: +41 79 799 62 98

Email: gerd.mann@psi.ch