



CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

Lugano, 11 luglio 2018

Memoria “petabyte” per l’Istituto Paul Scherrer presso il CSCS

In futuro, i dati raccolti dai grandi impianti di ricerca dell’Istituto Paul Scherrer (PSI) di Villigen saranno salvati e archiviati presso il Centro svizzero di calcolo scientifico (CSCS) di Lugano, grazie a una collaborazione fra i due istituti.

Salvare e archiviare dati rientra nella normale amministrazione del CSCS. Tuttavia, trasferire da un punto all’altro della Svizzera attraverso cavi in fibra ottica un intero archivio di dati ultradecennale di grandi progetti di ricerca, passati e presenti, diventa una vera e propria impresa logistica. È quanto dimostra un progetto dell’Istituto Paul Scherrer di Villigen e del CSCS di Lugano. Il PSI e i suoi ricercatori utilizzano da circa 15 anni il supercomputer del CSCS, costola del Politecnico federale di Zurigo (ETH Zürich). Ora il PSI ha deciso di archiviare presso il CSCS anche i suoi dati, generati dai grandi impianti di ricerca del nuovo laser a raggi X SwissFEL e da Swiss Light Source (SLS). Complessivamente vengono trattati ogni anno fino a 20 petabyte di dati. Per intenderci: se si volesse salvare un petabyte di dati su DVD, ne servirebbero 200’000.

Trasferimento dei dati sicuro

Il trasferimento dei dati generati presso il PSI come risultato diretto degli esperimenti scientifici avviene grazie a un progetto, curato e portato recentemente a termine con successo dai collaboratori del CSCS e dal reparto IT del PSI. A questo scopo PSI e CSCS hanno elaborato un processo con cui trasmettere, archiviare, ritrovare ed eliminare dati in tutta sicurezza: una specie di manuale d’uso su come confezionare dati per un invio sicuro, alla stregua di un pacchetto da spedire. Nelle apposite “istruzioni di imballaggio” hanno integrato nel processo di trasferimento, memorizzazione e migrazione dei dati un workflow manager intelligente (un gestore del flusso di lavoro) con il compito di coordinare l’operazione. Contemporaneamente il processo di archiviazione è stato dotato di un sistema in grado di catalogare sistematicamente i dati, semplificando, quindi, l’accesso agli stessi. Mediante un collegamento di rete ridondante, allestito appositamente a questo scopo, tra PSI e CSCS è ora possibile trasmettere 10 gigabit di dati al secondo. Per CSCS e PSI questa nuova soluzione rappresenta un contributo essenziale per soddisfare i crescenti standard in ambito di Open Research Data.

Gli iniziatori del progetto sono soddisfatti del suo esito positivo. «Questo progetto è per noi il logico perfezionamento dei servizi che offriamo ai nostri utenti», spiega Thomas Schulthess, professore all’ETH, nonché direttore del CSCS. «Al CSCS, inoltre, grazie alla nostra infrastruttura di ricerca, disponiamo delle necessarie risorse anche per analizzare all’occorrenza questi dati. Raggruppare le attività di memorizzazione, analisi e trattamento dei dati si traduce in un risparmio di tempo ed energia». Anche il PSI è soddisfatto della buona riuscita di questa cooperazione. «Nell’era dei big data, alla luce dei progetti di ricerca ad alta densità di dati, per noi trasferire i dati al CSCS è una logica conseguenza che consente un risparmio di costi rispetto a una soluzione locale», spiega il Dr. Gerd Mann, direttore generale IT al PSI. «Grazie alla cooperazione con il CSCS, le risorse IT del PSI possono concentrarsi sul sostegno della valutazione scientifica dei dati e su nuovi progetti».

Il CSCS mette a disposizione degli utenti una Tape Library per l’archiviazione dei dati con una capacità di memoria pari a 120 petabyte. Entro il 2022, il PSI trasferirà circa 85 petabyte di dati al CSCS, in cui verranno archiviati. Nello specifico, circa 35 petabyte risultano dagli esperimenti



CSCS

Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

del SwissFEL e 40 petabyte da SLS. Nel corso dei prossimi cinque anni, se necessario, la capacità di memoria sarà aumentata fino a 2000 petabyte, in modo da poter soddisfare il crescente fabbisogno in Svizzera di memorizzazione e archiviazione dei dati.

FOTO:

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

Contatto CSCS

Dr. Valentina Tamburello, Communications Officer

Telefono: +41 78 964 36 40

Email: communication@cscs.ch

Dr. Michele De Lorenzi, Deputy Director

Telefono: +41 91 610 82 08

Email: michele.delorenzi@cscs.ch

Contatto PSI

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

Telefono: +41 79 799 62 98

Email: gerd.mann@psi.ch