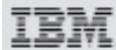


A university research institute reaches new levels of performance with IBM System Storage DS4700, GPFS and Tivoli Storage Manager



IBM Client reference



The client: Bioinformatics at Biozentrum, University of Basel

The Bioinformatics at the Biozentrum is part of the University of Basel. The Biozentrum is a research institute that specializes primarily in biochemistry, biophysical chemistry, microbiology, structural biology, cell biology, pharmacology and neurobiology. It is dedicated to excellence in research and teaching and is focused on the



investigation of biological nanomachines and communication systems in health and disease.

The Bioinformatics research demands powerful and complex computing capabilities, for example, to simulate and model three dimensional structures of proteins, and has the needs for flexibility to adapt to the

different project requirements. The existing compute and storage infrastructure was no longer flexible enough and could not scale to the increasing requirements of new research projects. The client has a large variety of applications and relies heavily on open source applications and self-written codes such as Perl, Python, C, C++ or Java.

To meet the different needs in a complex research environment and to manage the exponential growth of data, Bioinformatics at the Biozentrum recently implemented several new IBM System Storage DS4700 and System x3650 Servers with GPFS and IBM Tivoli Storage Manager software for Backup and Space Management.

Highlights

The replaced and expanded infrastructure now supports new and complex experiments in Bioinformatics research while using more than 150 different bioinformatics applications, databases, self-developed software and scientific pipelines with completely new levels of convenience and system availability.

Client challenge

When a Bioinformatics research group decided to begin a new project that required 10TB of disk space, the infrastructure team did not require only high-speed Fibre Channel/Serial Advanced Technology Attachment (FC/SATA) disks, but also a clustered file system based on General Parallel File System (GPFS).

At that point the client's infrastructure included a small cluster of IBM System x servers, cluster nodes in a mixed GPFS and a network file system (NFS) environment which had to be expanded.

In addition, the institute wanted to add a hierarchical storage management (HSM). Up to 50 % of the data in the existing storage was older than 1 year. By improving storage management and moving older data to a second and third storage tier, the Bioinformatics at the Biozentrum expected to reduce the volume of data on its primary GPFS storage system.

The solution

To replace and expand its infrastructure, the client implemented 10 new IBM System x3650 servers together with 12 System Storage DS4700 and EXP810 which multiplied the available disk space to 70TB. GPFS spans over all disk arrays to form a unified system where filesystem spaces can be easily created and transparently moved between different disk systems and types. Distributing data over many disks increases performance needed for the compute cluster where 750 CPU Cores access the same filesystem for data processing. Main applications are the institute's scientific pipelines and processing steps of raw data acquired devices such as deep sequencing or automated microscopes.

The Bioinformatics also implemented IBM Tivoli Storage Manager 6.1 which has native accesses to GPFS and provides a faster backup. Additionally, Tivoli space Manager for Space Management is used to integrate tape technology into the data lifecycle management of GPFS.

The upgrade of the storage and backup environment constitutes the final step in an overall renewal and expansion of the Bioinformatics research storage and compute infrastructure at the Biozentrum.

The benefits

By implementing the new System x, System Storage hardware, GPFS and Tivoli software, the Bioinformatics at the Biozentrum created a powerful, flexible and stable GPFS environment that meets the various demands of a complex university research environment. The solution allows the institute to support and manage its exponential data growth, while the load-balancing, high-availability and failover features of the GPFS infrastructure enables system maintenance and updates without an interruption of service. The Tivoli solution helped to improve data lifecycle, management and protection for research critical data while ensuring rapid recovery time, data reliability and integrity.

Contact:

Daniel Jehle
Senior Client Representative
Public Sector Switzerland
Phone: 41-58 333 83 54
Fax: 41-58 333 83 54
Email:
daniel.jehle@ch.ibm.com



© Copyright IBM Corporation 2009 Alle Rechte vorbehalten

IBM und das IBM Logo sind eingetragene Marken der International Business Machines Corporation in den USA und/oder anderen Ländern.

Marken anderer Unternehmen/Hersteller werden anerkannt. Vertragsbedingungen und Preise erhalten Sie bei den IBM Geschäftsstellen und den IBM Business Partnern. Die Produktinformationen geben den derzeitigen Stand wieder. Gegenstand und Umfrage der Leistungen bestimmen sich ausschliesslich nach den jeweiligen Verträgen.

Die vorliegende Veröffentlichung dient ausschliesslich der allgemeinen Information.