



The University of Bonn boosts genetics research

With a high-performance cluster computing solution from IBM

Overview

The need

To gain new insights into medical genetics, a research Institute within the University of Bonn needed to replace its high-performance computing platform, which had reached the end of its useful life.

The solution

Deployed a new Linux cluster on IBM BladeCenter® with LS42 and HS22 servers; implemented IBM General Parallel File System (GPFS™), two IBM System Storage® DS3400 disk arrays and an IBM System Storage DCS3700.

The benefit

Gained a compact and efficient solution that is the right fit for the Institute and offers extremely high performance for complex computational tasks. The solution is also highly reliable and scalable.

The Institute of Medical Biometry, Informatics and Epidemiology at the Faculty of Medicine in the University of Bonn is a world-leading research team, working in the field of medical genetics research. The Institute makes significant use of computational models and carries out complex statistical and analytical calculations on large data sets. To help gain new insights into important research topics, the institute had set up a small cluster solution with 14 nodes. As the number and scope of projects and the amount of data grew by orders of magnitude, this solution became increasingly outdated. The Institute looked for a cost-efficient yet powerful computational platform that would enable more complex operations at higher speed.

With limited space available in its server room, the Institute required a high-performance computing system with a small physical footprint. The second key requirement was for high-speed connectivity between the computational cluster and the data storage environment. Finally, the Institute needed a highly scalable file system that could handle the enormous—and growing—amounts of data that genetics researchers expect to be able to manipulate.

The institute assessed various solutions, then turned to IBM for assistance. “The only file system we knew on the market that would be capable of scaling as required was IBM GPFS [General Parallel File System],” says Waldemar Spitz, System Administrator at the Institute of Medical Biometry, Informatics and Epidemiology. “After previous positive experience with IBM in high-performance computing, we were convinced that they would again come up with the best solution for us.”



“The only file system we knew on the market that would be capable of scaling as required was IBM GPFS.”

—Waldemar Spitz, System Administrator at the Institute of Medical Biometry, Informatics and Epidemiology, the University of Bonn

High-performance architecture from IBM

IBM proposed a solution based on the IBM BladeCenter H chassis with IBM LS42 blade servers, each equipped with four six-core processors, and IBM HS22 blade servers based on the next generation of intelligent server processors, the Intel® Xeon® processor 5600 series, featuring industry-leading performance and energy efficiency with the ability to intelligently adapt to the current workload.

To satisfy the institute’s networking and storage performance needs, IBM recommended IBM System Storage DS3400 devices, connected via Fibre Channel to IBM System x3650 M2 servers, set up as IBM General Parallel File System I/O servers, which are linked via InfiniBand connections to the high-performance computing cluster.

“Together with IBM, we developed a high-performance computing concept that will provide enough processing power and a scalable storage infrastructure to drive our research projects for the next few years,” says Waldemar Spitz. “The performance and scalability of IBM GPFS were key factors in our decision.”

Rapid deployment

The Institute deployed a total of 26 IBM BladeCenter HS22 servers and eight IBM BladeCenter LS42 servers, providing 34 cluster nodes with 504 processor cores in total, supported by 1.8 TB main memory. Two IBM System Storage DS3400 arrays provide a total of 40 TB of storage capacity for the IBM BladeCenter cluster—and the Institute is actively adding more capacity.

As the institute has used the SUSE Linux Enterprise Server operating system for more than 15 years, this was also the operating system of choice for the new IBM cluster solution. Additionally, IBM implemented the comprehensive cluster management and provisioning software Extreme Cloud Administration Toolkit (xCAT), helping the institute to improve cluster management and system administration.

The new cluster solution was set up and configured within two months, and was seamlessly integrated into the existing IT landscape without any issues. Thanks to the efficient and compact design of IBM BladeCenter, the Institute was able to adopt the new technology without putting pressure on the limited resources of its server room.

“The BladeCenter is a well-designed platform, and the concept of shared cooling and networking creates an efficient and compact solution,” says Waldemar Spitz.

Solution components

Hardware

- IBM® BladeCenter® H
- IBM BladeCenter HS22
- IBM BladeCenter LS42
- IBM System x®3650 M2
- IBM System x3550 M2
- IBM System Storage® DS3400
- IBM System Storage EXP3000
- IBM System Storage DCS3700
- IBM System Storage DCS3700 Expansion Unit
- Intel® Xeon® processor 5600 series

Software

- IBM General Parallel File System (GPFS™)
- Extreme Cloud Administration Toolkit (xCAT)
- SUSE Linux Enterprise Server

Services

- IBM Global Technology Services
-

Exploiting the latest technology

Today, approximately 20 researchers in two teams use the computing resources provided by the IBM solution. The Institute also makes the resources available to guest researchers from around the world at certain times. New research enabled by the IBM cluster covers all aspects of genetically complex diseases and population genetics.

The researchers not only benefit from the massively increased performance, but also experience fewer interruptions. “The new IBM cluster solution runs in a more stable manner, especially under high load,” says Waldemar Spitz. “In the past, we noticed more interruptions when processing jobs were running on the old cluster for months.”

The cluster solution from IBM has also made management much easier for system administrators. In the past, the team used customized scripts to manage its cluster. Today, the Extreme Cloud Administration Toolkit provides standardized tools and methodologies for cluster administration, increasing the systems management efficiency at the Institute. This is a particularly important benefit, because it frees up researchers from routine administration and enables them to focus on science.

“We are very happy with the highly competent hardware support team from IBM,” says Waldemar Spitz. “If any issue arises, they provide assistance rapidly, helping us to improve systems availability.”

Valuable research

The IBM BladeCenter cluster processes significant volumes of useful medical research data, evidenced by the fact that the Institute is now expanding its storage environment with an IBM System Storage DCS3700 array and an IBM System Storage DCS3700 Expansion Unit to 400 TB net capacity. Designed for applications with high-performance streaming data requirements, the IBM DCS3700 offers optimal space utilization, low power consumption and high performance. With up to 60 SAS drives in just 4U of rack space, it can reduce operational costs for capacity-intensive applications. And with up to 4000 MB/s in sustained drive reads, the IBM DCS3700 storage system is equally adept at delivering throughput to bandwidth-intensive applications.

The use of GPFS will make it easier for the Institute to continue expanding the storage environment, as the file system has practically unlimited scalability in terms of today’s hardware. Looking to the future, if the Institute chooses to expand the computational facilities also, it will be able to do so quickly and easily by simply slotting additional blades into the BladeCenter chassis.

For more information

Contact your IBM sales representative or IBM Business Partner, or visit us at: ibm.com/systems/bladecenter

For more information about the University of Bonn visit: uni-bonn.de



© Copyright IBM Corporation 2012

IBM Corporation
Systems and Technology Group
Route 100
Somers, New York 10589

Produced in the United States of America
February 2012

IBM, the IBM logo, ibm.com, BladeCenter, GPFS, System Storage, and System x are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at ibm.com/legal/copytrade.shtml

Intel, the Intel logo, Xeon and Xeon Inside are trademarks of Intel Corporation in the U.S. and/or other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions. It is the user's responsibility to evaluate and verify the operation of any other products or programs with IBM products and programs.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

Statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Actual available storage capacity may be reported for both uncompressed and compressed data and will vary and may be less than stated.



Please Recycle

