

SUNY Buffalo's Bioinformatics Research Accelerates With The World's Largest Academic HPC Linux Cluster

Overview

- SUNY required massive computer power to support computationally intensive biological research for its new Center of Excellence in Bioinformatics in Buffalo.
- Dell's HPC Linux clusters provided SUNY with a fast, reliable, and scalable alternative to costly supercomputers.

Challenge

- SUNY had a fixed budget, ruling out the option to purchase supercomputers.
- Dr. Skolnick, the Center's Director, wanted the ability to grow the processing capability of the system as larger problems were tackled.

Solution

- Platform LSF® software
- 1,900 Dell PowerEdge 1650 servers and 100 PowerEdge 2650 servers

Results

- Platform LSF enables more than 5 trillion complex calculations and exhaustive comparisons per second accelerating the center's discovery initiatives from years to months.
- SUNY researchers can accomplish work previously reserved for multi-million dollar supercomputers and mainframes, at a fraction of the cost.

The new Center of Excellence in Bioinformatics at the University of Buffalo—a campus of The State University of New York (SUNY)—required massive processing power to support its planned research in computational biology. As a facility entirely funded by the State of New York and the City of Buffalo, the cost of supercomputers, which are traditionally used to power complicated and data intensive discovery projects, were prohibitive. Fortunately, high performance computing (HPC) clusters built with commodity components were a cost-effective and extremely reliable alternative to deliver the processing power required.

Due to the large amount and complex nature of the data generated by The Center of Excellence in Bioinformatics' researchers, the Center at SUNY required a fast, reliable, and scalable HPC cluster that could accommodate at least 10 TB of storage to perform millions of protein-folding simulations and calculations.

Dell designed and implemented an HPC cluster of 2,000 dual-processor Dell PowerEdge servers and deployed Platform LSF to automate and manage the complex computational workloads across the Linux clusters. Today, the server cluster, which is capable of performing more than 5 trillion complex calculations and comparisons per second, enables the Center at SUNY to accelerate its discovery initiatives.

Business Challenge

Given the Center's limited capital expenditure budget, Dr. Jeffrey Skolnick, director of the Center of Excellence in Bioinformatics at SUNY, required a fast, reliable, and scalable alternative to the costly supercomputers that are typically deployed in life science environments. According to Dr. Skolnick, the price/performance ratio of low-cost, industry standard servers deployed in a clustered environment would deliver the high levels of computing power required to process the terabytes of data involved in computational biology.

Although the Center at SUNY could not operate without the 'virtual supercomputer', the massive processing power would be ineffective without a software solution to automatically distribute and manage application workload across the planned server cluster. After conferring with scientists at the Center for Computational Research (CCR) at the University of Buffalo, a facility which supports computationally intensive research for about 50 research groups at SUNY, Dr. Skolnick came to the conclusion that the open source workload management solutions used by CCR would not scale to meet the demands of the planned 2,000-node server cluster. Additionally, an open source product would not provide the level of support needed both during and after the implementation.

"Platform LSF's unlimited scalability and ability to automate and manage the complex computational workloads across our Dell Linux cluster has enabled our researchers to exploit and analyze diverse types and sources of biological data in a shorter period of time, and at a fraction of the cost that would otherwise be possible."

Dr. Jeffrey Skolnick

Director,

The Center of Excellence in Bioinformatics at SUNY

Solution

To address its requirement for a fast, reliable, and scalable HPC cluster, the Center of Excellence in Bioinformatics at SUNY deployed more than 2,000 Dell PowerEdge servers (with dual Intel processors) running Red Hat Linux. The key, however, to cluster operation is scheduling. Although other open source solutions had been investigated by the Center at SUNY, Platform LSF offered unlimited scalability - Platform LSF can support over 100 clusters, more than 2,000 CPU's and 500,000 active jobs - to address the Center's current and future workload and cluster requirements.

The Center at SUNY deployed 4,000 Platform LSF licenses for the important task of maximizing the full potential of its extensive HPC Linux cluster. Today, Platform LSF provides the Center's researchers with on-demand access to the 2,000 server nodes, and automatically prioritizes and balances the high volumes of complex computational workloads.

By combining the HPC Linux cluster with Platform LSF, more than 5 trillion complex calculations and exhaustive comparisons can be performed per second, accomplishing work that was previously reserved for multi-million dollar supercomputers and mainframes, at a fraction of the cost. According to Dr. Skolnick, the amount of data to be analyzed would take approximately 2,000 years to analyze on a single computer with one processor. By using the cluster, he expects to complete the same data analysis in just six months.

Given the magnitude and computing capability of the Center's HPC Linux cluster, it is acknowledged today as one of the largest Linux-based clusters dedicated to bioinformatics, one of the largest ever at a U.S. educational institution, and is ranked 22nd out of the top 500 Supercomputing sites in the world according to the Linpack Top500 benchmark (a global ranking, based on the results of a computer's actual performance against a common benchmark).

Platform LSF, a key component of the Dell-powered HPC Linux cluster, is an important tool in furthering the Center at SUNY's vision to enable the development of new medical treatments through use of parallel and grid computing techniques.

Customer Site

The Center for Computational Research (CCR) at the University of Buffalo (www.bioinformatics.buffalo.edu)

The Center of Excellence in Bioinformatics at the University of Buffalo is a newly established campus of The State University of New York (SUNY). The Center merges high-end technology, including supercomputing and visualization, with expertise in genomics, proteomics, and bioimaging, to enable major contributions in science and health care. Through a combination of computational science and high-throughput experimental biology, the Center can enable new medical treatments by developing and exploiting state-of-the-art algorithms for data acquisition, storage, management, and transmission, as well as innovative parallel and grid computing techniques.



About Platform Computing

Platform Computing's intelligent, practical enterprise grid software solutions help organizations optimize IT resources to Accelerate Intelligence™. We plan, build, run and manage grids that link IT to core business objectives, and help our customers improve service levels, reduce costs and enhance business performance. With industry-leading partnerships and a strong commitment to standards, we are at the forefront of grid software development, propelling over 1,600 clients toward powerful insights that create real, tangible business value. For more please visit www.platform.com.



World Headquarters

Platform Computing Inc.
3760 14th Avenue
Markham, Ontario
L3R 3T7 Canada
Tel: 905 948 8448
Fax: 905 948 9975
Toll-free tel: 877 528 3676
info@platform.com

United States

Boston: 781 685 4966
Detroit: 248 359 7820
Reston: 703 251 4850
Newport Beach: 949 798 5654
New York: 212 672 1770
San Jose: 408 392 4900

Asia-Pacific

Beijing: +86 10 62381125
Hong Kong: +852 2869 5687
Tokyo: +81 3 5326 3105

Europe

Düsseldorf: +49 (0) 2102 610390
London: +44 (0) 1256 370500
Paris: +33 (0) 1 41 10 09 20

For more information, visit www.platform.com/contactus

Copyright © 2005 Platform Computing Corporation. ® TM Trademarks of Platform Computing Corporation.
All other logo and product names are the trademarks of their respective owners, errors and omissions excepted. Printed in Canada. Platform and Platform Computing refer to Platform Computing Inc. and each of its subsidiaries.