Blueprint Initiative Samuel Lunenfeld Research Institute, Mount Sinai Hospital

Sun Center of Excellence for Systems Biology

Institution/Organization

Lunenfeld Research Institute, Mount Sinai Hospital, Toronto

Vertical Market
Higher Education / Research

Key Challenges
• How to optimize network performance to link a number of large grids from around the world;
• Building a system with the ability to scale from a beginning level of about 25,000 protein interactions to some 200,000 interactions.

Solution
• To effectively run the Bimolecular Interactions Network Database (BIND), Sun developed a solution using Sun Fire servers, Sun Ray Ultra-Thin Clients, Sun StorEdge, Java Enterprise System software and the Solaris Operating Environment.

Business Results
• Blueprint North America's BIND database boasts the highest number of protein interactions on one database worldwide.
• BIND has accumulated more than 25,000 protein interactions and is emerging as the industry's primary source for protein data.
• The net effect of those protein interactions is the acceleration of drug discovery to combat diseases such as HIV and cancer.

The Blueprint Initiative at the Samuel Lunenfeld Research Institute in Toronto's Mount Sinai Hospital is operating and curating the Biomolecular Interaction Network Database (BIND). The Research Institute brings together top scientists in bioinformatics and proteomics to enumerate and study proteins and their interactions, which is leading advances in researching and treating HIV, cancer and other diseases. The Institute is the world's largest protein interaction database research facility.

Sun provides the Blueprint Initiative with all the technology tools required to plan, build, run, and manage large scale computing grids. These grids give researchers running the BIND database the ability to optimize network performance by harnessing online resources more effectively, including CPU cycles. Ultimately, it means users can test, translate and verify genomic and proteomic information more quickly, which will speed breakthroughs in identifying disease-related proteins and drug development.

Challenges

Blueprint North America's mission is to integrate the world's biomolecular assembly knowledge and support a complete understanding of cellular life through high-performance computing simulation.

The cornerstone of Blueprint North America's Biomolecular Data Center is the BIND database, which holds data from experiments carried out by scientists around the world. BIND was conceived by considering and utilizing the world's most comprehensive integrated bioinformatics standards, including those used by the U.S. National Center for Biotechnology Information (NCBI), for storing biomolecular sequence, taxonomy, and structure and literature information. The database is freely available to any researcher, company or organization via the web.

To provide the computing support for the database, it was necessary to overcome several challenges, including:
• How to optimize network performance to link a number of large grids from around the world;
• Building a system with the ability to scale from a beginning level of about 25,000 protein interactions to some 200,000 interactions.

Solution

To effectively run the Bimolecular Interactions Network Database (BIND), Sun developed a solution which called for multiple Sun Fire™ V880 servers, Sun Fire V480 servers, Sun Enterprise™ 280R servers, Sun Ray™ Ultra-Thin Clients, Sun Fire V60 servers, a Sun StorEdge™ 3910 with 24 TB of storage, a Sun StorEdge L700 tape library, a Sun Fire V1280 server, and a storage archive management file system.

The solution included running Sun Java™ Enterprise System (JES) software and the Solaris™ operating environment. Additional data analysis tools includeed SQHound and Blast.

Benefits of Sun

By clustering the Sun servers, Blueprint could offer highly available, horizontally scalable processing power, and lower total cost of ownership. The system now satisfies the entire high-performance computing needs of the Samuel Lunenfeld Research Institute, which is home to Blueprint, and is also used by the University of Toronto Proteomics and Bioinformatics research program.

"Sun's grid solution is enabling this proteomics revolution by providing the level of power, flexibility and scalability the Blueprint Initiative requires to keep pace with the explosion of proteomic data," noted Stephane Boisvert, President of Sun Microsystems of Canada.

Sun's active role in the life sciences community brings additional value to the Blueprint Initiative. Sun's experience in this community includes the knowledge of how to work with local government granting agencies.

Results

Currently, Blueprint North America's BIND database boasts the highest number of protein interactions on one database worldwide. BIND has accumulated more than 25,000 protein interactions and is emerging as the industry's primary source for protein data. In conjunction with Blueprint Asia and Blueprint Europe (which are being established), the BIND database is expected to increase to a total of 200,000 interactions within the next five years.

The net effect of those protein interactions is the acceleration of drug discovery to combat diseases such as HIV and cancer.

"With the computing support provided by Sun Microsystems, combined with our public archive of data from leading scientists and academics from around the globe, we'll significantly reduce the time to deliver valuable results in disease treatment," said Dr. Christopher Hogue, principal investigator of the Blueprint Initiative. "Our information management and analysis capabilities as well as our deliverables exceed those of any proteomics company in the world," he said.

Signalling the overwhelming success to date, Blueprint Initiative North America and Sun recently opened the first global Sun Center of Excellence in Systems Biology at the Samuel Lunenfeld Research Institute.

Sun Solution Components:

- Sun Fire V880 servers
- Sun Fire V480 servers
- Sun Enterprise 280R servers
- Sun Fire V60 servers
- Sun Fire V1280 server
- Sun Ray Ultra-Thin Clients
- Sun StorEdge 3910 (24 TB of storage)
- Sun StorEdge L700 tape library
- Sun Java Enterprise System (JES)
- Solaris Operating Environment