

# Genome Institute of Singapore

## Sun Microsystems selected to power Singapore's national Genomic Research Project

**“Sun has a reputation for being robust and worry free, this has held true in our environment. Sun has also been able to deliver greater value in terms of flexibility and efficiency that greatly enhances the performance of our work, plus the fact that many of the cutting-edge algorithms for this domain are being developed in the academic arena in which Sun is extremely strong. Put together, these factors make a strong case for our selecting Sun solutions over the other vendors. We are very happy to have chosen Sun for our initial IT server setup,” Heidi Dowst, Scientific Computing Group Leader, GIS.**

Sun Microsystems once again takes a major stake in high-performance computing application by providing the computing platforms that power Singapore's national flagship Genomics research infrastructure. In an implementation with the Genome Institute of Singapore (GIS), Sun systems are being used to carry out state-of-the-art research projects in genetic sequencing analysis, genotyping, bioinformatics, proteomics, and expression array technologies. Sun's end-to-end solutions are selected based on their ability to deliver seamless data-flow and fast file-access capabilities, network and resource management, as well as a strong presence in the academic arena where Sun platform is the platform of choice for developing Genomics research tools.

### **About the Genome Institute of Singapore (GIS)**

The Genome Institute of Singapore (GIS) is an initiative of Singapore's Agency for Science, Technology and Research (A\*STAR) to create a national infrastructure for genomics research. The institute brings scientists from different disciplines together, training new talents in the field, and acts as a force for creating collaborative genomics studies and research in the country.

The goal of GIS is to become an international center for biomedical sciences and its relevant industries. Since its founding in 2001, it has attracted over 120 scientists, fellows, and managers from North America, Europe, Asia, and Singapore, and plans to engage up to 250 scientists by 2003. The center has instituted a variety of research programs including biology, computational biology, and population genetics.

The explosive growth of the GIS between 2001 and 2002, and the targeted expansion in the next year has meant a sudden increase in the workload for the center's computing infrastructure. This results in a highly complex computing infrastructure that has to meet the requirement of scaling rapidly to keep up with leaps in computational demands. Furthermore, work carried out at GIS is known for high computational demands. An analysis of DNA sequencing, for instance, could take days, if not weeks to complete. GIS projects such as Proteomics, DNA sequencing, Microarrays, and complex algorithms all put a considerable strain on the center's IT infrastructure.

In addition, GIS also faced the problem of having to access from a remote location public databases that exist in other countries. Searches that were being carried out by GIS researchers required access to massive databases such as Ensembl, Genbank, and SwissProt, which are all located outside of Asia. GIS's access to them is limited through competition with other research groups for a relatively small overseas data pipe. A decision was reached to download the data from those databases locally, which amounted to almost 700GB of data, in addition to the high-throughput data generated by GIS's own projects.

“Knowing that Bioinformatics analysis requires large amount of computing power, as well as additional demand from our own high-throughput laboratories where gigabytes of data is accessed and created on a regular basis, we knew that we needed a really high-performing system to fulfill all our needs. However, as a government funded project, we are very conscious of the value we get from our investments, so a vendor must be able to show us real additional value or we would not be able to purchase the solutions,” said Heidi Dowst, Scientific Computing Group Leader, GIS.



#### **Company**

Genome Institute of Singapore

#### **Industry**

Education/Research

#### **Applications/Solutions:**

- Solaris™ Operating Environment

#### **Products/Solutions**

- Sun Fire™ 4800
- Sun StorEdge™ T3
- Sun Fire™ V880
- Sun Enterprise™ 220R

#### **Key Business Challenges**

- Ability to handle rapid expansion plans that result in leaps in computational demands
- Ability to handle high-throughput as well as high resource demands of advance research techniques and applications
- Capacity to store, access and process data from public data repositories as well as data generated from internal projects
- Raise work productivity from research to production

#### **Key Business Solutions**

- Implement high-performance and technical computing (HPTC) infrastructure
- Derive greater performance through Sun's technology in network and system management
- Leverage on StorEdge technology for reliable back-up and fast file management system
- Implement Sun solutions that are native platform for most state-of-the-art research applications

#### **Key Alliance**

- Frontline Technologies

## Laying the Foundation for High-Performance and Technical Computing (HPTC) Infrastructure

GIS began the selection of a high-performance computing platform in mid-2001. Several vendors bid for the project, and Sun Microsystems took the lead as the top choice early in the process. GIS selection committee was looking for high-performance and technical computing capabilities, but they also wanted solutions that delivered real value in terms of making their work more efficient and productive.

Sun Microsystems was able to deliver those values by its end-to-end solutions that provided flexibility and efficiency in network and resource management. Sun's open Computing technology, for instance, enables easy integration, thereby ensuring that the entire network is optimized for performance throughout all times. Secondly, Sun's HPTC solutions are able to realize seamless data-flow throughout the infrastructure while providing fast file-access capabilities from database to workstations. Finally, Sun was also selected for its presence in higher research institutions. Sun Solaris is the most familiar platform for most of GIS's researchers, and most of the tools they wanted to use were written for the platform. Choosing Sun solutions would enable them to use those research tools in their native computing environment.

"We chose Sun Microsystems because Sun has a long history of being involved in the education and research, especially in the life science industry. Their early involvement in projects like the PDB project and the Alliance for Cellular Signaling demonstrated their commitment to this scientific field. We are therefore very comfortable that SUN Microsystems would be an excellent partner in our building process," Dr Edison Liu, Executive Director, GIS.

Implementation began in June 2001. To handle the rapidly expanding computational needs, GIS quickly invested in two SunFire 4800 servers. Sun Microsystems and partner, Frontline Technologies collaborated to put the systems in place. The systems were initially used for development and testing, but GIS was able to achieve smooth operations through the help of Sun and Frontline Technologies, and moved the systems into production-level operations very quickly.

Additionally, GIS also installed four Sun T3 SAN storage bricks to house the massive data that was expected to arrive when the center downloads the databases of Ensembl, Genbank, and SwissProt. The T3's serve as back-up and main database for storing the downloaded data as well as the large amount of data that will be generated from GIS's own operations. Sun's V880 and E220R servers completed the line-up for building a high-performance computing network.

Sun solutions that were deployed for the GIS project include:

- Sun Fire 4800
- Sun StorEdge T3
- Sun Fire V880
- Sun Enterprise 220R

## Building HPTC Capabilities Quickly

The GIS implementation was completed by November 2001. By choosing Sun solutions, the research institute was able create an enterprise server room with fully developed high-performance computing capabilities in a very short time and began handling advance research activities almost immediately. Today, Sun servers are used for a variety of projects, including an automated sequence analysis pipeline, the development and implementation of an Ortholog Synteny algorithm, and the development and implementation of an Estrogen Promotor Mining Suite.

"Sun has a reputation for being robust and worry free, this has held true in our environment. Sun has also been able to deliver greater value in terms of flexibility and efficiency that greatly enhance the performance of our work, plus the fact that many of the cutting-edge algorithms for this domain are being developed in the academic arena in which Sun is extremely strong. Put together, these factors make a strong case for our selecting Sun solutions over the other vendors. We are very happy to have chosen Sun for our initial IT server setup," concluded Dowst.

Today, Sun servers are used for a variety of projects, including an automated sequence analysis pipeline, the development and implementation of an Ortholog Synteny algorithm, and the development and implementation of an Estrogen Promotor Mining Suite.

**HEADQUARTERS** SUN MICROSYSTEMS, INC. 4150 NETWORK CIRCLE, SANTA CLARA, CA 95054 USA,  
PHONE: +1-800-555-9SUN OR +1-650-960-1300 INTERNET: [www.sun.com](http://www.sun.com)

### SALES OFFICES

AFRICA (NORTH, WEST AND CENTRAL): +9714-3366333 • ARGENTINA: +5411-4317-5600 • AUSTRALIA: +61-2-9844-5000 • AUSTRIA: +43-1-60563-0 • BELGIUM: +32-2-704-8000 • BRAZIL: +55-11-5187-2100 • CANADA: +905-477-6745 • CHILE: +56-2-372-4500 • COLOMBIA: +571-629-2323 • COMMONWEALTH OF INDEPENDENT STATES: +7-502-935-8411 • CZECH REPUBLIC: +420-2-3300-9311 • DENMARK: 45 4556 5000 • EGYPT: +202-570-9442 • ESTONIA: +372-6-308-900 • FINLAND: +358-9-525-561 • FRANCE: +33-01-30-67-50-00 • GERMANY: +49-89-46008-0 • GREECE: +30-1-618-811 • HUNGARY: +36-1-202-4415 • ICELAND: +354-563-3010 • INDIA: +91-80-5599595 • IRELAND: +353-1-8055-666 • ISRAEL: +972-9-9513465 • ITALY: +39-039-60551 • JAPAN: +81-3-5717-5000 • KAZAKHSTAN: +7-3272-466774 • KOREA: +822-3469-0114 • LATVIA: +371-750-3700 • LITHUANIA: +370-729-8468 • LUXEMBOURG: +352-49 11 33 1 • MALAYSIA: +603-264-9988 • MEXICO: +52-5-258-6100 • THE NETHERLANDS: +31-33-450-1234 • NEW ZEALAND: +64-4-499-2344 • NORWAY: +47-2202-3900 • PEOPLE'S REPUBLIC OF CHINA: BEIJING: +86-10-6803-5588 CHENGDU: +86-28-619-9333 GUANGZHOU: +86-20-8755-5900 SHANGHAI: +86-21-6466-1228 HONG KONG: +852-2202-6688 • POLAND: +48-22-8747800 • PORTUGAL: +351-21-4134000 • RUSSIA: +7-502-935-8411 • SINGAPORE: +65-438-1888 • SLOVAK REPUBLIC: +421-7-4342 94 85 • SOUTH AFRICA: +2711-805-4305 • SPAIN: +34-91-596-9900 • SWEDEN: +46-8-631-10-00 • SWITZERLAND: GERMANY: 41-1908-90-00 FRENCH: 41-22-999-0444 • TAIWAN: +886-2-2514-0567 • THAILAND: +662-344-6888 • TURKEY: +90-212-335-22-00 • UNITED ARAB EMIRATES: +9714-3366333 • UNITED KINGDOM: +44-1-276-20444 • UNITED STATES: +1-800-555-9SUN OR +1-650-960-1300 VENEZUELA: +58-2-905-3800 • OR ONLINE AT [SUN.COM/STORE](http://SUN.COM/STORE)



We make the net work



© 2002 Sun Microsystems, Inc. All rights reserved. Sun, Sun Microsystems, the Sun logo, Sun Fire, Sun StorEdge, SunSpectrum Gold, SunSpectrum Platinum, Java and Solaris are trademarks or registered trademarks of Sun Microsystems, Inc., in the United States and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the United States and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.