



## PRESS RELEASE

20 Dec 2005

### **National Grid Proves to be a Powerful Tool for Biomedical Research**

***Grid Computing enables Singapore scientists to tackle massive genomics data, and opens new possibilities for researchers in pushing the envelope of biomedical sciences.***

Scientists at the Genome Institute of Singapore (GIS) led by Drs Ruan Yijun and Patrick Tan Boon Ooi recently completed a pilot grid research programme to study metagenomic sequence data of microorganisms found in Singapore. The research team has been using the latest DNA sequencing technology to study uncultured microorganisms residing in the human guts and natural habitats, and overcame the major challenge of analyzing and identifying the thousands of bacterial species through the computing resources made available by the National Grid.

“The sequencing process resulted in huge amounts of data. Initially we tried to match a pilot data set of 20,000 DNA sequences against a huge database of some 3 million known sequences, using a single run on isolated computing resources in our institute. It took us an entire month,” recounted Dr Patrick Tan. “We needed to shorten this process significantly and it soon became clear that a totally new type computing infrastructure was required to handle data on this scale.”

With the support of the National Grid Office, the analysis was repeated on the Tera-scale Campus Grid at NUS (TCG@NUS), tapping into pre-existing compute resources across the National University of Singapore. The entire query was remarkably shortened to less than two days.



“Grid computing was 15 times faster compared with the earlier processing time for a single run on an isolated system. The grid system also provided us with higher capacity and high fault tolerance.” said Dr Patrick Tan.

“By tapping into the existing network of compute resources, Grid computing is a cost-effective way to enable large-scale research computation that was previously not possible or too time consuming to perform.” explained Mrs . Tan Chee Kiow, Deputy Director of NUS Computer Centre and a member of the National Grid Governance Council. For more details on Grid Computing, please refer to attached Annex A.

With the success of this pilot grid-enabled trial, the GIS team plans to scale up its study on micro-organisms in the air involving analysis of data sets about ten times larger. The team also has plans to do similar analysis of other environments.

“Grid computing has the potential to develop into the next generation tool for solving more complex biomedical challenges,” said Dr Patrick Tan, who also chairs the Life Sciences Virtual Grid Community set up by the National Grid Office to foster grid-based applications in the life sciences. “We have forged links amongst A\*STAR research institutes, NUS and other research organizations. I encourage more researchers in the community to tap into and contribute towards this powerful computing grid.”

The National Grid Office will also be launching another project to use grid resources for molecular modelling in drug screening for anti-cancer targets. They will be speaking to schools to encourage their contribution of PC resources in their school for this project. “Virtual drug screening is a process which allows you to save lots of time and cost when designing compounds against selected targets such as cancer-related or diabetes related molecules. One has to “dock” thousands of putative drug compounds to your target before a reasonable solution can be found. This is a computationally taxing process and the use of the National Grid can help to obtain solutions in a reasonable frame of time. We can then carry out experiments to verify the findings for a focused



group of compounds found through the virtual screening,” said Dr. Prasanna Ratnakar Kolatkar from GIS.

“The use of grid resources provides more opportunities for collaborative ties among the various research organizations. Researchers can now come together to conceptualize and tackle large scale scientific questions that cannot be done using isolated computing resources of individual organizations,” said Mr. Richard Lim, Chairman of the National Grid Steering Committee.

---

For enquiries, please contact the following:

National Grid Office

Choo Thong Tiong  
Head (Promotion)  
Tel: 68745583  
Email: [thongtong@ngp.org.sg](mailto:thongtong@ngp.org.sg)

National University of Singapore

Tan Chee Chiang  
Manager, SVU/Academic Computing, NUS Computer Centre  
Tel: 65166211  
Email: [ccetancc@nus.edu.sg](mailto:ccetancc@nus.edu.sg)

Genome Institute of Singapore

Daniel Lum  
Senior Officer  
Office of Corporate Communications  
Tel: 64788013  
Email: [lumd@gis.a-star.edu.sg](mailto:lumd@gis.a-star.edu.sg)



### About the National Grid Office

The NGO was established by the Agency for Science, Technology and Research in January 2003 to promote Grid Computing and to develop a cyber-infrastructure that steers Singapore towards a Grid-enabled economy where computing resources, services and intellectual property can be provisioned securely on a high-speed network.

The roles of the National Grid are:

1. To develop a cyber-infrastructure for science and engineering research and education; and
2. To promote the use of Grid Computing for research, academic, commerce and industry.

The NGO strives to achieve the vision for the National Grid by the following means:

- Formulate the framework & policies;
- Plan & develop a secure platform;
- Adopt common open standards;
- Encourage the adoption of Grid Computing;
- Demonstrate the commercial viability of compute-resource-on-tap; and
- Lay the foundation for a vibrant Grid Computing economy.

For more information, please visit [www.ngp.org.sg](http://www.ngp.org.sg)

### About the Genome Institute of Singapore

The Genome Institute of Singapore (GIS) is a member of the Agency for Science, Technology and Research (A\*STAR). Established in 2001, the research institute's mission is to be a world-class genomics institute and a centre for genomic discovery. GIS pursues the integration of technology, genetics, and biology towards the goal of individualized medicine. The genomics infrastructure at GIS is utilized to train new scientific talent, to act as a bridge between academic and industrial research, and explore scientific questions of high impact. For more information, please visit [www.gis.a-star.edu.sg](http://www.gis.a-star.edu.sg)

### About Tera-scale Campus Grid @ NUS

TCG@NUS is a Grid computing development project to harness idle compute cycles of large number of PCs on campus, with the aim to create a new computing capability that enables large scale computation. Currently more than 1000 PCs are grid-enabled and there is a plan to expand the PC Grid to 3000 nodes in the next two years. The new computing capability allows researchers to do more science and to push the boundary of research computation further. For more information, please visit [http://www.nus.edu.sg/comcen/svu/about\\_PCGrid.html](http://www.nus.edu.sg/comcen/svu/about_PCGrid.html)



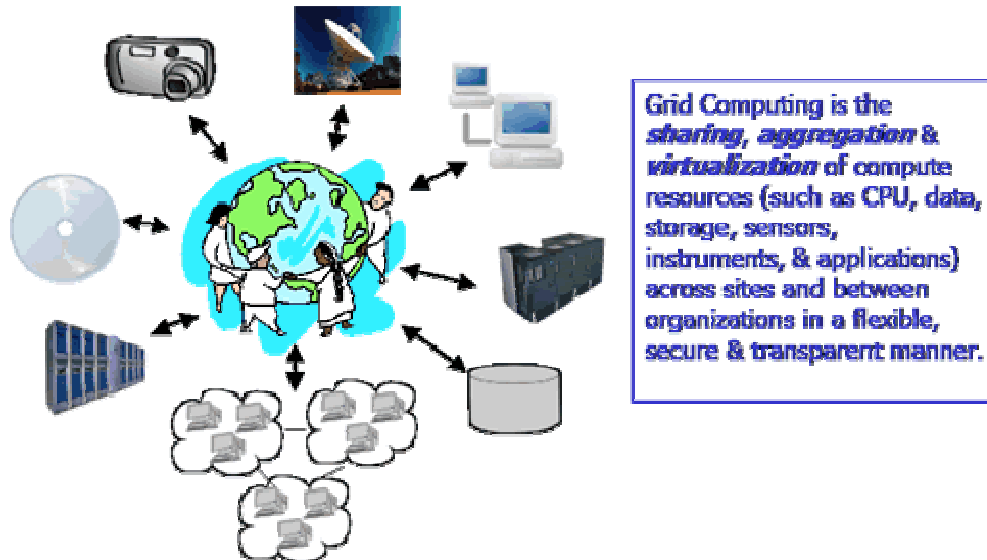
Genome Institute  
of Singapore

National Grid  
**NG**  
SINGAPORE



**NUS**  
National University  
of Singapore

## Annex A



### How Grid Computing works

Grid computing is a massive smoothing technique for asset allocation. A Grid makes compute resources (such as CPU, data, storage, instruments, and applications) securely and transparently accessible to authorized users and applications, across sites and between organizations.

### Characteristics of Grid

A Grid tends to have numerous resources that are heterogeneous in nature. The resources are typically geographically apart and owned by various distrusting organizations with different resource management policies. The Grid supports collaboration and collectively delivers a service that is non-trivial.

### Advantages of Grid Computing

The value from deploying a Grid is dependent upon the specific end user application.

The benefits that are typically achieved from aggregating compute resources include:

- Reduced hardware and software costs.
- Reduced ICT complexity and systems management costs.
- Increased end user productivity and effectiveness.
- Accelerated time to market through increased collaboration, faster computation and better quality analysis.