National Supercomputing Centre (NSCC) Singapore e-newsletter

NEWSBYTES

May 2021

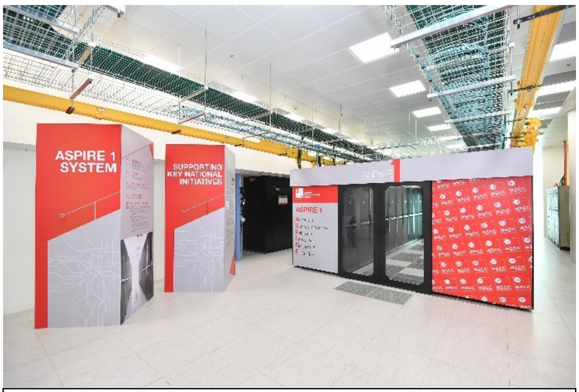
In this Issue…

CORPORATE NEWS

Singapore scientists will soon have a new national supercomputer for research!

NSCC has awarded the tender for the development of the next generation national supercomputer system which will upgrade Singapore's HPC capabilities for the research community. The development also includes the national storage and research network infrastructure.

The next generation national supercomputer for Singapore will be a green, warm water-cooled system – one of the first known deployments of such a system in a tropical environment. When operational the supercomputer is expected to provide an aggregate of up to ***10 PFLOPS of raw compute power*** and is ***eight times more powerful*** than the current ASPIRE1 supercomputer. ASPIRE1, which was commissioned in 2016, has been running at near full capacity in support of local advanced research that requires high-end computing resources. The new system is the first in a series of supercomputers that will be deployed in phases from now till 2025 to expand and upgrade Singapore's high-performance computing (HPC) capabilities for the research community here.



The *ASPIRE1* – Singapore's first petascale supercomputer. The new supercomputer will be 8x more powerful than *ASPIRE1* and provide an aggregate of up to 10PF of raw compute power.

The upgraded national supercomputing resources will support and strengthen local research at universities, research institutes, government agencies and companies across a wide array of scientific fields.

Supercomputers can be used to augment and enable tools such as AI, deep machine learning and quantum computing as well as amplify research in areas like weather and climate monitoring, genomics, engineering, offshore and marine, urban planning and biomedical science, among many others. The national research infrastructure upgrade is part of a SGD200M National Research Foundation grant that was set aside by the Singapore government and announced at the SupercomputingAsia (SCA) conference in March 2019.

"Supercomputers have proven their value to the Singapore research community in many ways. Examples include simulating the dispersion of COVID-19 droplets to mitigate spread, helping urban planners develop our most advanced townships as well as advancing weather and climate monitoring research," said Mr Peter Ho, Chairman of NSCC's Steering Committee. "The upgrading of the supercomputer resources is another milestone in Singapore's research roadmap and will provide the necessary HPC resources that will allow Singapore scientists to unlock even greater research potential."

Associate Professor Tan Tin Wee, Chief Executive of NSCC added, "Based on the HPC demands garnered from Singapore researchers, the new system will provide the necessary resources to meet their growing supercomputing needs, and to enable more of such significant scientific breakthroughs at the national and global level."

The new supercomputer will comprise nearly 900 CPU and GPU computing nodes of different configurations, and more than 100,000 computing cores. The system is designed to cater to a range of different users across various scientific fields - a reflection of the national research infrastructure mission of NSCC. The yet-to-be named supercomputer is expected to be commissioned and operational by early 2022.

Want to know more about the new supercomputer?

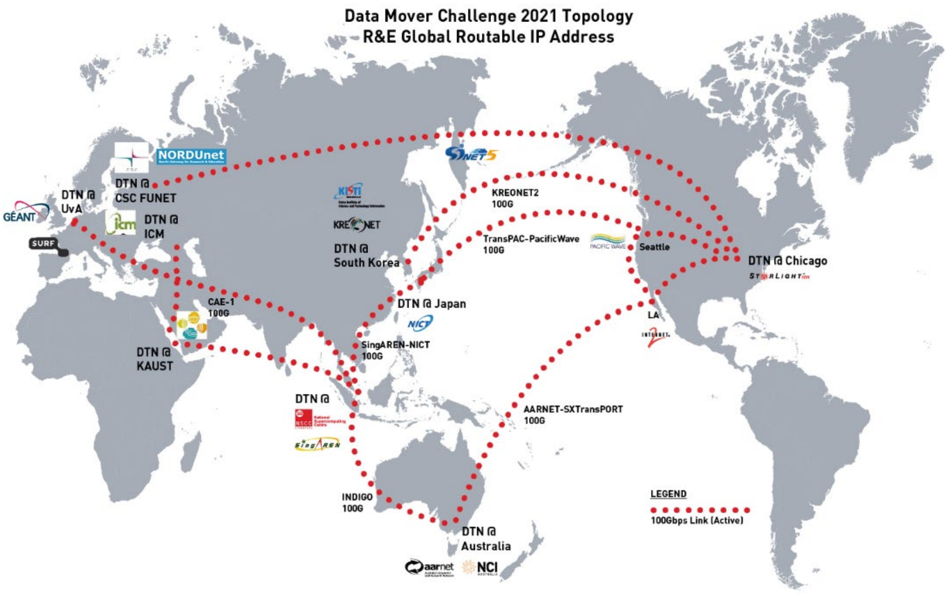
If you are interested to know how you can benefit from the upgrades and increased HPC resources, drop us an email at e-news@nscc.sg to find out more!

For more examples about how supercomputing is being applied in Singapore research, please visit www.nscc.sg/case-studies/.

Back to main content list

The Data Mover Challenge (DMC) 2021 is open!

From now till <u>31 May 2021</u>, all organisations, companies, research institutions, academia, researchers, post-graduate students and undergraduate students can submit their applications to take part in the global competition.



The international Data Mover Challenge (DMC) is an annual competition that aims to bring together experts from industry and academia in a bid to test their software and solutions for transferring huge amounts of research data. DMC seeks to challenge international teams to come up with the most advanced and innovative solutions for data transfer across servers located in various countries that are connected by 100Gbps international research and education networks.

The challenge focuses on optimising point-to-point data transfers between sites – a crucial step forward in advancing research collaboration and sharing. Participants from all over the world will compete by deploying the best software tools on Data Transfer Nodes (DTNs) that are set up within existing international networks across the globe.

After the successful teams have been selected and notified, the DMC21 will run from **1 August to 31 October 2021**. The winning team will be announced at the SupercomputingAsia 2022 (SCA22) Conference in March 2022 in Singapore. The winning team's leader will be invited for the DMC21 award presentation and solution showcase at SCA22 from 1-3 March 2022 in Singapore*.

Register here or head over to www.nscc.sg/data-mover-challenge-2021/ for more information on DMC 2021.

**Award presentation and solution showcase may be done virtually depending on prevailing pandemic circumstances.*

Back to main content list

The **EU-ASEAN High-Performance Computing (HPC) Virtual School 2021: System Design and HPC Applications** seeks 60 selected participants from ASEAN Member States to learn about the fundamentals of HPC design and applications to critical domains, such as COVID-19, and urgent computing and climate science from international luminaries from Europe, ASEAN and Japan through formal and practical courses.

The virtual school is a hands-on virtual curriculum taught by foremost international experts in HPC technology and its applications to computational sciences. It is hosted by National Science and Technology Development Agency (NSTDA) Supercomputer Center (ThaiSC), Thailand.

From **5 to 9 July 2021**, selected participants will receive tutorials on top supercomputers such as LUMI and Meluxina, new EuroHPC world-class supercomputers being deployed in the EU, and Riken Fugaku Japan, the number one supercomputer in the world.

Head over to www.hpcschool.net/ for more information.

Back to main content list

Studying the potential effect of storm surges and tsunamis on coastal regions in the South China Sea

A research team from NUS is harnessing high performance computing to quantify the uncertainty of earthquake characteristics and tropical cyclone storm surges to estimate its effects on coastal inundation in the region.

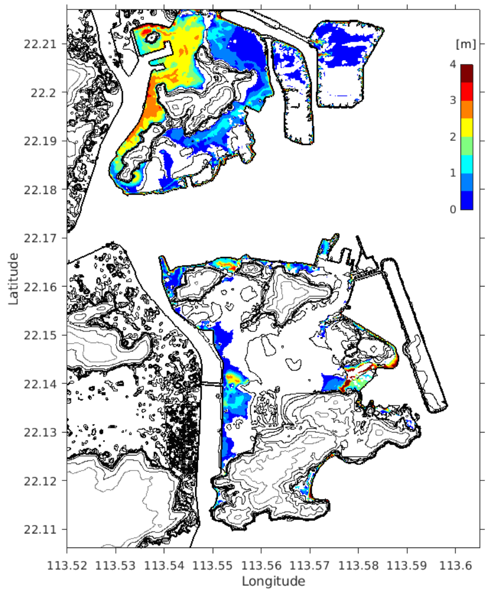
Intensified tropical cyclone activities in the western North Pacific have become increasing threats to coastal cities in the context of global climate change. Storm surges superimposed with astronomical tides often cause severe flooding in many coastal cities in South China Sea region.

The potential for coastal inundation can also be linked to potential *tsunami* hazards associated with earthquakes originating from the Manila trench. Research and assessment in these areas have become increasingly important for the region. These two types of coastal disasters have been studied independently in the past since they are caused by different source mechanisms in nature.

A team of reseachers at the Department of Civil and Environmental Engineering at the National University of Singapore are using NSCC's supercomputing resources to investigate the potential compound impact of storm surges and *tsunamis* on Macau and Hong Kong through the numerical modelling of the hydrodynamic processes involved. The team is looking to develop a comprehensive regional coastal hazard database by synthesizing historical tsunami and earthquake events associated with the Manila trench as well as a methodology for quantifying the uncertainty of earthquake characteristics in the Manila trench to estimate its effects on coastal inundation in the region.

"Numerical simulation of the hydrodynamics on the scale of the South China Sea region covering the entire typhoon process over several days is extremely computationally intensive, especially when we have to look at a variety of scenarios. Running the simulations in parallel on NSCC's HPC resources significantly saves our time on the acquisition of numerical data, so that we can speed up our research and focus on the analysis."

Wang Jinghua
Research Fellow
Department of Civil and
Environmental Engineering,
National University of Singapore



The figure above shows the maximum inundation water depth in Macau subject to the concurrent storm surges and *tsunami* hazards.
Credit: Department of Civil and Environmental Engineering, NUS

To find out more about the NSCC's HPC resources and how you can tap on them, please contact e-news@nscc.sg.

Back to main content list

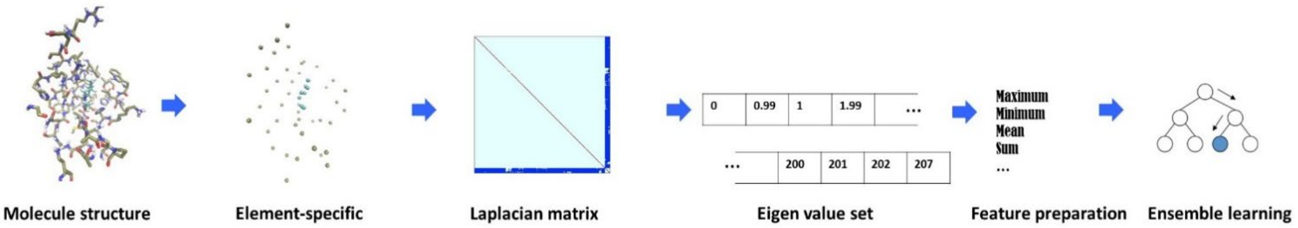
Using Machine Learning to identify and improve protein-ligand binding affinity predictions

Researchers from NTU are utilising NSCC's supercomputing resources to improve the performance of learning models in molecular data analysis.

AI is expected to play an increasingly important role in the years to come, especially in biological studies. As a major part of AI, neural networks cannot achieve such success without the support of enormous data.

The huge advancements in biological sciences and technologies has led to the accumulation of unprecedented amounts of biomolecular data. For example, in a protein data bank, there are about 150,000 three-dimensional biomolecular structures (Berman, Westbrook et al. 2000). There is an abundance of available biological structures, data analysis methods and models, including data mining, manifold learning and graph or network models. Leveraging the data, topological data analysis (TDA), for example, can potentially provide great promise in the big data era and have become increasingly popular in bioinformatics and computational biology in the past two decades.

Data-driven learning models are among the most important and rapidly evolving areas in chemoinformatics and bioinformatics. Featurization, or feature engineering, is key to the performance of machine learning models in material, chemical, and biological systems. As such, a group of researchers at the School of Physical and Mathematical Sciences at Nanyang Technological University Singapore are using high performance computing to develop a new molecular representation framework, known as persistent spectral (PerSpect), and PerSpect based machine learning (PerSpect ML) for protein-ligand binding affinity prediction. The proposed PerSpect theory provides a powerful feature engineering framework. PerSpect ML models demonstrate great potential to significantly improve the performance of learning models in molecular data analysis.



The workflow of PerSpect-ML.
Credit: School of Physical and Mathematical Sciences, NTU

"NSCC provides sufficient computing resources, including CPU, GPU, and storing space. In our research, the most demanding resource is CPU, and for the whole project, we consume around 100,000 CPU hours. We also use GPU in testing our CNN models, and NSCC's resources provides the Nvidia Tesla K40 as well as more advanced DGX1 specific for deep learning system."

Meng Zhenyu
Research Fellow
School of Physical and Mathematical Sciences
Nanyang Technological University Singapore

High performance computing plays a pivotal role in the team's daily work. The project needs to process large databases which contains thousands of entries, the databases needs to be divided into several pieces and parallel computing has to be employed to treat each part. Additionally, some algorithms are time/memory-consuming and computing resources with multiple cores and large memory are needed to run them.

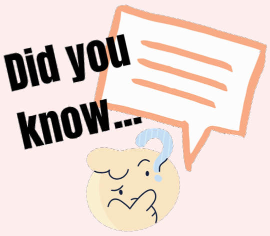
To find out more about the NSCC's HPC resources and how you can tap on them, please contact e-news@nscc.sg.

Back to main content list

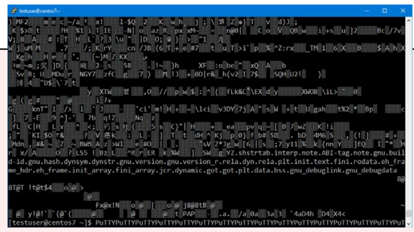
TECHNICAL NEWS

Terminal gibberish after displaying binary file

How do you avoid it?



Opening a binary file using cat, tail and head commands will cause gibberish in shell terminal. This can be avoided by maintaining best practices while working in shell terminal.



In the Linux world there are two general kinds of files - ASCII and binary ones. ASCII files are text files such as PBS Pro job files (suffix .pbs), Python (suffix .py), input files for various applications (suffix .inp) and similar. ASCII files can be readily edited with a text editor of choice and viewed using commands like less, more, cat or tail.

Binary files by contrast cannot be edited with a text editor or viewed using the commands mentioned above. Using cat or tail will fill your screen with pages of gibberish and likely cause your terminal to become unresponsive and crash. This can be avoided by maintaining best practices as stated below while working in shell terminal.

a) Always check the file type before viewing its content

Use file command to check the type of file and make sure it is not a binary file before viewing the file content.

$ file <filename>
$ file -i <filename>

For example:
$ file binaryfile
binaryfile: **ELF 64-bit LSB executable**, x86-64, version 1 (SYSV), dynamically linked (uses shared libs), for GNU/Linux 2.6.32, BuildID[sha1]=c8ada1f7095f6b2bb7ddc848e088c2d615c3743e, stripped
$ file -i binaryfile
binaryfile: application/x-executable; charset=**binary**

$ file textfile
textfile: **ASCII text**
$ file -i textfile
textfile: text/plain; charset=**us-ascii**

b) Use less and more commands instead of cat, tail or head command

By default, less and more commands will check the file type and filter a binary file.

$ less binaryfile
"binaryfile" may be a binary file. See it anyway?

$ more binaryfile
******** binaryfile: Not a text file ********

For more information and FAQs on ASPIRE 1, please visit:

https://help.nscc.sg

Back to main content list

THE LAST BYTE...

<SHARED CONTENT>

Shared articles and news from the HPC world.

Australian government to spend $64m on Pawsey Square Kilometre Array supercomputer

Funding will come over a decade, joining a number of supercomputing efforts for the SKA.

The Australian government will spend $64.4 million on a new supercomputer in Perth to process and analyse data from the Square Kilometre Array (SKA) radio telescope. The world's largest public science data project, the SKA will be huge in scope, cost, and output. Receiving stations will extend at least 3,000 kilometers (1,900 mi) from a concentrated core, and span across both Australia and South Africa. Read more at Data Center Dynamics here.

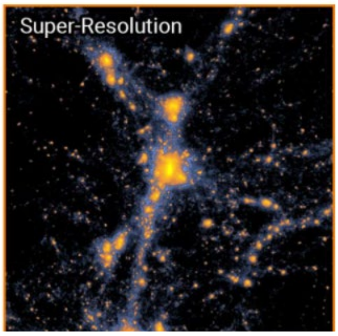
Credit: UWA

Back to main content list

Machine Learning accelerates cosmological simulations

Using neural networks, researchers can now simulate universes in a fraction of the time, advancing the future of physics research.

A universe evolves over billions upon billions of years, but researchers have developed a way to create a complex simulated universe in less than a day. The technique, published in this week's Proceedings of the National Academy of Sciences, brings together machine learning, high-performance computing and astrophysics and will help to usher in a new era of high-resolution cosmology simulations. Read more at Science Mag here.



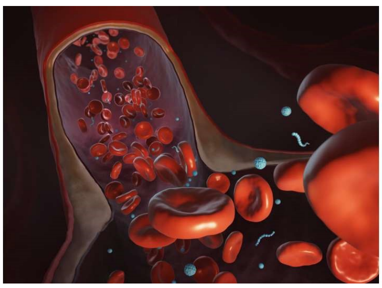
Credit: Y. Li et al./Proceedings of the National Academy of Sciences 2021

Back to main content list

Targeting tumors with nanoworms

Drugs and vaccines circulate through the vascular system reacting according to their chemical and structural nature.

The effectiveness of a medicine — and how much is needed and the side effects it causes — are a function of how well it can reach its target. Ying Li, an assistant professor of mechanical engineering at the University of Connecticut, harnesses the power of supercomputers to simulate the dynamics of nanodrugs in the blood stream, design new forms of nanoparticles, and find ways to control them. Read more at Phys Org here.



Credit: Phys Org

Back to main content list

Powering Innovation
Supercomputing in Asia
National Supercomputing Centre (NSCC) Singapore
1 Fusionopolis Way, Connexis South, #17-01 Singapore 138632