

NEWSBYTES

April 2024



In this issue...

Corporate News

- 1 Deepening engagement with the local HPC community at NSCC Singapore HPC User Group Event 2024
- 2 NSCC Singapore hosted a workshop on Artificial Intelligence (AI), Machine Learning (ML) and Large Language Models (LLM) to upskill the local HPC community
- 3 3rd Call for Projects: Submit your request for NSCC Singapore's Supercomputing Resources
- 4 Privacy-Preserving Data Fusion for Traffic State Estimation
- 5 An Eye on AI: 5 Major Breakthroughs Fuelled by HPC

Shared News

- 1 NVIDIA to help Japan build hybrid quantum-supercomputer
- 2 Intel Introduces Gaudi 3, Elevating HPC and AI with Increased Speed and Scalability
- 3 AI-powered 'digital twin' of Earth could make weather predictions at super speeds



CORPORATE NEWS

Deepening engagement with the local HPC community at NSCC Singapore HPC User Group Event 2024

The Singapore HPC User Group Event 2024, held on 5 April 2024, provided a dynamic platform for the High Performance Computing (HPC) community to converge, exchange insights, and explore the latest advancements shaping the industry.

Jointly organized by the National Technological University (NTU) and the National Supercomputing Centre (NSCC) Singapore, this year's event took place at NTU, offering 64 attendees a comprehensive overview of cutting-edge HPC technologies and their diverse applications. Throughout the event, participants had the opportunity to delve into a spectrum of topics elucidated by both esteemed academic and industry experts.



Interested to have your research published in NSCC Singapore's NewsBytes?

We are looking for *guest writers / contributors* to be part of our e-newsletters, which are sent out to a subscriber base of more than **7,500** monthly.

If you are interested in contributing content to our NewsBytes, drop us an email at e-news@nscg.sg and we'll be in touch with you!

During the session, Assistant Professor Chan Wai Lee from NTU shared more about how his team use HPC for numerical AeroSpace research while Dr. Kanchan Ajit Ulman and Dr. Lai Mingrui from NUS provided insights about ab initio many-body calculations on excitonic physics and its applications.



Full-house attendance at Singapore HPC User Group Event 2024.

Moreover, industry giants such as Intel, Altair, NVIDIA, and AMD shared their perspectives and advancements in the HPC domain. Intel discussed the omnipresence of AI in its presentation "AI Everywhere with Intel," while Altair elaborated on "harmonizing power and flexibility: orchestrating hybrid HPC workloads". NVIDIA emphasized energy and power efficiency for applications on its platform, while AMD delved into the realm of Quantum simulation on Instinct GPUs.

The event concluded with a vibrant networking session, fostering connections among HPC professionals spanning various sectors, including academia, industry, research organizations, and national platforms.

If you are interested in attending similar future events at NSCC Singapore, subscribe to our mailing list [here](#).

[Back to Main List](#)

NSCC Singapore hosted a workshop on Artificial Intelligence (AI), Machine Learning (ML) and Large Language Models (LLM) to upskill the local HPC community

In a dynamic collaboration with Intel Technologies aimed at fostering innovation within the HPC community, the full-day workshop was conducted at Fusionopolis on 18 April 2024.

27 attendees from esteemed institutions including A*STAR, IMDA, NTU, NUS, Govtech and various research centres, benefited from the enlightening workshop conducted by Anish Kumar, the AI software engineering manager from Intel. With a wealth of expertise in the field of AI software, he guided the participants through an immersive experience, shedding light on Intel's latest advancements in AI technology.



Attendees from various esteemed institutions learning about the latest advancements in AI technology.

The workshop delved into a spectrum of topics designed to equip attendees with the knowledge and skills necessary to navigate the evolving landscape of AI, ML and LLM technologies. Attendees were introduced to Intel Gaudi AI Accelerator that was designed for GenAI and Large Language Models. Through hands-on practice on fine-tuning the system on Intel Gaudi2, Llama2, Intel Xeon and LangChain, the attendees gained many insights into the diverse capabilities of these frameworks, enabling them to tackle complex AI challenges with confidence.

By providing a platform for industry experts to share their knowledge and best practices in using the latest technologies, NSCC Singapore plays a vital role in keeping professionals up to date on the rapidly evolving landscape of AI-driven innovation.

We will be constantly updating our website with the upcoming workshops, so do visit our [website](#) regularly to be in the loop of the latest developments.

[Back to Main List](#)

3rd Call for Projects: Submit your request for NSCC Singapore's Supercomputing Resources

NSCC Singapore's Call for Projects is open, allowing researchers to gain access to Singapore's latest supercomputing resources.

Supercomputers have been proven to be an invaluable resource in the development of cutting-edge technologies in critical science areas such as AI, climate science, biomedicine and more.

The NSCC Singapore is pleased to announce the launch of the 3rd Call for Projects on ASPIRE 2A which would allow researchers to gain access to supercomputing resources to aid their research.

Details	For Research	For Education
Eligibility	NSCC Stakeholder Organisations – A*STAR, NTU, NUS, SUTD	Institutes of Higher Learnings (IHLs) – SIT, SMU, SUSS, NP, NYP, RP, SP, TP
Types of Projects Accepted	<ol style="list-style-type: none"> 1. National initiative 2. Government-funded 3. Industry collaboration 4. General project 	Education-related project
How to Apply	NSCC Projects Portal	

NSCC Singapore will also be holding an optional clinic session for users who require guidance on determining the types of resources required and calculating the actual resources to request for on **2 May 2024, from 2pm to 4pm**. If you are interested in attending the session, please register [here](#).

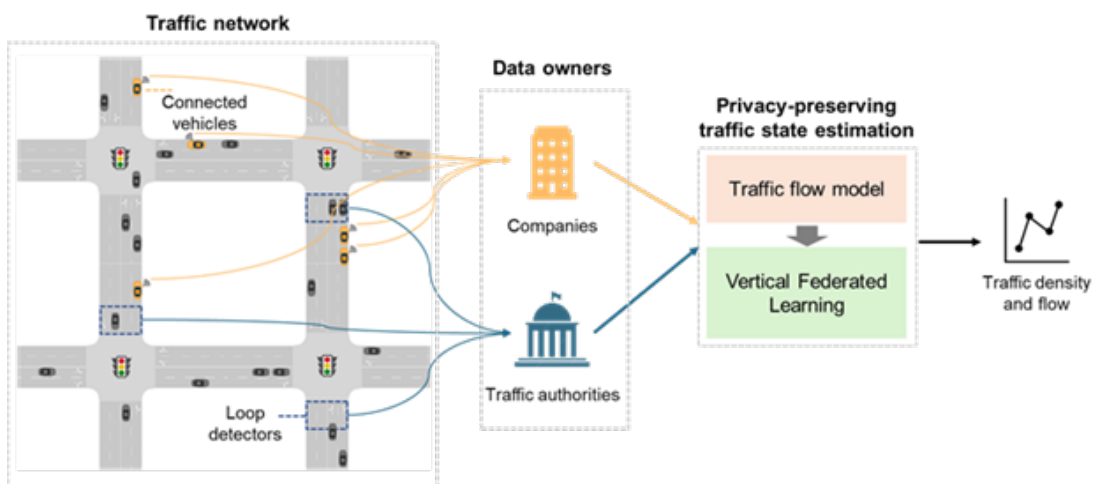
The application period for this Call for Projects is from **17 April 2024, 10am to 17 May 2024, 11.59pm** and the resource allocation period is tentatively from 1 September 2024 to 30 June 2025.

For more information, please refer to the document: [Call for Projects Information and Instructions](#).

[Back to Main List](#)

Privacy-Preserving Data Fusion for Traffic State Estimation

Researchers from the National University of Singapore (NUS) are proposing a privacy-preserved data fusion algorithm for traffic state estimation (FedTSE) to facilitate collaboration and data sharing between multiple data owners, such as traffic authorities and mobility companies.



Transportation systems are undergoing rapid transformation, characterized by the expansion of ridesharing services and the development of connected automated vehicles. These paradigms have enormous potential to facilitate human-centred transportation operations, not only by improving traffic efficiency but also by ensuring responsible data usage (e.g., privacy protection, fairness, etc.).

In the current phase, researchers from the National University of Singapore (NUS) investigated traffic state estimation by merging trajectory data provided by these paradigms with traditional roadside detector data owned by traffic authorities. While such data fusion has been widely demonstrated by researchers and practitioners to significantly enhance traffic operations, existing algorithms rarely consider the fact that sharing such trajectory data can raise privacy concerns for both individual travellers (e.g., their origins and destinations) and mobility companies (e.g., their operation algorithms).

To address the privacy concerns, the researchers proposed a privacy-preserving data fusion algorithm for traffic state estimation, leveraging a combination of federated learning (FL) and traffic flow theory.

The Research

The aim is to develop a privacy-preserving data fusion algorithm for traffic state estimation that accurately estimates key traffic states such as flow, density, and speed for better traffic operations while safeguarding the privacy of mobility companies that provide the trajectory data.

To achieve this, the researchers first utilize the promising framework of FL, which enables multiple parties to collaboratively train a model without exchanging private data. The researchers formulate the traffic state estimation problem as a vertical FL problem and build their algorithm on a recently developed framework, reducing communication overhead through local gradient updates and easily integrating with graph neural networks.

Next, the researchers proposed a physics-informed FL approach that integrates traffic models with FL to improve data efficiency. This ensures the applicability of the proposed FedTSE in common traffic state estimation scenarios with limited ground-truth availability. Physics-informed deep learning integrates physical models into learning-based approaches to improve the data efficiency in the training process and/or to preserve the desired physical properties of the trained models.

The Technology

Vertical federated learning: It enables data fusion between two entities without the need to disclose their raw data.

Traffic flow modelling: It ensures that the estimated traffic states satisfy key constraints defined by traffic flow dynamics.

HPC Resources: This project was allocated 3,000 GPU hours from NSCC Singapore to train their models.

The Impact

Case studies demonstrated that the models developed by the researchers can preserve the privacy of data owners with minimum impact on estimation performance, significantly outperforming baselines where mobility companies only share partial or no data due to privacy concerns. Moreover, by protecting privacy, mobility companies can be more incentivized to participate actively in data fusion and use higher-resolution data, thereby enhancing estimation accuracy.

On the practical side, the proposed strategy can help eliminate the data silos in the transportation industry by promoting collaboration between stakeholders, allowing them to share data safely and trustfully.

“This research is among the few pioneer works that address the cross-company privacy concerns among various transportation entities interested in collaborating and sharing heterogeneous datasets. With the help of NSCC Singapore’s HPC resources, such privacy-preserving FL framework can be extended to data-fusion-based transportation applications in the area of traffic flow prediction and traffic control.”

Assistant Professor Yang Kaidi
Principal Investigator
Department of Civil and
Environmental Engineering



To find out more about how NSCC Singapore’s HPC resources can help you, please contact e-news@nscg.sg.

[Back to Main List](#)

An Eye on AI: 5 Major Breakthroughs Fuelled by HPC

Backed by immense computing power, breakthroughs in AI are transforming multiple facets of society, from the way we deliver patient care to how we harness renewable energy resources like solar power.



Today, enormous amounts of computing power are used to build generative AI (GenAI) models, which are trained on terabytes of data and have parameters in the billions—with some models reaching the trillion-parameter mark. At the same time, HPC is also becoming more democratized as cloud computing creates easier access to innovative AI tools and gives enterprises the ability to create their own AI-driven solutions.

Coming full circle, supercomputers are levelling up their own capabilities, powering AI-driven design tools that accelerate the development of increasingly sophisticated semiconductor chips.

Since OpenAI's ChatGPT tool exploded onto the scene in late 2022, a flurry of large language models (LLMs) has been released as tech giants duke it out to create GenAI capable of accomplishing more complex tasks with greater accuracy.

Across the Pacific, China's GenAI space is booming. Many Chinese tech firms have built their own proprietary LLMs to give their existing products and services a GenAI-powered refresh, as well as to provide AI-based business solutions on the cloud. As of October 2023, the country's tech sector has produced at least 130 LLMs—40 percent of the global total.

[Read more](#)

[Back to Main List](#)

B THE LAST BYTE...

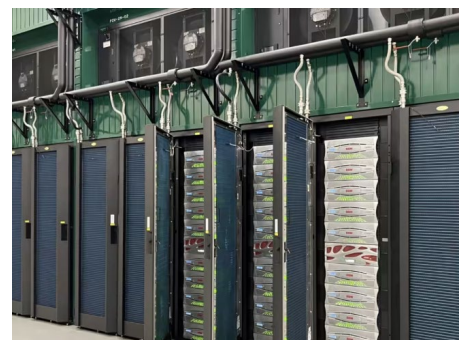
<SHARED CONTENT>

Shared articles and news from the HPC world.

NVIDIA to help Japan build hybrid quantum-supercomputer

Japan's government-backed technology institute will work with Nvidia to build a hybrid supercomputer that offers quantum computing capability for use by researchers and companies.

The National Institute of Advanced Industrial Science and Technology is building a quantum-AI hybrid cloud system called ABCI-Q, as part of the country's quantum computing initiative. Nvidia, the top designer of artificial intelligence chips based in Silicon Valley, is already supplying graphic processing units to the ABCI-Q but will also provide quantum computing software via a cloud service. The system will become available for use for a fee from as early as fiscal 2025. The technology institute envisions such applications as drug research and logistics optimization. Read more [here](#).



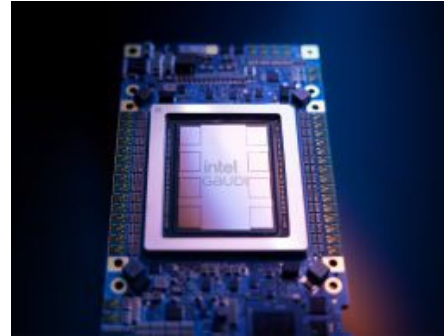
Credit: Nikkei Asia

[Back to Main List](#)

Intel Introduces Gaudi 3, Elevating HPC and AI with Increased Speed and Scalability

Intel introduced the Intel Gaudi 3 accelerator to bring performance, openness and choice to enterprise generative AI (GenAI), and unveiled a suite of new open scalable systems, next-gen products and strategic collaborations to accelerate GenAI adoption.

With only 10% of enterprises successfully moving GenAI projects into production last year, Intel's latest offerings address the challenges businesses face in scaling AI initiatives. The Intel Gaudi 3 AI accelerator will power AI systems with up to tens of thousands of accelerators connected through the common standard of Ethernet. Intel Gaudi 3 promises 4x more AI compute for BF16 and a 1.5x increase in memory bandwidth over its predecessor. The accelerator will deliver a significant leap in AI training and inference for global enterprises looking to deploy GenAI at scale. Read more [here](#).



Credit: HPC Wire

[Back to Main List](#)

AI-powered 'digital twin' of Earth could make weather predictions at super speeds

An AI-driven supercomputer dubbed Earth's 'digital twin' could help us avoid the worst impacts of climate catastrophes headed our way.

The technology could help prevent some of the catastrophic impacts of disasters such as typhoons and flooding. The intensive data-crunching system could also give us a more detailed view of the future effects of climate change and reveal clues about how to mitigate it. The platform, called Earth-2, contains a suite of machine learning technologies developed by NVIDIA, all of it driven by supercomputers to provide the immense processing power needed to do the job. Read more [here](#).



Credit: Live Science

[Back to Main List](#)



Powering Innovation
Supercomputing in Asia

National Supercomputing Centre (NSCC) Singapore

1 Fusionopolis Way, Connexis South, #17-01 Singapore 138632