NARLabs 國家實驗研究院
國家高速網路與計算中心
National Center for High-performance Computing

NCHC N-WAY GPU Bootcamp

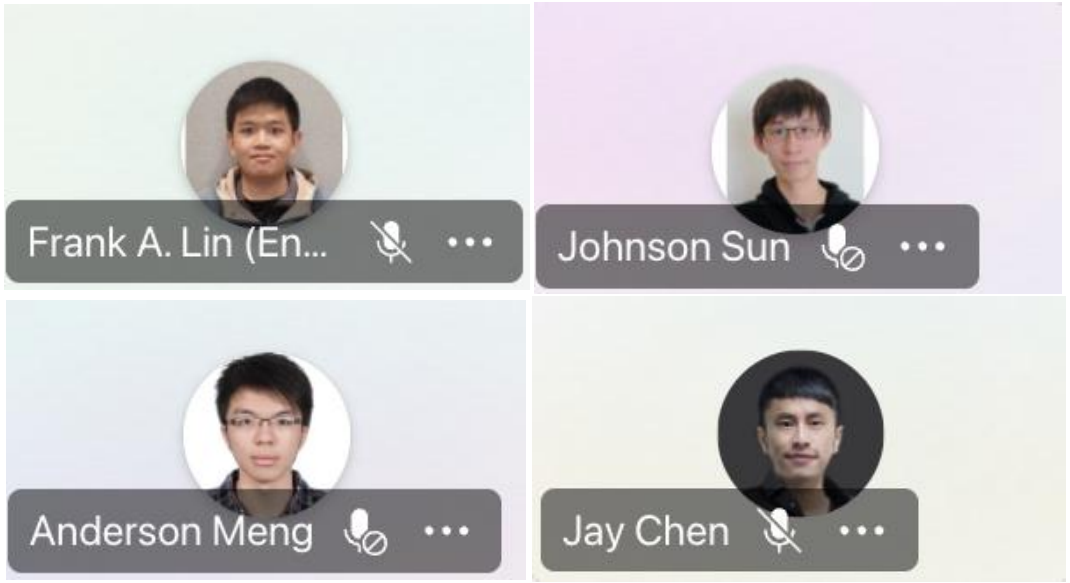
Sep 24-25, 2024

NCHC X OpenACC X NVIDIA

Team Roster

Leo Chen

Instructors



Teaching Assistants

NCHC N-WAY GPU Bootcamp

https://www.openhackathons.org/s/siteevent/a0CUP00000L44xb2AB/se000344

OPEN HACKATHONS

Events Attendees Mentors About Resources

Search... Log In

NCHC N-WAYS TO GPU PROGRAMMING BOOTCAMP

September 24-25, 2024
Application Deadline: September 3, 2024
Hybrid Event

Event Overview

Together with NVIDIA and OpenACC organization, the NCHC (National Center for High-performance Computing) will host a bootcamp on September 24 and 25. Day 1 of the bootcamp will be conducted online, and Day 2 will happen in-person.

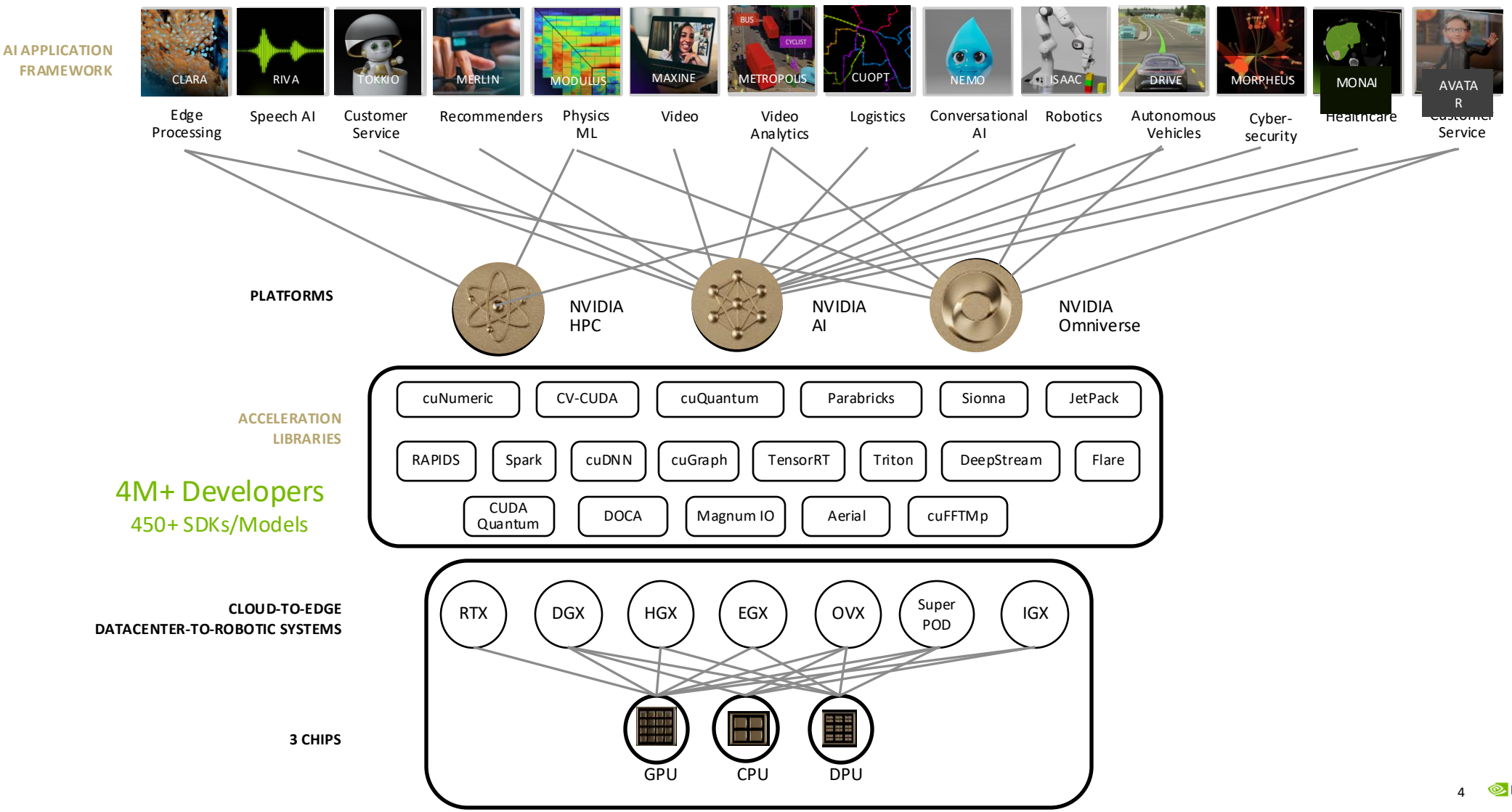
The N-Ways to GPU Programming Bootcamp covers the basics of GPU programming and provides an overview of different methods for porting scientific applications to GPUs using NVIDIA® CUDA®, OpenACC, OpenMP offloading, and/or standard languages. Throughout the bootcamp, attendees will learn how to analyze GPU-enabled applications using NVIDIA Nsight™ Systems and participate in hands-on activities to apply these learned skills to real-world problems.

This in-person Bootcamp is a hands-on learning experience where you will be guided through step-by-step instructions with teaching assistants on hand to help throughout.

***Note:**

- The event will be conducted in Mandarin
- Day 1 (September 24) will be online and is open to everyone. However, Day 2 (September 25) will happen in-person in Taiwan at NAR Labs (Hsinchu Headquarters) No. 7, R&D 6th Rd., East Dist. Hsinchu City, Taiwan, R.O.C. 30076

Platforms for Discovery



NVIDIA Developer Program

NVIDIA 開發者計畫

Program Benefits:

Tools

- 550+ exclusive SDKs and models
- GPU-optimized software, model scripts, and containerized apps
- Early access programs

Training

- Research papers, technical documentation, webinars, blogs, and news
- Technical training and certification opportunities
- 1,000s of technical sessions from industry events On-Demand

Community

- NVIDIA developer forums
- Exclusive meetups, hackathons, and events

Special Program (Present to Jan 2024)

- Join NVIDIA Developer program now, you will get one NVIDIA Training

Join the Community

NCHC-NVIDIA Joint Lab

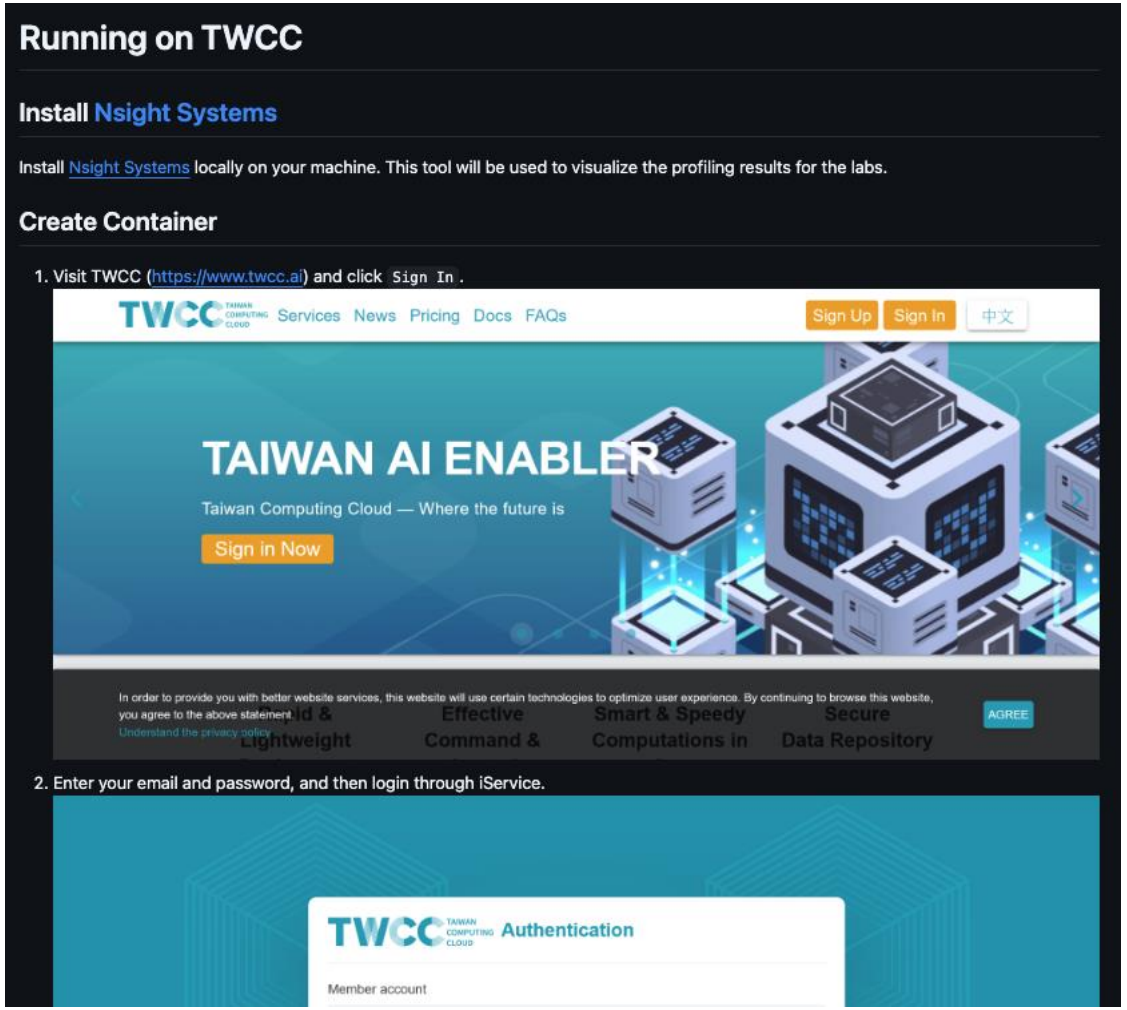
https://github.com/nqobu/nvidia/tree/main

Repository for NCHC-NVIDIA Joint Lab

- tutorials - NVIDIA online courses/tutorials in AI/HPC
- 20210412 - NVIDIA Techniques Sharing 2021
- 20210706 - NCHC Techniques Sharing 2021
- 20211202 - NVIDIA Techniques Update 2021
- 20211221 - AI+HPC: 利用 NVIDIA Modulus 實踐 PINN 於物理模擬
- 20220415 - NVIDIA Techniques Update 2022
- 20220530 - NCHC-NVIDIA Techniques Sharing 2022
- 20220629 - PINN 與 NVIDIA Modulus 實作訓練營
- 20221111 - Quantum Computing Workshop / 量子計算模擬實作
- 20230413 - NVIDIA Techniques Sharing 2023
- 20230517 - NVIDIA Techniques Briefing: NVIDIA Federated Learning
- 20230525 - AI for Science: NVIDIA Modulus 及 NVIDIA Omniverse 實作
- 20230727 - N-Way to GPU Programming Bootcamp / 多 GPU 程式設計訓練課程
- 20230821 - NVIDIA Techniques Salon 2023: Programming the NVIDIA Superchip
- 20231207 - NCHC Open Hackathon 2023
- 20240410 - NCHC Quantum Computing Bootcamp 2024 - NVIDIA CUDA-Q and cuQuantum
- 20240506 - AI for Science: NVIDIA Modulus, NVIDIA Omniverse, and NVIDIA Earth-2
- 20240508 - NCHC Techniques Sharing 2024
- 20240626 - NCHC AI for Science Bootcamp 2024 - NVIDIA Modulus 物理模擬計算
- 20240806 - NCHC End-to-end LLM Bootcamp 2024 - NVIDIA NeMo 大型語言模型框架
- 20240924 - NCHC N-Way Bootcamp 2024 – NVIDIA GPU 加速運算

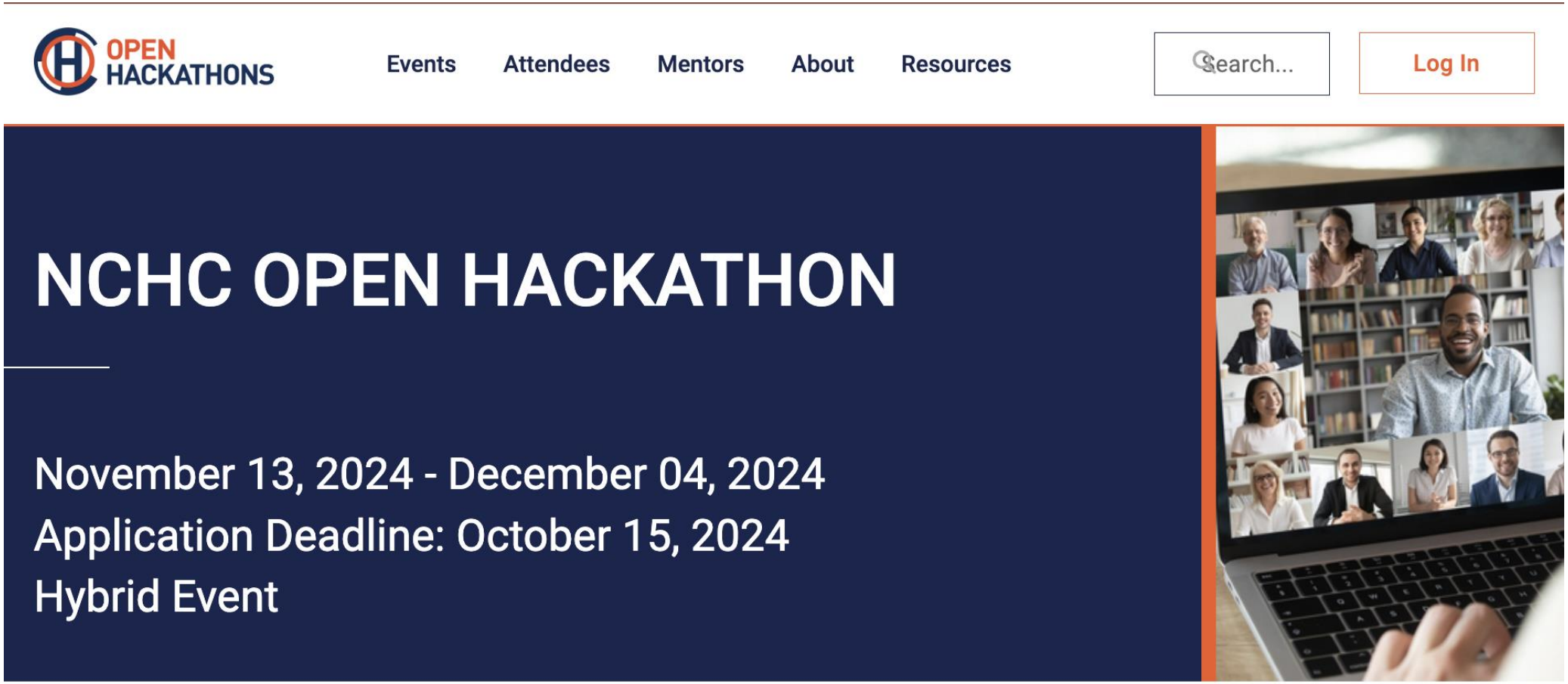
TWCC Tutorials

https://github.com/j3soon/nways_accelerated_programming/blob/main/README_TWCC.md



2024 NCHC Open Hackathon Registration

https://www.openhackathons.org/s/siteevent/a0C5e000008dX2GEAU/se000298



OPENACC – CELEBRATING 11 YEARS

Building Community.





Ecosystem Development	Training/Education	OpenACC Specification

OPEN HACKATHONS
OpenACC

Open Hackathon Objectives

Connect	Accelerate	Celebrate
Developers & Mentors Apps & Acceleration	Speedup Energy Efficient	Publication, Co-Paper Blogs and Talks

2021 NCHC Open Hackathon

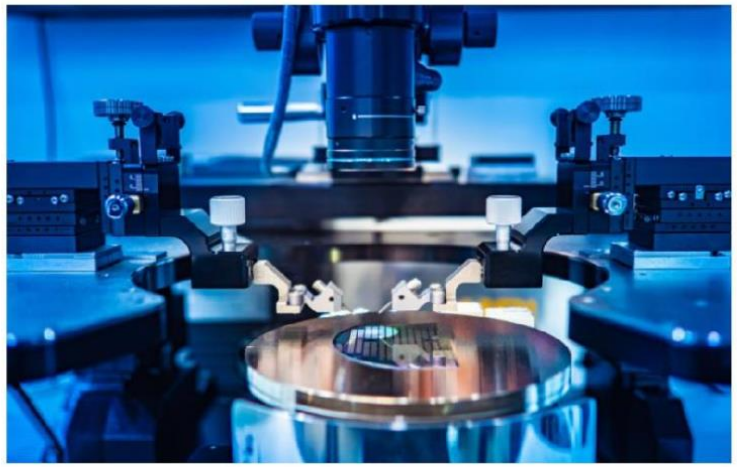
【NVIDIA 在台灣辦了一場黑客松】看 GPU 把地球科學、工廠產線排程等不同任務，通通加速！

NVIDIA 在台灣舉辦了一場熱血線上黑客松，從產業升級到基礎科學研究通通可以加速！

NVIDIA 2022-01-05

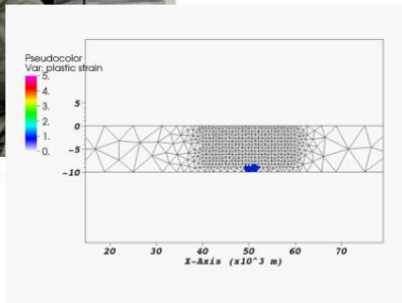
智慧製造的關鍵項目，產線智慧排程如何優化？

由林群惟博士、蘇粲程博士所帶領的團隊「AI Scheduler」，首度參與本屆黑客松，就透過 GPU 加速找到未來商品優化的方向。



GPU 加速深化台灣地球科學研究成果，提升全球學術圈重要性

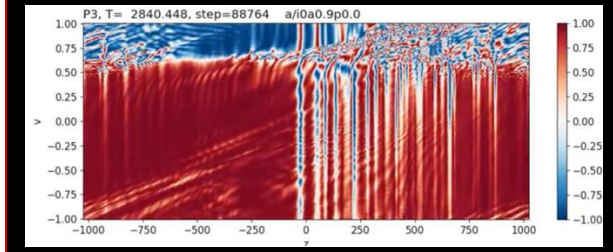
譚老師團隊照片，譚謂（圖中）帶領團隊參與本次 NVIDIA 黑客松，找到地質研究的運算新方法



https://buzzorange.com/techorange/2022/01/05/nvidia-2021-taiwan-gpu-hackathon/

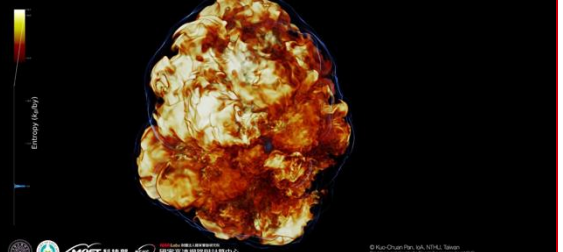
2022 NCHC Open Hackathon

COSEnu



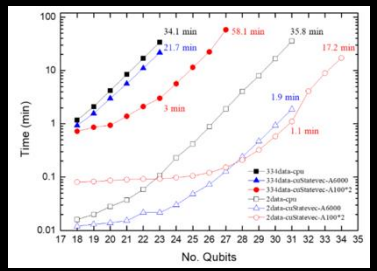
novel neutrino transport mechanism
Accelerate 7X by OpenACC @ GSI, SINICA, NCHC

CEAG



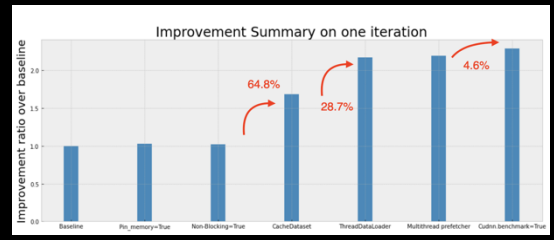
core-collapse supernova shock
Accelerate 10X by OpenACC @ NTHU

QKBA



predict tumor metastasize by QSVM
Accelerate 33X by cuQuantum @ NSRRC, NYCU, AU, ICL

GFEverything



medical semantic segmentation by HarDNet
Accelerate 2X by MONAI @ NTHU

OpenACC

2023 NCHC Open Hackathon

https://www.nchc.org.tw/Message/MessageView?id=3874&menutype=0&sitemenuid=5&mid=46

NARLabs 財團法人國家實驗研究院
國家高速網路與計算中心
National Center for High-performance Computing

::: | | | EN | A- A A+ |

核心服務 創新技術 科研成果 動態資訊 關於我們

看 OPEN 黑客松如何帶領了技術變革? DPU把網路，GPU把大型語言模型、大氣科學、量子電路模擬，通通加速！

2023.12.07

動態資訊

- 焦點新聞
- 活動專區
- 動態集錦
- 招標公告
- 出版品專區
- 專利/商標與技術移轉
- 影音專區
- 輿情回應

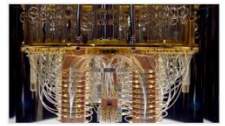
NARLabs 財團法人國家實驗研究院
國家高速網路與計算中心
National Center for High-performance Computing

::: > 科研成果 > 學研成果

學研成果

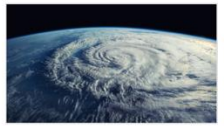
日期 開始日期 ~ 結束日期 關鍵字 關鍵字

ex：2019/01/01



2024.01.30

【2023 NCHC, NVIDIA, OpenACC 黑客松】-HPC、DPU及量子運算加速成果



2024.01.30

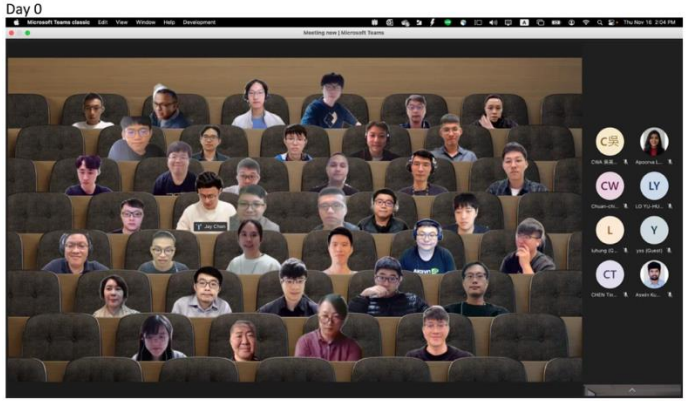
【2023 NCHC, NVIDIA, OpenACC 黑客松】-大氣科學應用加速成果

2024.01.30

【2023 NCHC, NVIDIA, OpenACC 黑客松】-人工智慧應用加速成果

2023 NCHC Open Hackathon

Day 0



Final Day – AM Sessions

Highlights: (*Note most benchmarks results come with specific conditions.)

- "468X vs CPU" – Team 2 haofan2023
- "3X by TensorRT-LLM on TAIDE model" – Team 12 NCHC-Speedrunning
- "126X, the CPU implementation is significantly less suitable in time critical applications." – Team 1 Schrödinger s cat
- "Speed up about 8x with only 25 lines modification. CPU cost 10+ months, we never think about re-generate mesh, now we can!" – Team 5 CWA mesh generation
- See below **Team Results/Outcomes** for more!

*Note most benchmarks results come with specific conditions.

Team Name	Mentors	Code Area of Focus	Domain	Languages /Libs	Results	Energy Efficiency	Comments
1 Schrödinger s cat	Reese Yun-Yuan	Quantum-Inspired Algorithm (QUBO)	HPC Optimization	C++ /JAX	126X	126X	The CPU implementation is significantly less suitable in applications where timeliness is critical, such as high-frequency trading or real-time route planning, which demand rapid decision-making within a limited timeframe.
2 haofan2023	Tian Frank Yun-Yuan	Quantum Circuit Simulation (QFT, QAOA)	HPC Quantum	C++	468X (vs 1 core)	65X	Benchmark up to 30 qubits. They could increase up to 39 qubits with special SSD cache optimization innovation.
3 NTHU-LSALAB	Sungta Erez	5G SBA (Service Based Architecture)	HPC Network Infra	C++ /DOCA	DPU Save CPU 15% reduce 11X latency	N/A	Utilizing the natively provided high-efficiency SR-IOV network architecture to increase the network speed of 5G Service. Preserving a significant amount of CPU processing resources for 5G SBA components.
4 NTUST CFD Lab	Bharat Shijie Kuan-Ting	3D-CFD (Direct Forcing Immersed Boundary, LES turbulence model)	HPC CFD	C++ /OpenACC	16.7X	46X	This marks another milestone for our in-house code as large-scale computations are required to simulate real-world cases. The significant speedup achieved is a substantial leap forward for us in reaching that goal.
5 CWA mesh generation	Leo Jay	Mesh generation for MPAS model	HPC CWO	Fortran /OpenACC	12X (vs pre-event)	22X	Speed up about 8x with only 25 lines modification. High-res mesh generation takes estimiated 10+ months, now <1 month with GPU, which is incredible. we never think about re-generate mesh, now we can!
6 CYCU BME	Eason	Otoscopy Diagnosis	AI Healthcare	/Pytorch TensorRT	1.5X	N/A	Model inference optimization implementation based on tensorRT.
7 CWA_GVER	Ming Kuan-Ting Jay	Global Ensemble model Verification	HPC CWO	python /JAX	44X (wo I/O)	N/A	Verification costs 12+ hours, which is not feasible with CPU only in future operation, not to mention higher resolution and more ensemble members.
8 WTMH	Ying-Kai	Arrhythmia Screening of Real-Time Single-Lead ECG	AI Healthcare	/Pytorch TensorRT Triton	40X	N/A	Achieved "instant" ECG analysis system by TensorRT+Triton.
9 氣興聯隊	Leo Jay	CWAGFS-TCo	HPC CWO	Fortran /OpenACC	1.8X (wo I/O)	1.6X	The first and only in-development gpu-accelerated NWP in CWA.
10 YSS	CK Tian Frank	Functional Encryption (BSGS) apply in Machine Learning Service	AI Data Privacy	C++	2.2X	3.1X	It is important for server to provide fast and secure service, Functional Encryption is one of the cryptosystem mechanisms used for data privacy In MaaS.
11 TXM-AI	Warren	X-ray Background Correction Model	AI Healthcare	/MONAI	30X	21X	This will save considerable acquisition time and enhance the efficiency of nano CT scans, ultimately accelerating progress in paleontology, biomedical research, materials science, energy, and other related fields.
12 NCHC-Speedrunning	Anthony Cliff	LLM inference with TensorRT-LLM on NCHC servers	AI LLM	/TensorRT-LLM	3X (bs=1)	N/A	Test on TAIDE Model, Taiwan ChatGPT project, based on LLaMA2-7B. could be further improved on H100 in near future.

NARLabs 國家實驗研究院
國家高速網路與計算中心
National Center for High-performance Computing

(extend from 2022 Open Hackathon)

cuTN-QSVM

https://arxiv.org/abs/2405.02630, 4th May 2024.
https://github.com/Tim-Li/cuTN-QSVM

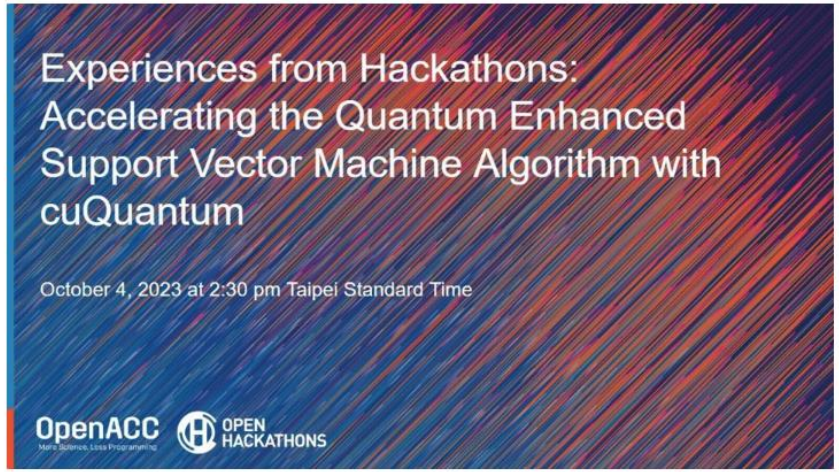
國家同步輻射研究中心
National Synchrotron Radiation Research Center

Imperial College London

Wednesday Oct 4, 2023

02:30 - 02:50 PM

Experiences from Hackathons: Accelerating the Quantum Enhanced Support Vector Machine Algorithm with cuQuantum



Speaker

Tai Yue Li
National Synchrotron Radiation Research Center (NSRRC)

cuTN-QSVM: cuTensorNet-accelerated Quantum Support Vector Machine with cuQuantum SDK

Kuan-Cheng Chen**
QuEST
Imperial College London
London, United Kingdom
kc2816@ic.ac.uk

Tai-Yue Li**
National Synchrotron Radiation Research Center
Hsinchu, Taiwan
li.timty@nsrrc.org.tw

Yun-Yuan Wang**
NVIDIA AI Technology Center
NVIDIA Corp.
Taipei, Taiwan
yunyuanw@nvidia.com

Simon See
NVIDIA AI Technology Center
NVIDIA Corp.
Santa Clara, CA, USA
ssee@nvidia.com

Chun-Chieh Wang
National Synchrotron Radiation Research Center
Hsinchu, Taiwan
wang.jay@nsrrc.org.tw

Robert Willie
Chair of Design Automation
Technical University of Munich
Munich, Germany
robert.wille@tum.de

Nan-Yow Chen
National Center for HPC
Narlabs
Hsinchu, Taiwan
nanyow@nchc.narl.org.tw

An-Cheng Yang
National Center for HPC
Narlabs
Hsinchu, Taiwan
1203087@nchc.narl.org.tw

Chun-Yu Lin
National Center for HPC
Narlabs
Hsinchu, Taiwan
lincy@nchc.narl.org.tw

README Code of conduct Apache-2.0 license

License Apache 2.0 arXiv 2405.02630

cuTN-QSVM: cuTensorNet-accelerated Quantum Support Vector Machine with cuQuantum SDK

Welcome to the official repository of cuTN-QSVM, featuring fast GPU simulators for benchmarking Quantum Support Vector Machines (QSVMs) and scripts for generating compatible quantum circuits for hardware execution. Facilitated by NVIDIA's cuQuantum SDK and the cuTensorNet library, this project integrates cutting-edge quantum computing technologies with high-performance computing systems, enhancing quantum machine learning's efficiency and scalability to new heights.

Project Overview

Quantum Support Vector Machines (QSVMs) utilize a quantum-enhanced approach to tackle complex, multidimensional classification problems, surpassing the capabilities of classical SVMs under certain conditions. However, prior to the advent of large-scale quantum systems, the scalability of simulating QSVMs on CPUs was traditionally limited by the exponential growth in computational demands as qubit counts increased. By employing NVIDIA's cuQuantum SDK and the cuTensorNet library, cuTN-QSVM effectively reduces this computational complexity from exponential to quadratic. This enables the simulation of large quantum systems of up to 784 qubits on the NVIDIA A100 GPU within seconds.

Mentored by NVIDIA Pika Wang, Leo Fang, Jay Chen

(extend from 2023 Open Hackathon)

Queen: A quick, scalable, and comprehensive quantum circuit simulation for supercomputing

https://arxiv.org/abs/2406.14084

Welcome to compare to our state-of-the-art simulator released in 2024 or offer paid consultation for performance improvement.

Chuan-Chi Wang
National Taiwan University
Taipei, Taiwan
d10922012@ntu.edu.tw

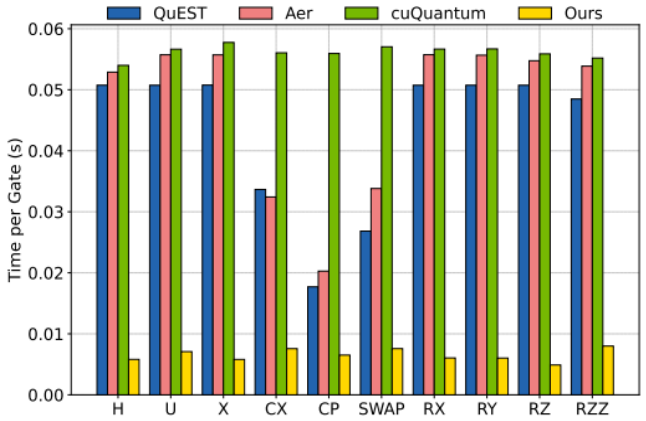
Yu-Cheng Lin
National Taiwan University
Taipei, Taiwan
r11922015@csie.ntu.edu.tw

Yan-Jie Wang
National Taiwan University
Taipei, Taiwan
yanjiewtw@gmail.com

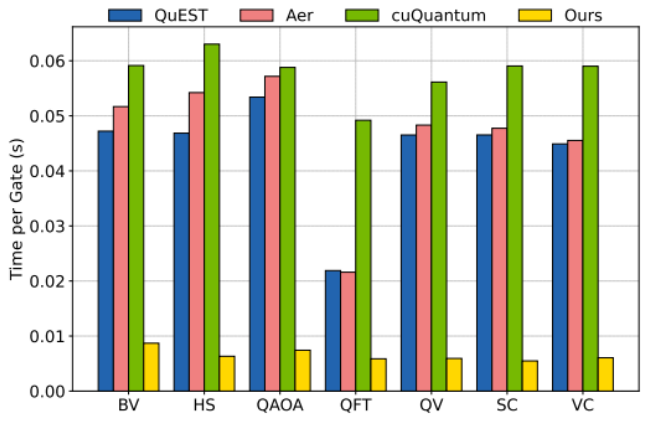
Chia-Heng Tu
National Cheng Kung University
Tainan, Taiwan
chiaheng@ncku.edu.tw

Shih-Hao Hung
National Taiwan University
Taipei, Taiwan
hungsh@csie.ntu.edu.tw

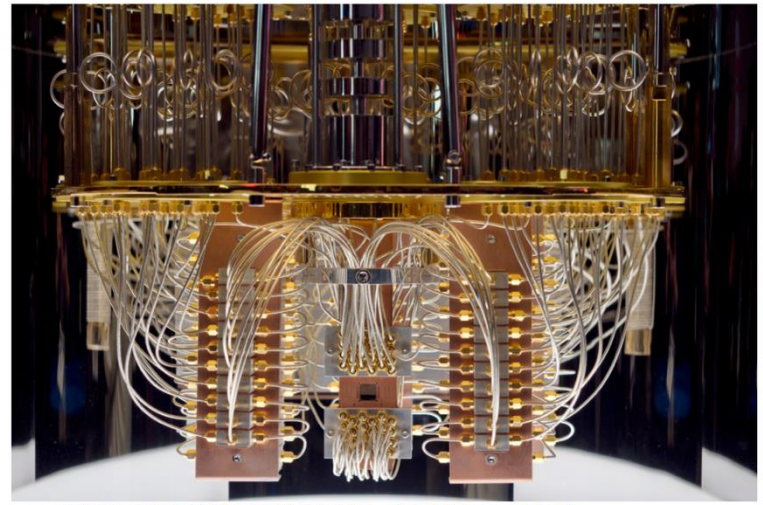
We thank the National Center for High-Performance Computing for providing access to the NVIDIA DGX-A100 workstation.



(a) Gate benchmark.



(b) Circuit benchmark.



haofan2023團隊成員來自臺灣大學資工系「洪士灝老師實驗室」，將量子演算法QAOA加速468倍！

— NVIDIA Mentors: Tian Zheng, Frank Lin, Yun-Yuan Wang

量子技術正以驚人的速度發展，預示著我們即將進入量子計算的時代。在這個過程中，量子電路模擬成為一個關鍵工具，它在量子硬體和軟體的開發中扮演著重要的角色，特別是在處理量子程式的編寫和驗證方面。傳統電腦的強模擬能夠獲得完整的量子狀態信息。這使得傳統電腦在構建量子系統方面變得不可或缺，尤其是在當前嘈聲較多的中等規模量子（NISQ）時代。

量子近似優化算法（QAOA）是一種常用的量子算法，用於通過近似解來解決組合優化問題。然而，在虛擬量子計算機上執行QAOA對於解決需要大規模量子電路模擬的組合優化問題而言，會遇到模擬速度較慢的問題。團隊使用數學優化來壓縮量子操作，並結合有效的位元操作進一步降低計算複雜性，透過GPU加速最高獲取468倍的加速效果！

Table 1: The elapsed time of 5-level QAOA (unit: second, double).

Qubit	CPU_{Single}	$CPU_{Mutiple}$	CPU_{Cache}	GPU_{Cache}	GPU_{All}
23	29.80	1.28 (23x)	1.28 (63x)	0.24 (120x)	0.06 (**341x**)
24	68.00	3.46 (20x)	3.46 (43x)	0.55 (123x)	0.12 (**382x**)
25	152.52	15.32 (10x)	15.31 (45x)	1.19 (127x)	0.23 (**404x**)
26	330.69	33.83 (10x)	33.83 (56x)	2.60 (126x)	0.56 (**417x**)
27	712.26	72.66 (10x)	72.66 (54x)	5.59 (127x)	1.08 (**427x**)
28	1556.87	156.52 (10x)	156.52 (54x)	11.96 (130x)	2.17 (**445x**)
29	3325.55	335.09 (10x)	335.09 (49x)	25.73 (129x)	4.45 (**451x**)
30	7226.46	718.33 (10x)	718.33 (47x)	55.20 (130x)	9.22 (**468x**)

更多資訊請看：https://github.com/nqobu/nvidia/raw/main/20231207/Team02.pdf

Mentored by NVIDIA Frank Lin, Tian Zheng, Pika Wang

(extend from 2023 Open Hackathon)

NCAR WRF/MPAS Users Workshop 2024

https://www.mmm.ucar.edu/events/workshops/wrf-mpas

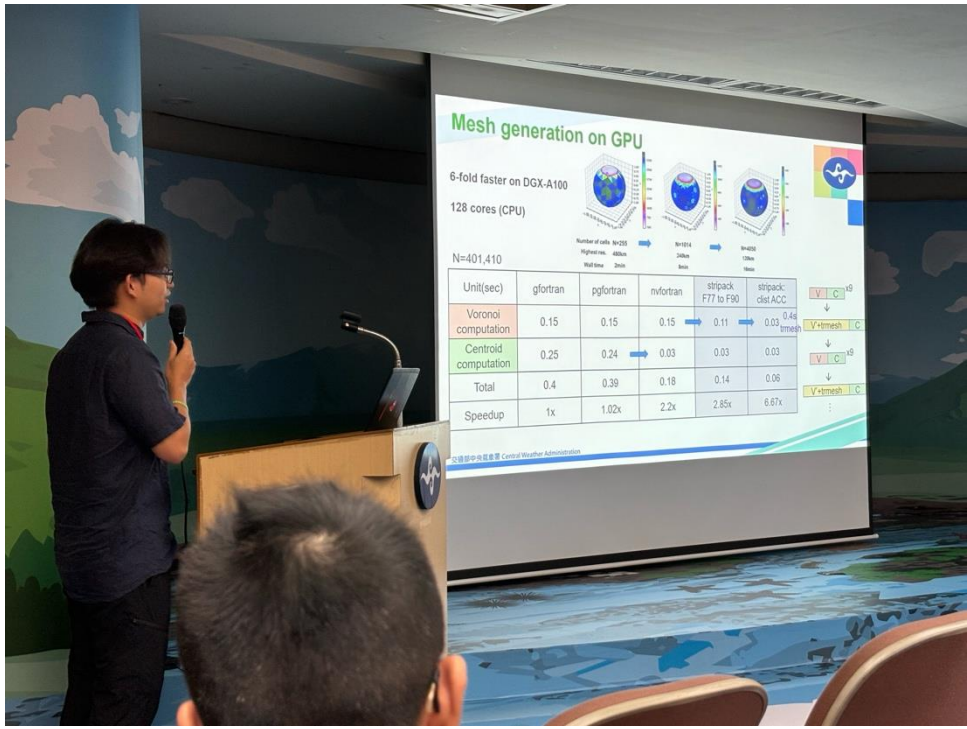
Advancements in Implementing the MPAS-A Regional Model at the Central Weather Administration

[1]Wu, Y.-J., [2]W. Wang,[3]C.-Y. Chen, [3]Y.-L. Chen, [1]S.-L. Huang [1]B.-S. Lin, [1]L.-F., Hsiao
[1]Central Weather Administration, Taiwan
[2]University Corporation for Atmospheric Research, Boulder, Colorado
[3]NVIDIA

In our ongoing efforts to enhance weather modeling capabilities at the Central Weather Administration (CWA), we have integrated several key components from the Weather Research and Forecasting (WRF) operational model into the MPAS-A Regional Model. Collaborating with National Taiwan University (NTU), we have developed the cloud physics scheme TCWA1, which has been integrated into the model framework.

One key focus of our work has been comparing the surface wind speeds simulated by MPAS-A with those from the Weather Research and Forecasting (WRF) model. Our analyses have revealed notable discrepancies between the two models, prompting the introduction of the topo_wind option from WRF. This addition aims to mitigate wind speed biases and improve the overall accuracy of our simulations, particularly in complex terrain regions.

The grid generation program plays a critical role in setting up the computational grid for weather simulations. By leveraging GPU acceleration, we achieved significant improvements in the performance of this program. The tailored GPU acceleration techniques, developed in collaboration with the mentors in NVIDIA workshop, allowed for faster data processing and computation, leading to a six-fold speed increase compared to the previous implementation.



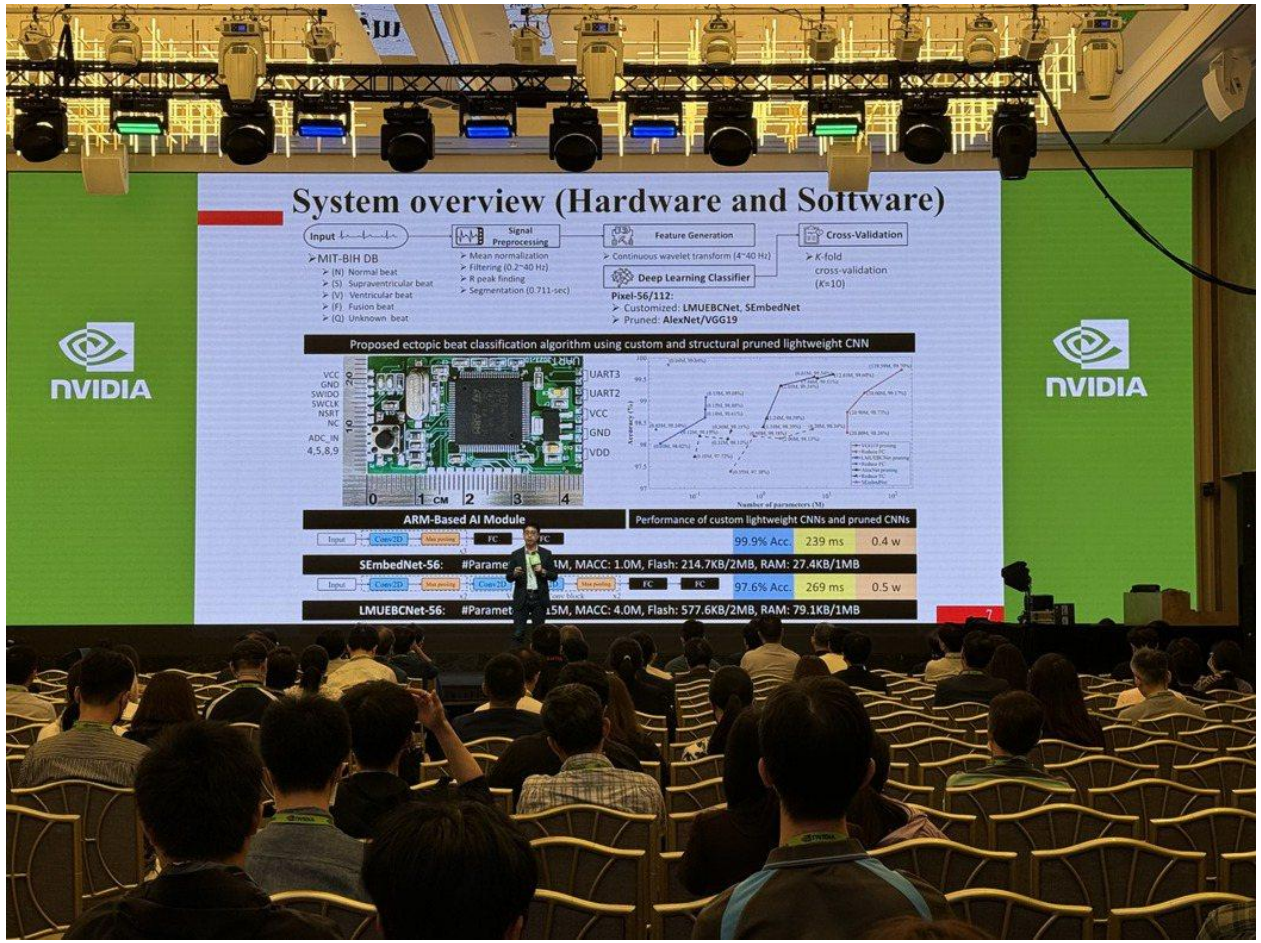
(Present in CWA Conference 2024)

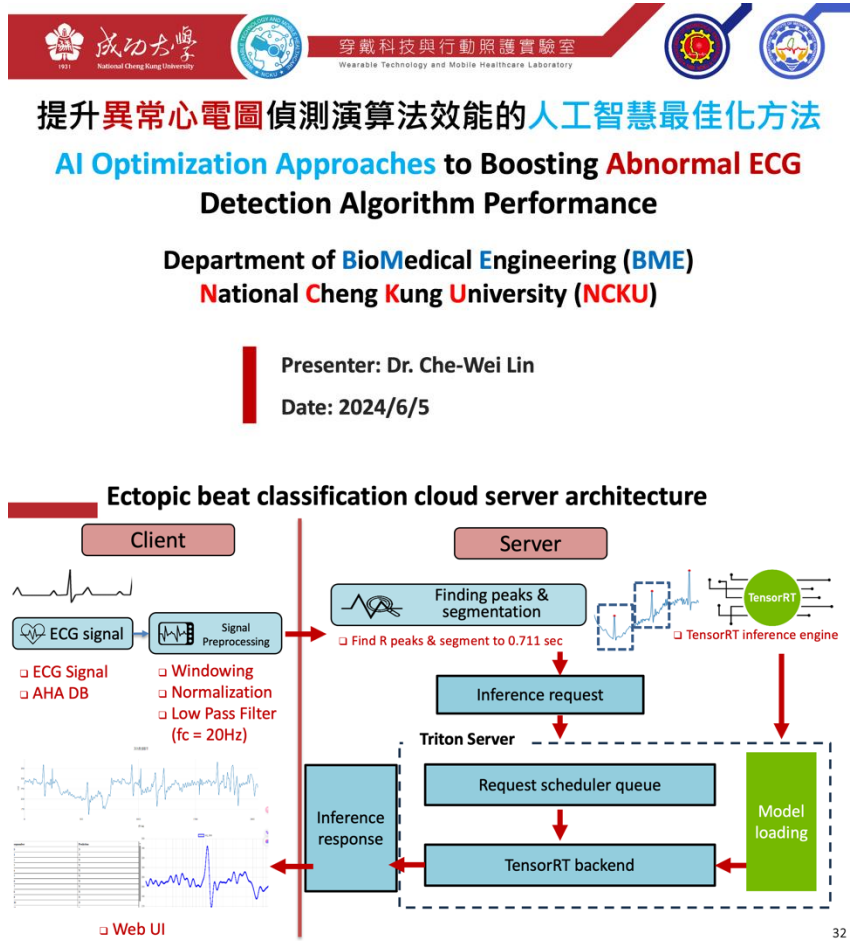
Mentored by NVIDIA Leo Chen, Jay Chen

(extend from 2023 Open Hackathon)

Computex 2024 AI Summit Talk

Prof. Che-Wei Lin @ NCKU-BME

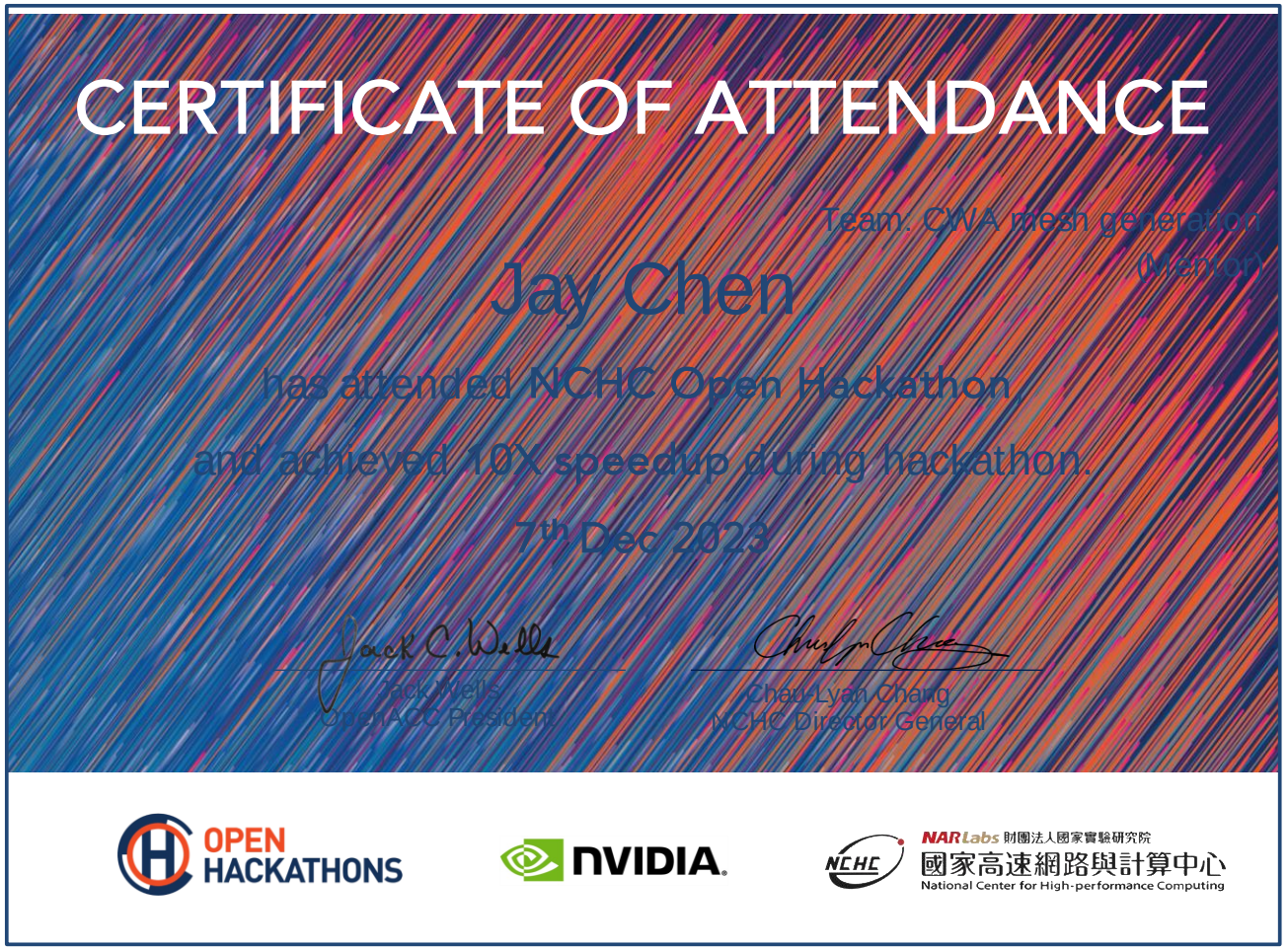




Mentored by NVIDIA Ken Liao

Certification

https://github.com/nqobu/nvidia/blob/main/20231207/Certificate_Of_Attendance.pdf



2024 NCHC Open Hackathon Registration

https://www.openhackathons.org/s/siteevent/a0C5e000008dX2GEAU/se000298

Agenda

**All times are in Taipei time zone*

Day 0: Kick-off Meeting: Wednesday, November 13, 2024 // 02:00 PM - 06:00 PM (Virtual)

- 02:00 PM - 02:15 PM: Welcome and event overview
- 02:15 PM - 03:15 PM: Team self-introduction and getting to know the mentor. (5 minutes per team)
- 03:15 PM - 04:00 PM: Introduction to computing resources
- 04:00 PM - 05:00 PM: Introduction to Nsight Analysis Tools
- 05:00 PM - 06:00 PM: Intra-team Discussion

Day 1: Wednesday, November 20, 2024 // 02:00 PM - 03:00 PM (Virtual)

- 02:00PM - 03:00PM: Scrum #1 (5 mins presentation per team)

Day 2: Wednesday, November 27, 2024 // 02:00 PM - 03:00 PM (Virtual)

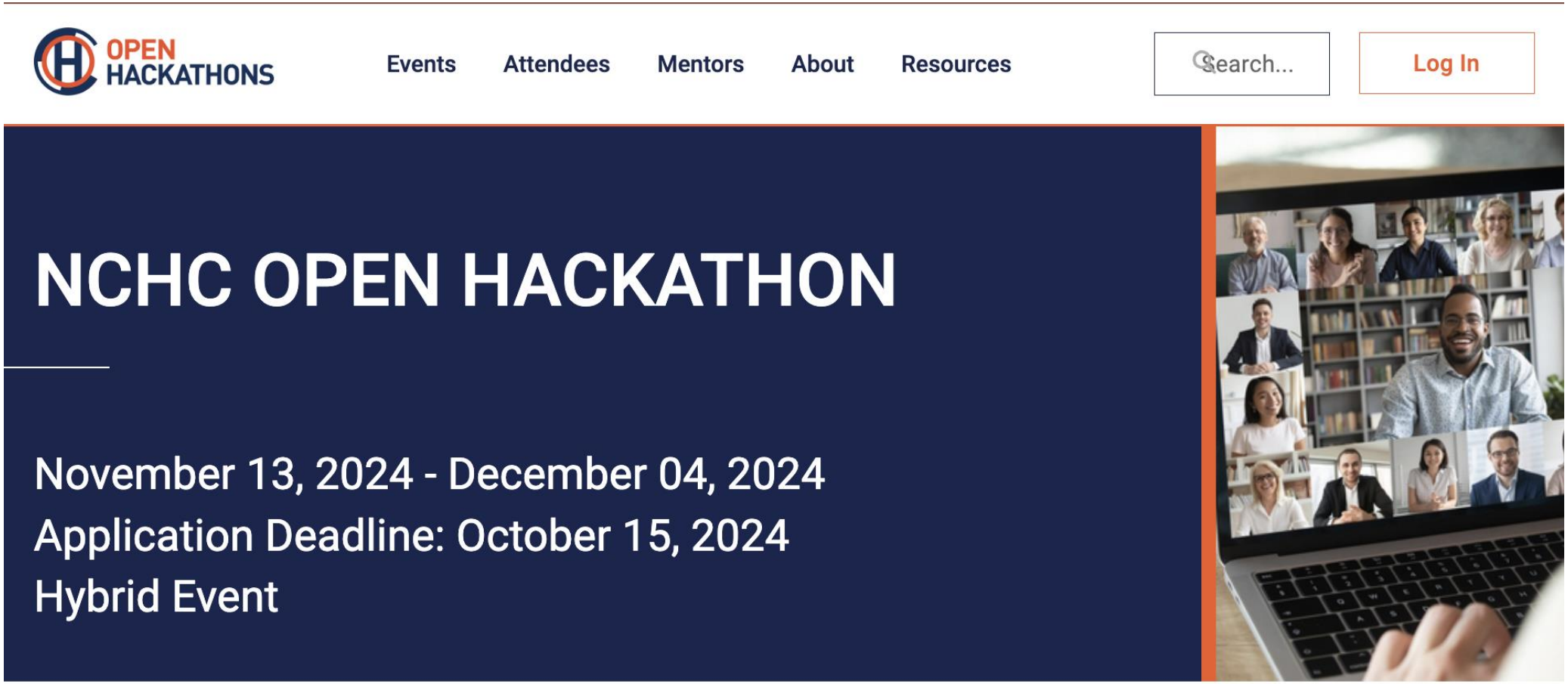
- 02:00PM - 03:00PM: Scrum #2 (5 mins presentation per team)

Day 3: Wednesday, December 4, 2024 // 10:00 AM - 04:00 PM (In-person)

- 10:00 AM - 10:30 PM: Welcome and event description
- 10:30 AM - 12:00 PM: Final presentation (12 presentation +3 QA minutes per team)
- 12:00 PM - 01:30 PM: Lunch time
- 01:30 PM - 03:00 PM: Final presentation (12 presentation +3 QA minutes per team)
- 03:00 PM - 04:00 PM: Wrap-up session

2024 NCHC Open Hackathon Registration

https://www.openhackathons.org/s/siteevent/a0C5e000008dX2GEAU/se000298



Feedback Survey

https://forms.office.com/r/c4fe5h73CQ

Your feedback is valuable to us! Please take a moment to share your thoughts on this event. Your input will help us enhance future events. Thank you for participating!