Scalable Quantum Chemistry Simulation Using DMET-VQE Framework with CUDAQ

2024 NCHC Open Hackathon

Team 7 **NoLab**

2024.12.04

Members

張仁瑀 (Ran-Yu Chang)
Arete Honors Program
NYCU

陳思頤 (Ssu-Yi Chen)
Graduate Institute of Electronics Engineering
NTU

張沂竣 (Yi-Chun Chang)
Arete Honors Program
NYCU

徐培哲 (Pei-Che Hsu)
College of Electrical and Computer Engineering
NYCU

劉珍君 (Zhen-Jun Liu)
Department of Physics
NTNU

鍾宇程 (Yu-Cheng Chung)
Department of Physics
NTNU

Mentors: Pika Wang, Ikko Hamamura, Tian Zheng

Motivation

Variational Quantum Eigensolver (VQE)

A powerful approach for simulating molecules and materials in the NISQ era.

Energy Expectation Value

$$E(\theta) = \langle \psi(\theta) | H | \psi(\theta) \rangle$$

Variation Principle

$$E_0 = \text{minimize} E(\theta)$$

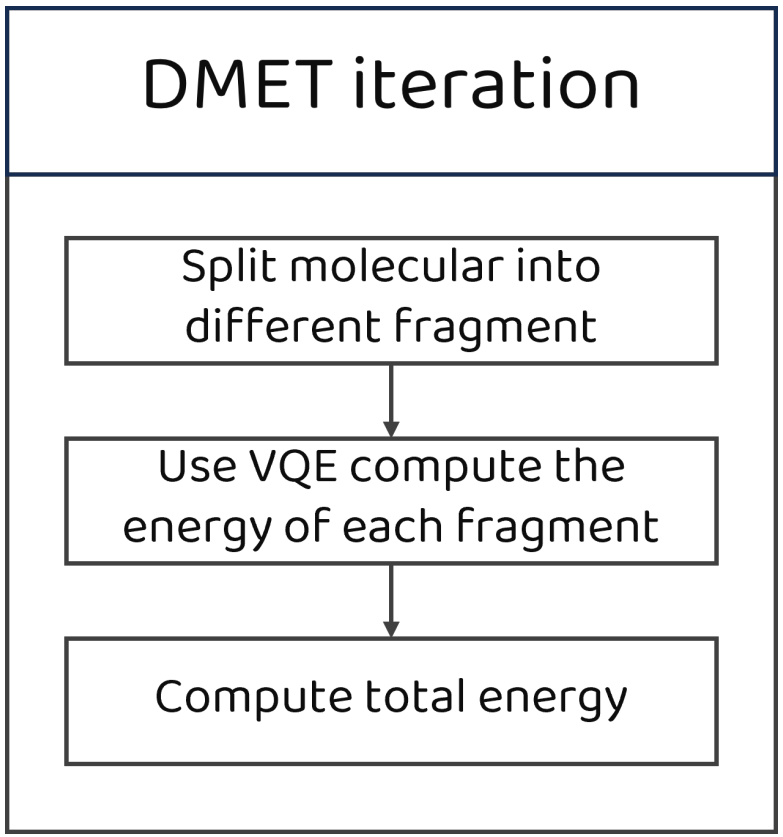
Challenge

Real Quantum Computer

- Limited by noise, scalability, and measurement overhead.

Simulation Challenges

- As qubit count increases, simulation complexity grows **exponentially**, making large systems challenging
- Examples:
  - **H_2 (4 qubits)** Easily simulable.
  - **H_{10} (20 qubits)** Challenging for classical simulators.



Our Approach

GPU simulation via CUDA-Quantum

- Single-GPU simulation
- Multi-GPU simulation

Density Matrix Embedding Theory (DMET)

Calculate the molecular **partially**

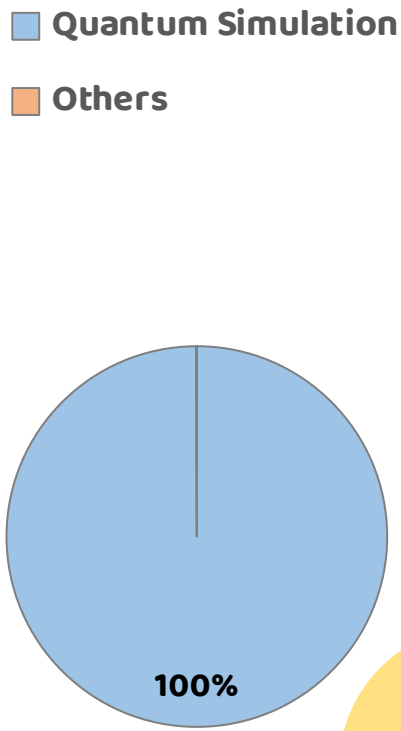
How Many Times We Speed Up

Profiling (CPU)

	Total Time (sec.)	Each Iteration Time (sec.)
H_2 (4)	2658.8	8.1
H_2O (14)	168225.3	841.2
CH_4 (18)	2037372.6	10186.8
H_{10} (20)	3264852.7	16324.2
C_2H_6 (32)	CAN'T RUN	CAN'T RUN

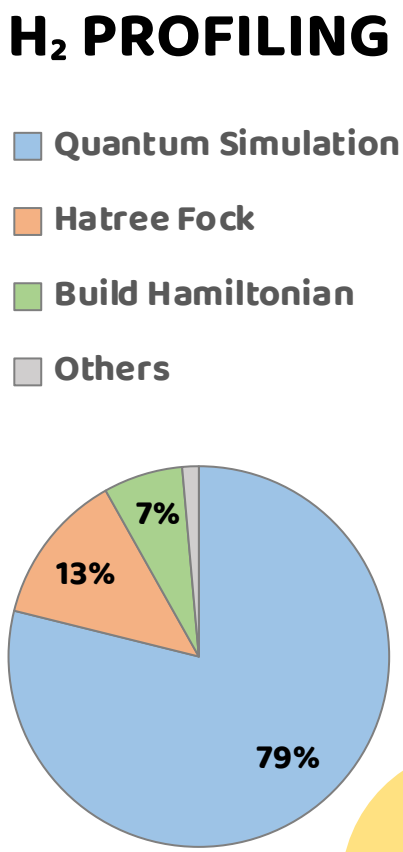
Blue: estimated time

H_2 PROFILING



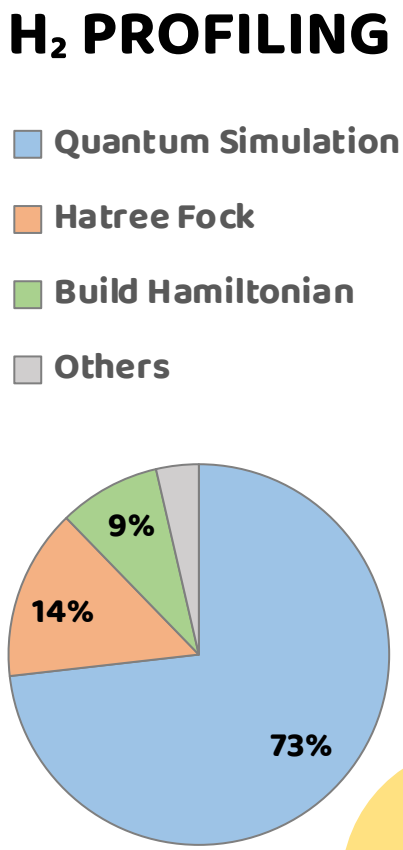
Profiling (Single-GPU)

	Total Time (sec.)	Each Iteration Time (sec.)
H_2 (4)	17.598	0.098
H_2O (14)	497.401	2.487
CH_4 (18)	5518.200	27.591
H_{10} (20)	11801.014	59.005
C_2H_6 (32)	CAN'T RUN	CAN'T RUN



Profiling (Multi-GPU)

	Total Time (sec.)	Each Iteration Time (sec.)
H_2 (4)	12.784	0.065
H_2O (14)	335.008	1.675
CH_4 (18)	4399.789	21.999
H_{10} (20)	9868.743	49.344
C_2H_6 (32)	CAN'T RUN	CAN'T RUN



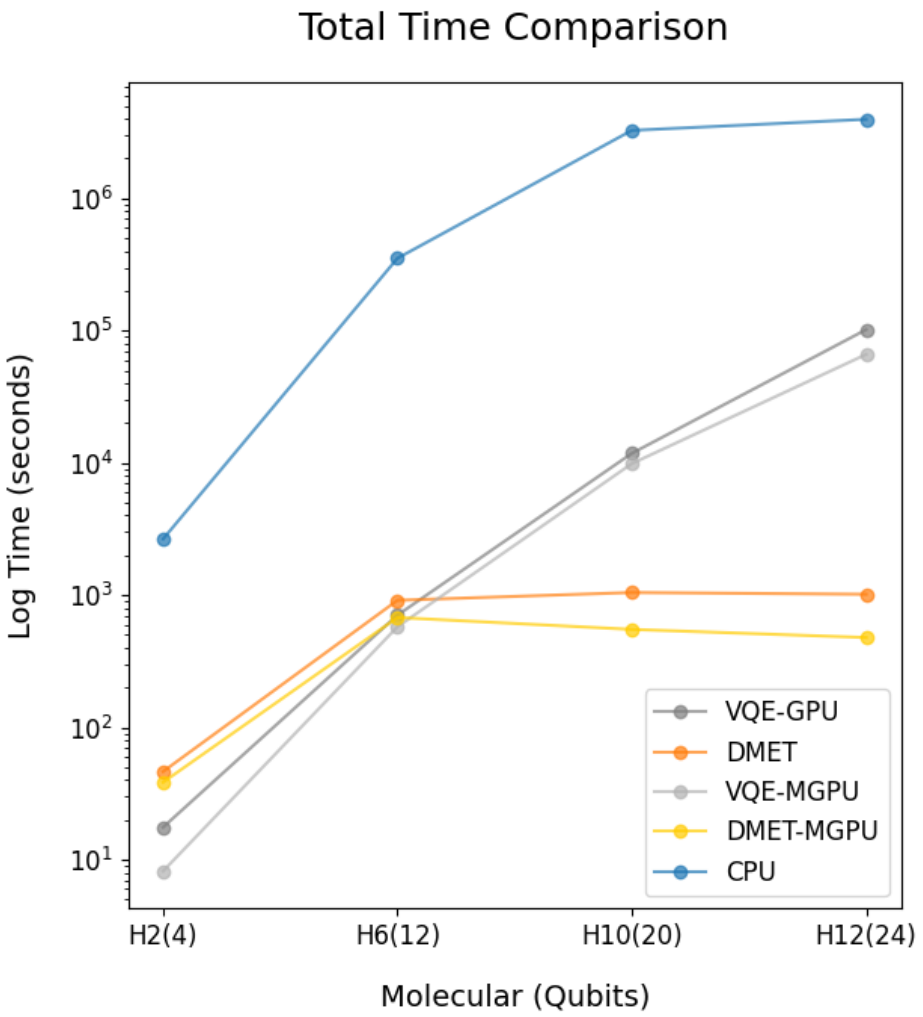
Profiling Tables

	CPU (sec.)	Single-GPU (sec.)	Multi-GPU (sec.)	Speed Up
H_2 (4)	8.165	0.098	0.065	125x
H_2O (14)	841.275	2.487	1.675	502x
CH_4 (18)	10186.863	27.591	21.999	463x
H_{10} (20)	16324.244	59.005	49.344	330x
C_2H_6 (32)	CAN'T RUN	CAN'T RUN	CAN'T RUN	CAN'T RUN

DMET Profiling (Single-GPU)

	Total Time (sec.)	Each DMET Iteration Time (sec.)	Number of DMET iterations
H_2O (14)	12370.513	1767.216	7
CH_4 (18)	11149.755	2229.950	5
H_{10} (20)	1047.118	87.260	12
H_{12} (24)	1015.835	101.584	10
C_2H_6 (32)	83968.001	6997.333	12

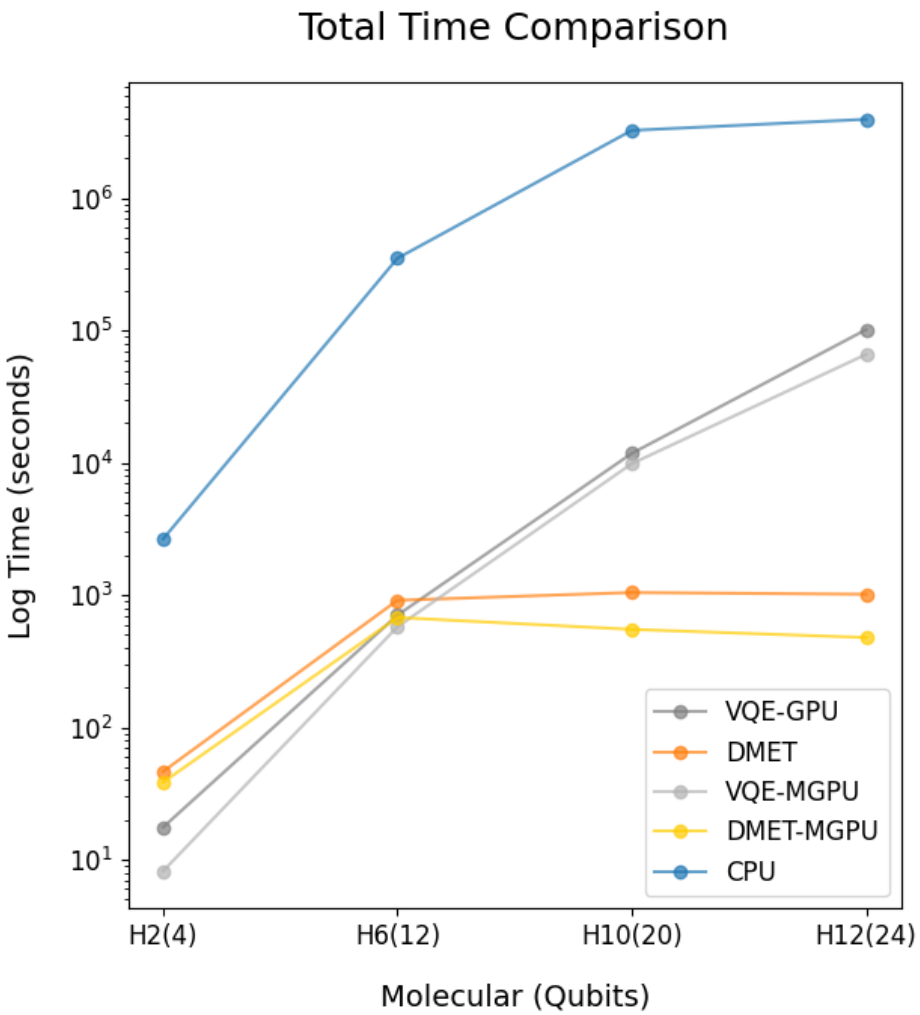
Time Comparison

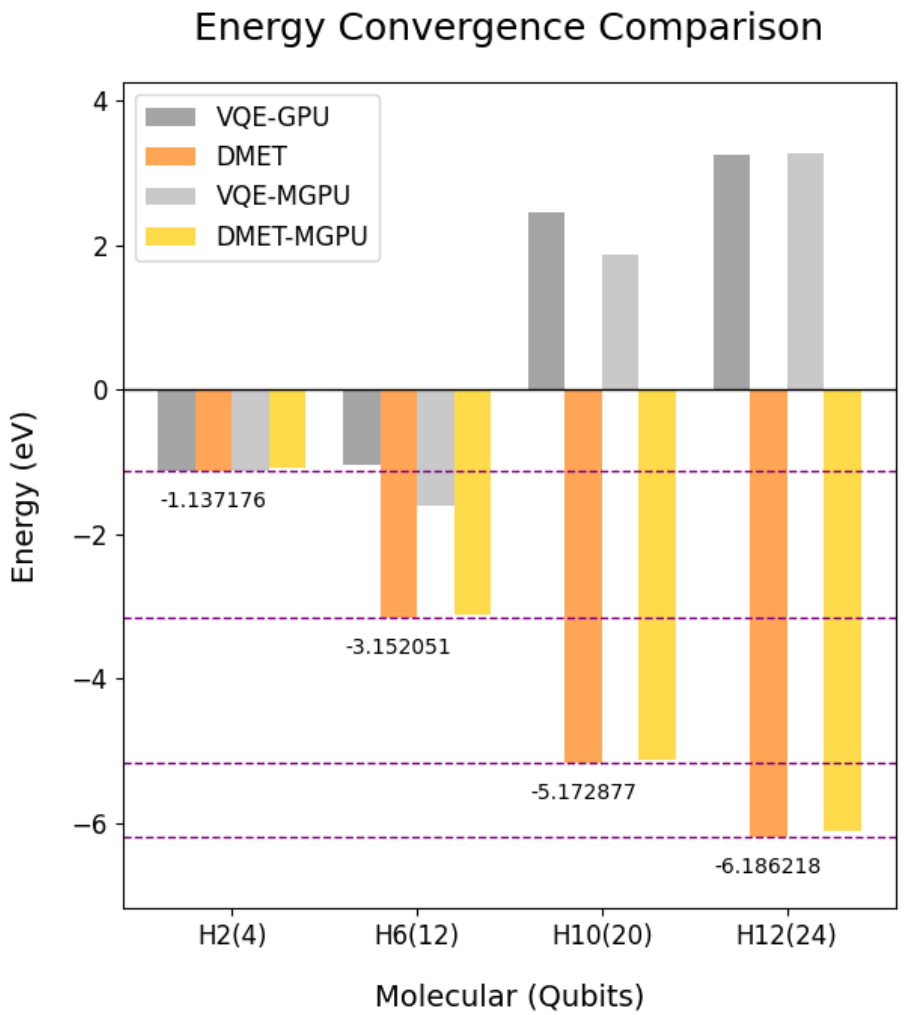


	VQE-GPU	DMET	VQE-MGPU	DMET-MGPU
H_2 (4)	151X	57X	324X	69X
H_6 (12)	499X	385X	611X	520X
H_{10} (20)	277X	3118X	330X	5925X
H_{12} (24)	39X	3898X	60X	**8282X**

The acceleration factor is estimated, the actual speed up will be greater

Time Comparison





Power Efficiency

	AMD Dual Rome 7742	3x A100 40GB	Power Savings
Compute Power (W)	189,796	6,500	183,296
Networking Power (W)	8,012	93	7,919
Total Power (W)	197,808	6,593	191,215
Compute Power (W)	189,796	6,500	183,296

1188

Metric tons of CO_2

256

Gasoline cars driven for a year

19632

Trees growing for 10 years

DMET Benifit

- For large molecules, it can **accelerate convergence** to **accurate energy**.
- It enables rapid computation for a number of **qubits exceeding the GPU limit**.
- The computation time is lower, and the convergence effect is better.

Additional Highlights

Packages

Cuda-Quantum

OpenQemist **(Our Modified Version)**

Qsharp

Qiskit

Contribution to the community

- Bridged the CUDA-Q kernel to the OpenQEMIST package (DMET).
- Rewrote code to make it compatible with modern versions of packages.

Problems encountered

1. Lack of rigorous comparison units due to constraints in computational resources, memory, and time.
2. Can’t split one GPU into MQPU for some molecules. It would raise a memory error.

Contribution

- Enhanced scalability: DMET-VQE surpasses the 32-orbital limit, simulating larger systems based on GPU resources.
- CUDAQ-based DMET-VQE: Facilitated systematic molecular comparisons, improving calculation accuracy and flexibility.
- Multi-GPU acceleration: Leveraged MGPU architectures for scalable DMET-VQE in high-dimensional systems.
- We shortened a **46-day task to just 17 minutes.**

Future Work

- DMET divides based on **orbitals** and the number of **qubits,** rather than by atomic species.
- DMET computations also utilize GPU processing.
- Experiment with **different partitioning methods**.
- **Multi Threading** to make calculation more efficient