

Chan-Yu Mou, En-Ming Huang, Pin-Yu Kuo, Fu-Chiang Chang, Pang-Ning Wu, Hsu-Tzu Ting, Jerry Chou

Department of Computer Science, National Tsing Hua University, HsinChu, Taiwan

## ARCHITECTURE

### CLUSTER OVERVIEW

#### Compute Node x2:

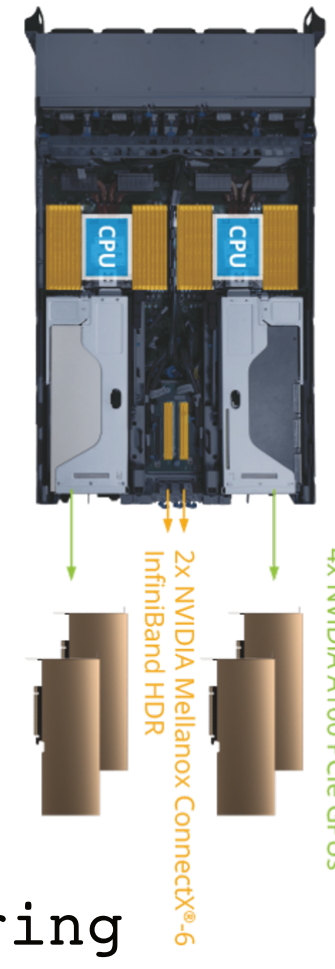
- High performance CPU
- Equipped with computing card

#### InfiniBand NIC:

- 200Gbps bandwidth HDR IB
- maximize IO performance

#### Gigabit Ethernet:

- 25Gbps bandwidth with MT4119
- use for management and monitoring



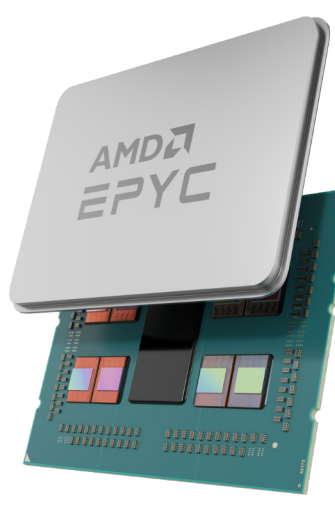
### HARDWARE

#### CPU: AMD EPYC™ 7003X

- 64 Cores per processor
- 2 CPUs on one node

#### GPU: Tesla A100

- 4 GPUs per Node
- 2TB/s memory bandwidth
- 300 Watts power consumption



#### NVLink Bridges:

- support communication between 4x A100
- each pair provide 600GB/s communication

#### RAM DDR4 3200 MHz 16GB x32:

- 8 DDR channels for full bandwidth

#### SSD Intel 750 SSD 400GB x2:

- RAID 0 for best performance

### SOFTWARE

#### OS:

- Ubuntu 20.04.3 LTS

#### Package Manager: Spack

- offer good environment for us
- save time on compilations
- provide more dependable options

#### Compiler:

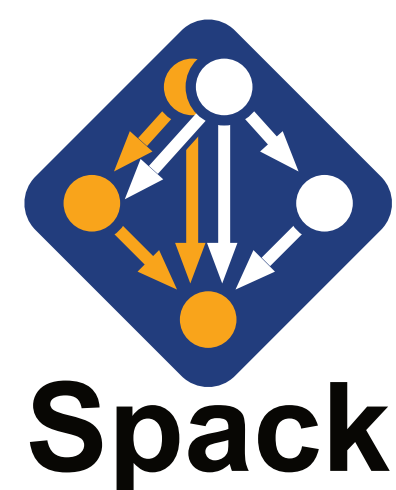
- GNU Compiler Collection 12.1.0
- GNU Compiler Collection 4.8.5
- Intel oneAPI 2022.1
- AOCC 3.2.0 / AOCL 3.1.0

#### GPU Software:

- CUDA 11.3, 10.2, 10.0
- HPC-SDK 21.5, 20.11

#### MPI:

- OpenMPI 4.1.0, 3.1.0
- Intel oneAPI-2022.1



### CLOUD

#### Configuration:

- set environment with our scripts
- create environments with Singularity

#### Installation:

- install software with Spack
- write Spack patches for custom installation

## OPTIMIZATION

### GENERAL METHODOLOGY

#### Profiling:

- Intel® VTune Amplifier, NVIDIA® Profiling Tools
- identify the performance hotspot of an application
- allow us to analyze the time consumption

#### Code improving:

- vectorize with SIMD, offload to accelerators
- reduce communications and computing
- cache optimization, branch prediction, data alignment

#### Compiling: enable modern compiler optimization

#### Tuning: mpitune, CPU affinity, NUMA binding

### LINKPACK & HPCG & IO500

#### LINKPACK:

- fully understand and tune settings
- automatically optimizing and running

#### HPCG:

- support Tesla® A100 to accelerate computation

#### IO500:

- optimize communication and IO performance

### REPRODUCIBILITY CHALLENGE - DACE

#### Checkpoint and restart: resume with checkpointing

#### Port to AMD architecture: reproduce experiment on AMD

#### Optimize for AMD:

- employ AOCC and AOCL to boost performance
- enhance the portability of npbench

### PHASTA

#### CFD: have a scientific idea of how PHASTA works

#### Profiling:

- get hints for optimization and compare different build configurations

#### Porting to Math Libraries:

- utilize BLAS to offload heavy calculations

#### Vectorize: employ hardware features

#### Automated Scripts: run and resume from checkpoints

### LAMMPS

#### Paper survey:

- single node: employ GPUs
- multiple nodes: CPU only for better scalability

#### Experimental result:

- Hyper-threading should not be applied on LAMMPS, no matter AMD or Intel (Tested on Intel Xeon and AMD Threadripper)

#### Bottlenecks:

- memory bandwidth
- previous work: MCDRAM (Multi-Channel DRAM) reduces the runtime by 90%

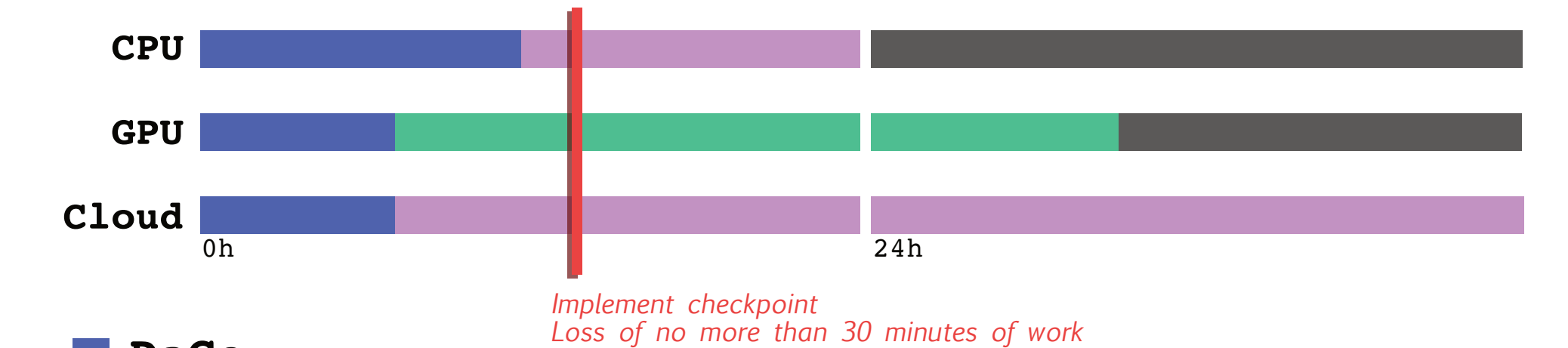
## PREPARATION AND TACTICS

### PREPARATION

In order to get familiar with the cluster environment and the SCC selected application, we meet our professor and coach every week to discuss in the problems we encountered.

In addition, additional HPC applications are selected to be installed and optimized for our server every few weeks. We are fully prepared and confident that we are ready for the fight!

### SCHEDULING IN 48H



#### DaCe:

- complete evaluation and start writing report
- use cloud to complete scalability experiments

#### LAMMPS:

- can be accelerated with GPU
- can overlap with PHASTA

#### PHASTA:

- allocate most of cloud cluster quota

#### Mystery:

- run after detailed research and analysis

### CONTROLLING

#### LDAP:

- LAM is a webfrontend for managing entries

#### NFS:

- allow machines to mount a disk partition on a remote machine

#### Dynamic Frequency Scaling:

- develop tools to scale CPU frequency on demand

#### AMD P-State Driver:

- finer grain frequency management with CPPC

#### CGroups resource management:

- prioritize critical applications

#### Power Shut-off:

- implement checkpoints on all applications

### MONITORING

#### Grafana:

- obtain realtime data from server
- make efficient decisions

#### IPMI:

- observe server resource consumption

#### Notifications:

- Push notifications to members' phones
- on power limit events
- on program is finished
- on program is aborted

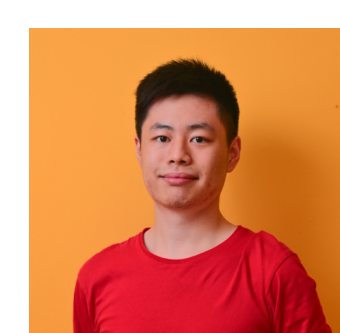
## DIVERSITY AND COLLABORATION

### TEAM DIVERSITY

Our team is comprised of five CS majors and one student from Tsing-Hua Experimental Education Program. All team members are first-time participants. Our team members are diverse in terms of education, gender, orientation, and motivation. One-third of the members are traditionally considered to be underrepresented in STEM fields.

Our goal is to create an SCC team that can represent all communities in Taiwan. We are proud to have these individuals from different backgrounds and provide their skills.

### COLLABORATION



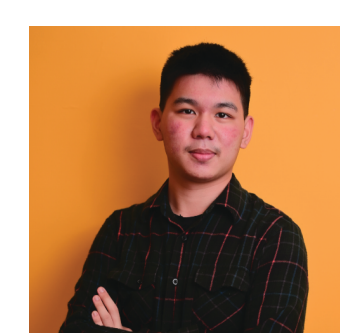
#### Chan-Yu Mou

- Major: Tsing-Hua Experimental Education Program
- Skillset: System management, Virtualization
- Working on: Leader, System Management, Poster



#### Fu-Chiang Chang

- Major: Computer Science
- Skillset: Distributed Training, Embedded System
- Working on: DaCe, PHASTA, Architecture



#### En-Ming Huang

- Major: Computer Science
- Skillset: Parallel Computing, CUDA Programming
- Working on: LAMMPS, DaCe, Presentation



#### Pang-Ning Wu

- Major: Computer Science
- Skillset: Algorithm, Quantitative Trading
- Working on: LAMMPS, Survey Research



#### Pin-Yi Kuo

- Major: Computer Science
- Skillset: Cybersecurity, Networking, Profiling
- Working on: PHASTA, Components of Cloud



#### Hsu-Tzu Ting

- Major: Computer Science
- Skillset: Data Analysis, Machine Learning, Linux
- Working on: PHASTA, Survey Research

## SUPPORTER

### OUR SCHOOL

National Tsing Hua University is a comprehensive research university offering a full range of degree program in science, technology, engineering, humanities, social science, and management. It consistently ranks as one of the premier universities in East Asia. <https://www.nthu.edu.tw/>

### Our Sponsor



Quanta Cloud Technology  
<https://www.qct.io/>



National Applied Research Laboratories  
National Applied Research Labs  
<https://www.narlabs.org.tw/>