# Team Peking University

*Yun Liang (advisor), Zhenxin Fu (advisor), Pengcheng Xu, Yueyang Pan, Yucheng Gu, Zhewen Hao, Fangyuan Yang, Zejia Fan*
*{ericlyun, fuzhenxin, jsteward, pyyjason, llgyc, haozhewen, yangfangyuan, zejia}@pku.edu.cn*

## About Peking University

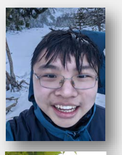
Peking University (PKU) is a major research university in China. Founded in 1898, Peking University has always been the pioneer of novation and improvement, playing a significant role "at the center of intellectual movement" in China. Peking University ranks as one of top academic institutions in China, Asia and worldwide. Here, top Chinese students seek opportunities for academic excellence and dedication to the society.

### HPC at Peking University

- Peking University's high performance computing university-level public platform was officially launched In 2018.
- Supercomputers "Weiming no. 1", "Weiming Teaching no.1", "Weiming Teaching no.2" and "Weiming Life Science no. 1" are in operation.
- The total number of computing cores on the public platform reached 11620, the peak computing capacity 999TFLOPS, and the storage capacity 9649.6TB.


## Team Members

We have seniors and juniors in our team, all obsessed with supercomputing.



**XU Pengcheng** *Team leader*
Senior, School of EECS
**Major:** Turing Class
**Skills:** System Building
**Focus:** HPL, HPCG, IO500 & cluster management

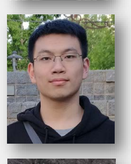
**YANG Fangyuan**
Senior, School of EECS
**Major:** Computer Science
**Skills:** EDA, Algorithms
**Focus:** Gromacs

**HAO Zhewen**
Junior, School of EECS
**Major:** Computer Science
**Skills:** Algorithms, Parallel Programming
**Focus:** Reproducibility task

**Prof. LIANG Yun** *Advisor*
Research Interest: High Performance Computing, Compilation and Architecture.
Publication in Top-tier conferences: HPCA, ISCA, MICRO, FPGA, DAC, etc.

**PAN Yueyang**
Junior, School of EECS
**Major:** Turing Class
**Skills:** SysOps, Networking
**Focus:** HPL, HPCG, IO500

**FAN Zejia**
Senior, School of EECS
**Major:** Computer Science
**Skills:** Pytorch, C++, python
**Focus:** Reproducibility task



**GU Yuchen**
Junior, School of EECS
**Major:** Turing Class
**Skills:** ACM/ICPC, Algorithms
**Focus:** CESM

**FU Zhenxin** *Advisor*
Research Interest: Natural Language Processing.
Publication in Top-tier conferences: AAAI, IJCAI, CIKM, EMNLP, etc.

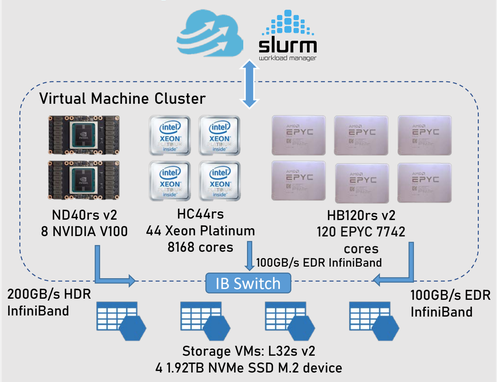
### Team Diversity

- Diversity of genders. We recruited members of different genders for our team and actively encourage both females and males to develop interest in computer systems, HPC and scientific computing by offering a short-term learning experience. We also provide opportunities to consult our advisors about career development.
- Diversity of birthplaces. The birthplaces of our team members cover half of Chinese land
- Diversity of dialects. While standard Mandarin is the official dialect, we also have team members speak Shandong Mandarin, Shanxi dialect and Wu Chinese.

## Hardware and Software Configuration

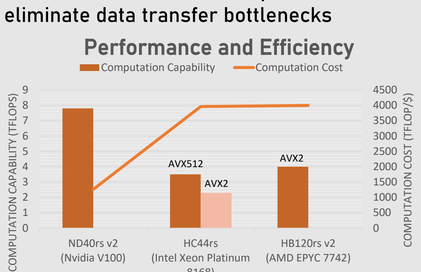
### Hardware Configuration

#### Cluster composition and architecture




#### Benefits and Reasons

- GPUs offer higher performance-cost efficiency, AMD-based CPU VMs have high cost efficiency. Intel CPUs are capable of dense computation.
- Sufficient performance for solving the problems while consuming budgets wisely with careful and detailed planning
- Maximized interconnection performance to eliminate data transfer bottlenecks



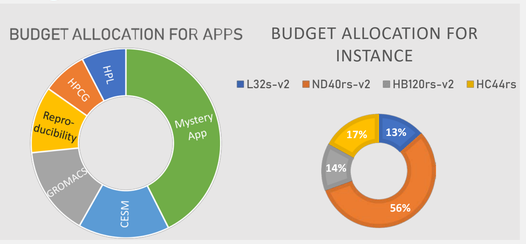

### Software Configuration

- CycleCloud CentOS 7 with EPEL repositories
  - Freely available Linux distribution that features enterprise-level stability. Wide adoption in the HPC community. Impressive hardware support for HPC-related software tools.
  - EPEL features extra packages commonly used by software development.
- NVIDIA CUDA 10.1 toolkit
  - A necessity to compile and run GPU programs.
- Mellanox OFED with UCX and OpenMPI 4.0
  - Official driver + user libraries for Mellanox EDR IB cards.
- AMD AOCC Compiler Suite
  - a high level of advanced optimizations, multi-threading and processor support.
  - highly optimized libraries.
- Intel Parallel Studio XE
  - Highly optimized code generation via ICC for Intel processors.
  - Fast math routines via MKL.

## Special Thanks to

## Resources Management

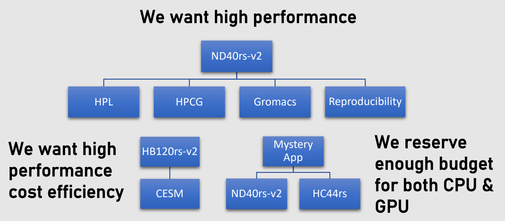
### Budget Plan




### System Administration Management

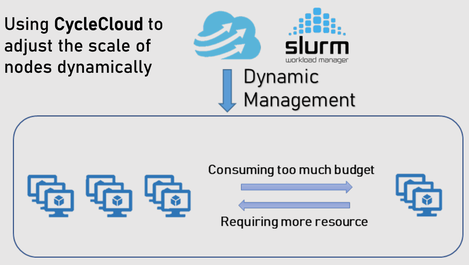
- Use Slurm for job scheduling and cloud resource management, coupled with a set of in-house scripts to monitor the nodes and cloud budget status and dynamically scale the cluster up or down.
- The storage nodes will stay online for the entire competition period, while worker nodes will be allocated on demand, providing great flexibility while alleviating idle cost.
- All the software and applications will be deployed beforehand to save time. Job scripts will be prepared ahead of time as well.

### Application Featured Management




### Dynamic Adjustment

Using **CycleCloud** to adjust the scale of nodes dynamically




## Our Preparation

### Learning Internally and Externally

- Take prerequisite classes for HPC match.
- Arrange regular technical tutorials to share and discuss where hosts also assign topic-related homework.
- Invite experts in academia and industry to deliver talks on HPC-related topics, free and open for everyone willing to promote HPC education inside the campus.

### Growth through Practice

- Homework assignments focusing on general strategies and techniques such as program compiling, profiling, Fortran, MPI, CUDA with deadlines.
- One of the team members will check the homework hand-ins and answering related questions.
- A GitHub organization is available for archiving code, discussion slides, and other files for open-sourcing and future reference.

### Responsibility Distribution

- Clarify everyone's responsibility. A technic leader is responsible for studying background information, profiling bottlenecks and proposing optimization methods for each application
- One student, as the system administrator, tunes the highest benchmark performance and low power consumption of the system hardware.
- Group meeting on VooV once a week to report progress and current obstacles in every application discussed.

### Support from our Vendors

- Provide training hardware and technical assistance:
  - Omnisky: access to HPC hardware for training as well as their expertise in the domain.
  - Peking University: university-sponsored access to online meeting platforms and regular training sessions in particular domains.
- Discussion on HPC software solutions with experts in industry.

## Optimizations and Strategies

### General Strategy

- Run profiling and find application hotspots and data transfer bottlenecks with ARM Forge, Intel VTune and NVIDIA Visual Profiler.
- Target-specific compiler options (-xHost, -march. –mtune and –arch) for optimal architecture code generation.
- Consider utilizing math kernel libraries, rewriting hotspots, hand-coding assembly, or porting the hotspot to acceleration devices.

### HPL / HPCG

- Gain experience from experiments.
- Tuning the parameters of HPL to change the problem size and get maximum performance according to the cluster configuration.
- Substituting out-of-date library with advanced ones and manually align MPI rank layout to solve performance bottleneck.

### IO500

- Try different possible combinations of parameters based on the configuration details of the cluster with a parallel version of IO500 to tune the system for maximum performance.

### CESM

- Run profiling for the application and find out performance-critical regions.
- Try different PE layouts to improve parallelism and choose the most efficient way.
- Use GPUs to offload some of the most time-consuming parts with heaviest computations.

### GROMACS

- Tried different CMake configurations, used GCC and MPI to build GROMACS and chose CUDA GPU configuring to accelerate.
- Profiled for all test components using NVIDIA Visual Profiler and found the hotspots.
- Created a shell script to fast deploy the application on the cloud clusters.

### Reproducibility Task

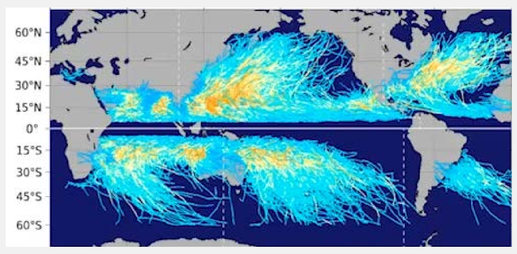
- Reproduce the results of the SC19 paper on datasets that grows exponentially in size and generate the output figures that show the scalability of the proposed system on several aspects with a one-click script.
- Study the code carefully to reproduce reliable results and experiment on tuning parameters to reach the best performance on the challenge datasets.
- Strictly schedule the usage of the nodes to ensure finishing a large amount of computation with the given budgets.

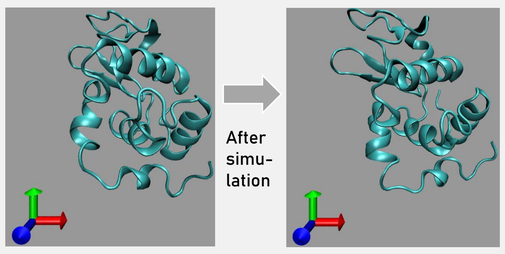
### Mystery Application

We've designed our cluster to cope with a wide range of problems:

- Sufficient CPU computing resources for conventional HPC problems
- Powerful GPU clusters for Big-Data / DL problems
- Enough high speed storage.

Besides hardware, we've also prepared various software toolkits for different possible situations.




## We will win!

- We have **passionate members** who cooperated together for the competition.
- The applications are well profiled and **tweaked for months**, so we have a deep understanding of different applications for the competition.
- Members are experienced in areas including **Parallel Computing**, **System Administration, Cluster Management** and **Computer Networking**.
- Well-designed **hardware and software configurations** are adopted.
- We have **extensive support** from vendors for training.