

About Peking University



Peking University (PKU) is a leading research university in China. Founded in 1898, it has been at the forefront of innovation and improvement and has played a significant role in shaping intellectual movements in the country. PKU is widely recognized as one of the top academic institutions in China, Asia, and worldwide, attracting top Chinese students who are dedicated to academic excellence and making a positive impact on society.

HPC at Peking University

- Peking University's high performance computing university-level public platform was officially launched In 2018.
- The total number of computing cores on the public platform reached 40436, the peak computing capacity is 5394 TFLOPS, and the storage capacity is 14783 TB.
- Peking University develops SCOW, an open-source supercomputer management portal and *CraneSched*, a supercomputer scheduler.



The Peking University Supercomputing Team

- The Peking University Supercomputing Team, established in 2015, achieved the first place in ASC 2023 and ASC 2024, and the second place in SCC 2020 and SCC 2023.
- The Peking University Supercomputing Team periodically holds events such as *HPCGame* and *HPC from Scratch* and arrange lectures to promote HPC as well as develop new members for our team.

1st HPCGAME

- 558 students from 87 universities made 13208 valid submissions.
- Both novices and experts in HPC can have fun and learn HPC knowledge in the game.



SCOW & CraneSched

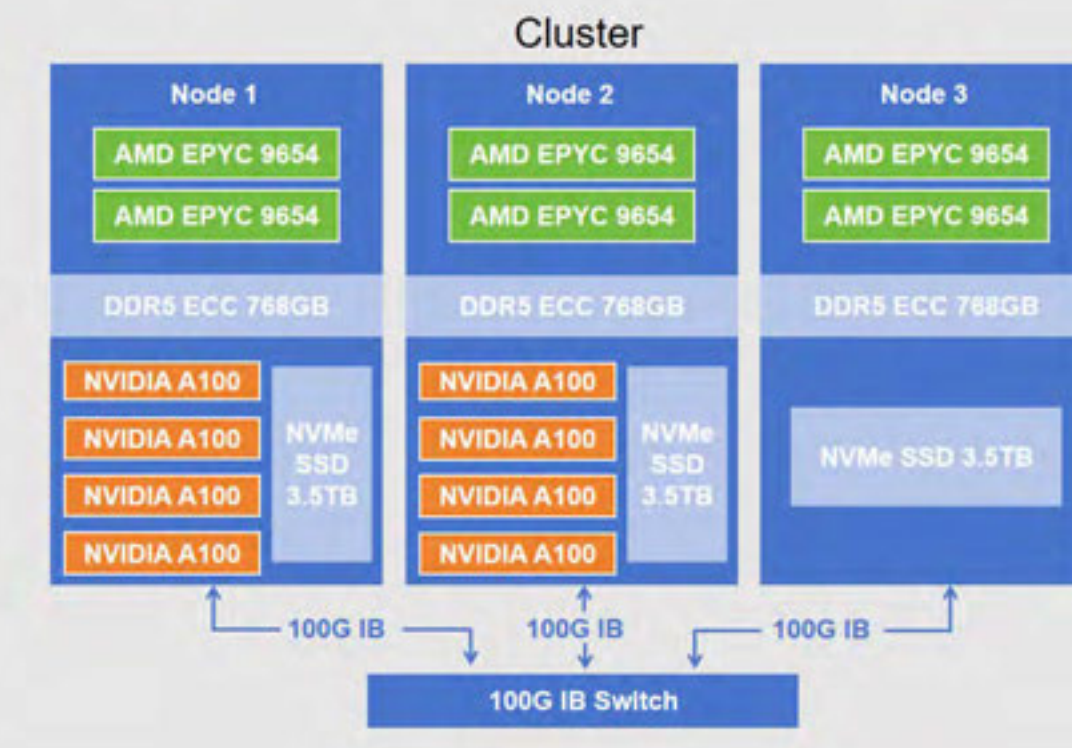
- SCOW**: a supercomputer portal and management system. Users can efficiently use computing resources simply through a modern web browser conveniently.
- CraneSched**: a high performance HPC and Cloud Computing fused Job Scheduling System.



Hardware and Software Configuration

Hardware Configuration

Cluster composition and architecture



Details and Reasons

- CPU**: Two 96-core AMD EPYC CPU provide performance bonus for highly paralleled applications while maintaining relatively low power consumption.
- GPU**: Four NVIDIA A100 GPUs per node offer high performance-per-watt efficiency, and provide an opportunity for GPU acceleration.
- Memory & Storage**: High-frequency DDR5 DRAM enables better bandwidth. High speed SSD provides high IOPS.
- Networking**: High-speed Infiniband network reduces data transfer overhead.

Software Configuration

- OS**: Rocky Linux v9 as OS (host and container)
 - Optimized performance, robust stability, and security.
 - Rocky's great compatibility ensures seamless integration of existing applications and tools.
- Compilers**: NVIDIA HPC-X, Intel oneAPI Toolkit, AMD AOCC
 - Generate highly-optimized programs.
- Scheduler**: CraneSched
 - Effectively manage hardware resources.



Strategies and Optimizations

Competition Strategy

Emergency Handling

- Sufficient personnel allocation for each application**. For each benchmark and application, at least three team members have a thorough and complete understanding of optimizations applied to it, and are capable of running it independently.
- Multiple runnable code versions**. For each application, we keep code and build scripts under different optimization options, allowing for quick switching in case of emergency.
- Always keep power and noise under control**. We have three team members familiarized with our power and noise management tools, and they will take shifts to be on site.

Time Saving

- Get scripts ready ahead of time**. We will have all possible scripts in advance to save time.
- Reasonable scheduling**. We have a comprehensive plan of scheduling team members to tasks. The reproduction will be prioritized to spill enough time for report writing. While running the other applications, idle members will be assigned with the mystery, enabling sufficient time to address the problems. We envision that, with our comprehensive preparation, we can achieve the maximum possible time efficiency.

Optimization

General Strategy

- Thoroughly understand the background knowledge and application scenarios.
- Use profilers to identify and optimize hotspots, porting to GPU if necessary.
- Couple applications with different compilers, compiler options, libraries, and runtimes.
- Optimize IO and MPI bottlenecks by analyzing results from micro benchmarks.
- Run applications on different hardware to determine the most suitable one.

HPL

- Use pre-compiled binaries from NVIDIA.
- Run HPL under different parameters thoroughly to gain experience about parameter tuning.
- Leverage InfiniBand, NVSHMEM, and GPUDirectRDMA to reduce communication overhead.
- Conduct on-site benchmarks to find the best parameters in the newly-configured cluster.

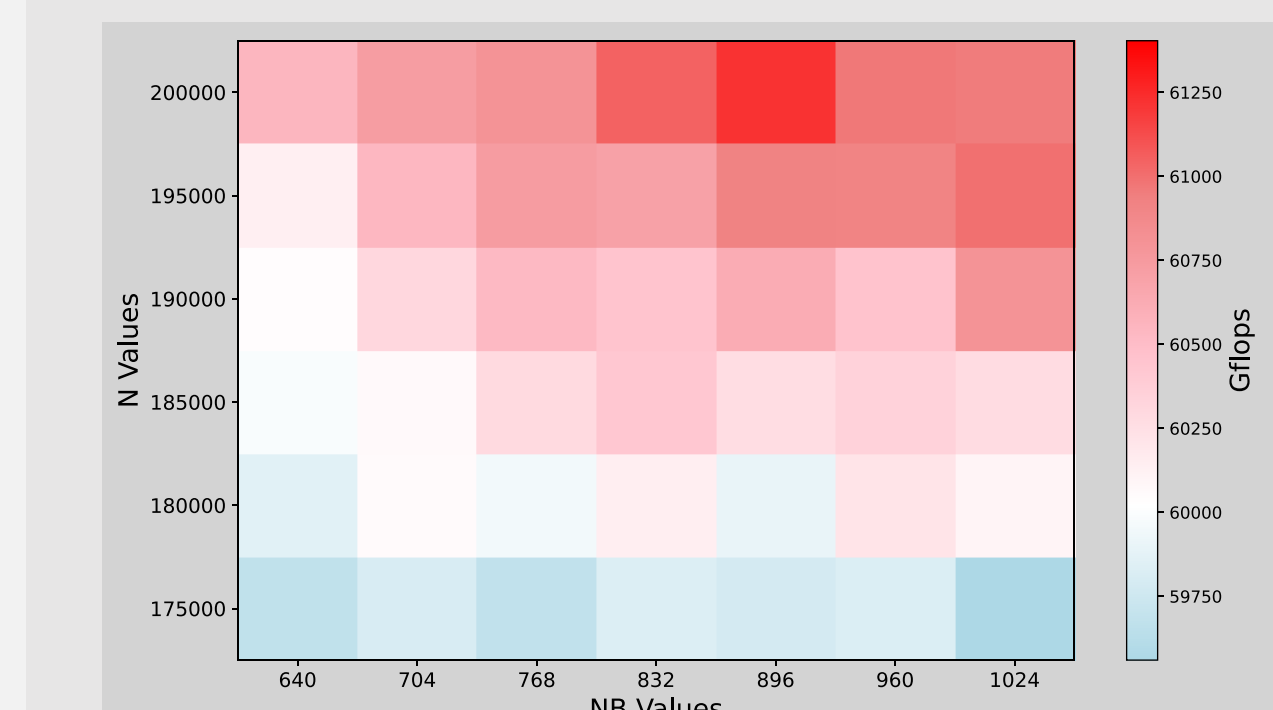


Image: The performance (GFlops) of HPL under different parameters (N and NB)

NAMD

- Use our custom toolkit to enable GPU resident mode on multiple nodes.
- Try different compiler, compiler options and hardware to find optimal configurations.
- Profile to find the hotspot.



Image: Profiler output of NAMD

MLPerf

- Employ int-8 quantization to accelerate inference speed while maintaining accuracy.
- Utilize NVIDIA's TensorRT as our graph and kernel level ML compiler and runtime.
- Use Ray to implement intra- and inter-node parallelization.

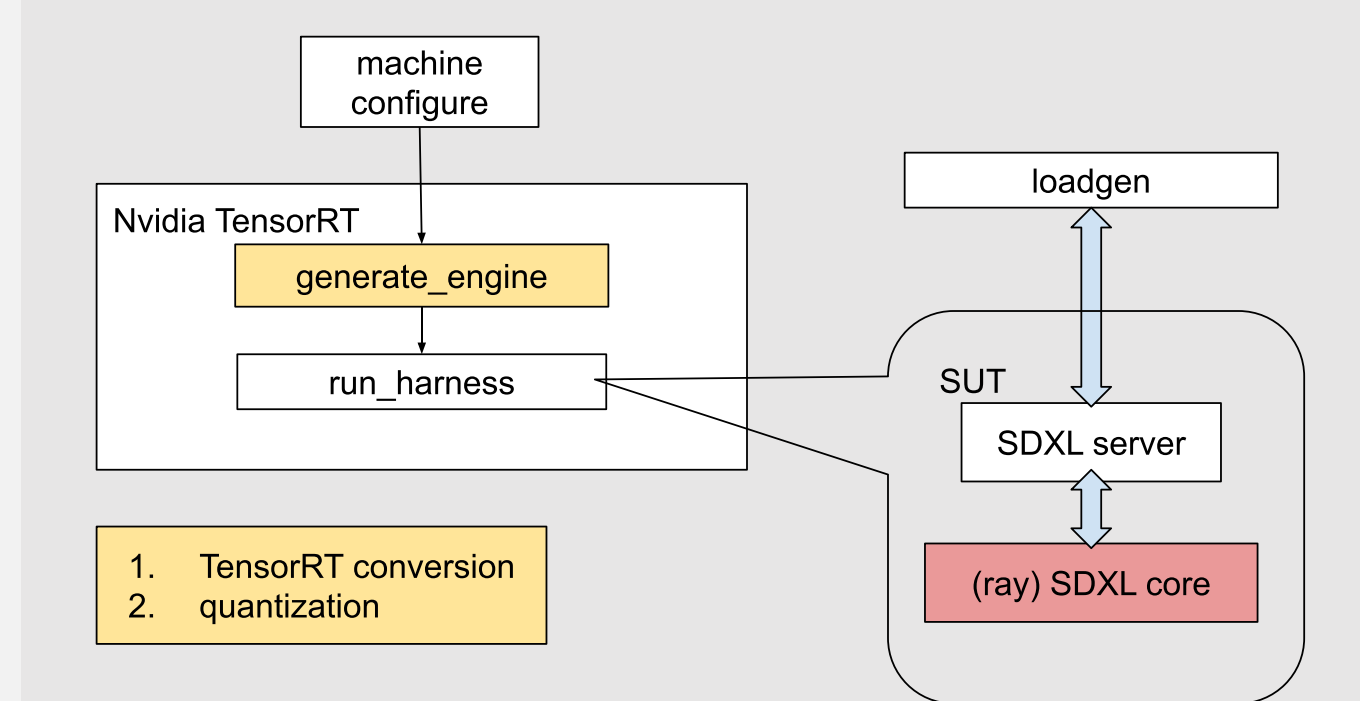


Image: Code architecture diagram of MLPerf

ICON

- Run full ESM configuration with MPI/OpenMP and/or GPU and compare performance.
- Measure MPI waiting time to improve load-balancing.
- Fix numerical issues and rounding error.

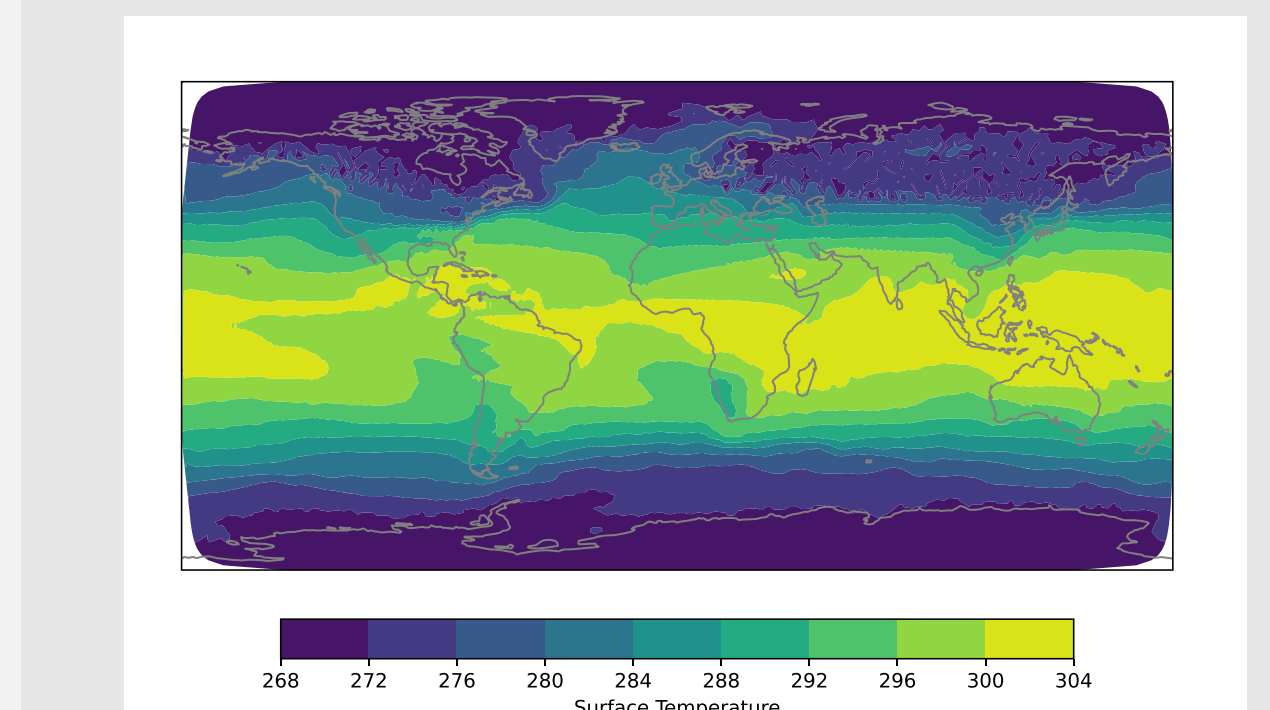


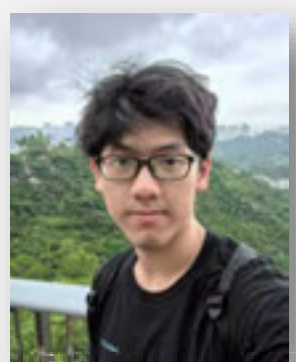
Image: Visualization of ICON's computation results

Reproducibility Task

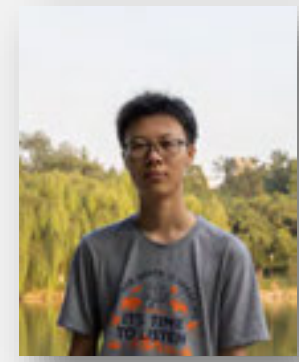
- In addition to the Ethernet setup provided in the paper, we tested performance improvement on high-performance InfiniBand network, adapting original scripts to fit our cluster setup.
- We applied *Datalife* to analyze other applications, effectively expanding the scope of the study.

Team Members

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



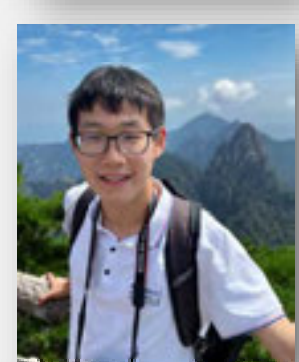
CAO Sinuo Senior, Team leader
Major: Information and Computation Science
Interests: Computer Networks, Parallel Programming
Focus: HPL, Cluster Management, Presentation



LIN Ruoyu Sophomore
Major: Life Science
Interests: Machine Learning Systems, Heterogeneous Computing, Computer Graphics
Focus: NAMD, MLPerf



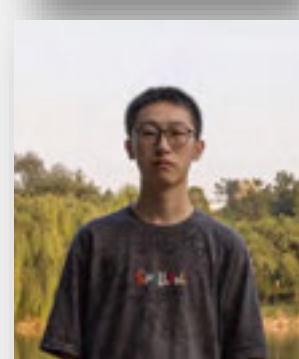
HUANG Di Junior
Major: Computer Science x Economics
Interests: Tuning, Quantitative Trading
Focus: System Tuning, Cluster Management



YANG Xinye Sophomore
Major: Computer Science
Skills: Algorithm design and implementation, Theory, Parallel Computing
Focus: ICON, NAMD



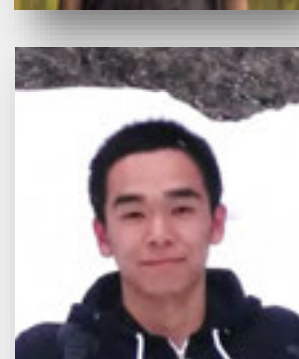
XU Lingyun Junior
Major: Artificial Intelligence
Skills: Deep Learning, System for Machine Learning, Recommendation Systems
Focus: MLPerf, Algorithm



ZHOU Yuxuan Sophomore
Major: Computer Science
Interests: Programming Language, Program Synthesis
Focus: Reproducibility Challenge, NAMD



FAN Chun Advisor
Chief Engineer, High Performance Computing Platform of Peking University
Research Interest: High Performance Computing, CFD.



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 individuals of all gender identities to develop interest in supercomputing.
- Diversity of majors and skills**. Our members come from various majors, including not only Computer Science students but also students from other majors that heavily utilize HPC. The members also bring together a diverse range of skills, covering various aspects of knowledge from hardware to software.
- Diversity of personality**. The personalities of our team members range from enthusiastic and talkative to calm and steady, allowing for great complementarity.
- Diversity of background**. We have members coming from provinces of different levels of development in China. Many members had little exposure to computer science before entering university. We provide ample learning resources and platforms for everyone, ensuring that everyone can overcome the barrier to studying high-performance computing.
- Diversity of grades**. We have students from different grades, including sophomore, junior and senior.

What Makes Our Team?

- Open discussion**. The team meet every Monday night face to face to discuss insights and issues.
- Clear responsibility**. Each team member clearly knows their responsibilities.
- Great team tools**. We use modern team management tools like Outline and Teams for convenient collaboration.
- Friendship**. We are not only colleagues in supercomputing teams, but also lifetime friends.



Power and Noise Management

Power Management

- Use our own self-developed power controller
 - Automatically adjusts fan speed, CPU and GPU frequencies according to predefined rule lists.
 - Provides estimated power consumption before changes are applied as a safeguard.
- Invent strategy-based semi-automatic tuning mechanism specifically designed for applications with bursts in power consumption such as HPL.



Image: The graphical interface of our power controller

Noise Management

- We have measured the noise level of our cluster and have a strong guarantee that it will never exceed the noise limit.
- Our power controller supports quick and easy adjustments of fan speed to manage noise levels.



Image: The power consumption curve of HPL