Southeast University

Supercomputing Club

University&Club

Southeast University

Southeast University is a world-renowned institutions of higher learning. It is a comprehensive and research-oriented university with engineering as its main feature. It has 13 disciplines ranked in the top 1% of the world in ESI, among which, its computer science ranks 9th in the world.

Supercomputing Club of SEU

In order to introduce HPC to more students, our university established the Supercomputing Club in 2022. The club has created an open-source supercomputing learning platform(CSWU-Challenge.github.io) for college students from scratch. Many students interested in supercomputing has joined the club, and are learning on the website. Additionally, the Big Data Center of SEU also provides HPC clusters for club members to prepare for supercomputing contests.

Team

ShunAn Zhu
Major: Software Engineering
Focus: Cloud Manage, IO500

Gongrui Zhang
Major: Artificial Intelligence
Focus: LAMMPS, PHASTA

Pengcheng Huang
Major: Cyber Security
Focus: Reproduction&Mystery

Xiaolong Ma
Major: Computer Science
Focus: Cloud Manage, LAMMPS

Zhongheng Guan
Major: Artificial Intelligence
Focus: PHASTA, HPL&HPCG

Linxiao Yu
Major: Cyber Security
Focus: Reproduction&Mystery

Team Diversity

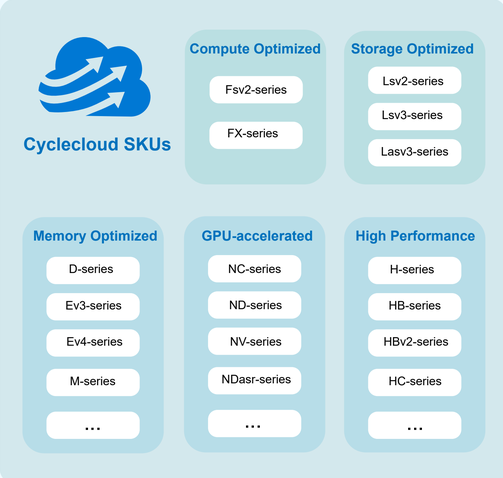
Gender Diversity. Our team consists of five males and an mtf (male to female). As the saying goes that 'Work is easier when men and women are working together', this diversity is undoubtedly a huge advantange for our team.

Birthplace Diversity. The birthplaces of team members cover Jiangsu Province, Guizhou Province, Yunnan Province and so on. People from different areas of China commonly have their unique skills and characteristics, which makes the team creative and active.

Major Diversity. Our members are from different college, covering most areas of computer science. Some of them are also experts in other fields, including Chemistry, Math, and Macromolecular Dynamics.

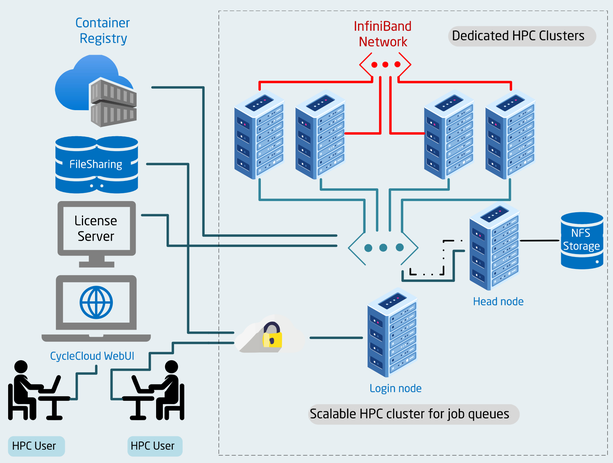
Cloud Configuration

SKUs choice



Usage	SKU	Description	Cost
LAMMPS	NC48ads_A100_v4	48*vCPU 440GB Mem 2*A100(80GB)	$205.69
PHASTA	ND40rs v2	40*vCPU 672GB Mem 8*V100(32GB)	$286.42
IO500	L80as v3	80*vCPU 640GB Mem 10*NVMeSSD(1.92TB)	$149.76
HPL&HPCG	ND96asr A100 v4	96*vCPU 1900GB Mem 8*A100(80GB)	$870.30
Storage	HB60rs	60 AMD EPYC 7551 processor cores 228GB Mem	$109.44

Management Strategy



Software Configuration

CycleCloud Ubuntu 20.04: A Linux distribution based on Debian and composed mostly of free and open-source software.

MadFS: A supercomputing burst buffer file system which has amazing IO performance.

NVIDIA HPC SDK: For maximizing the performance and portability of HPC applications in Nvidia GPUs.

Intel OneAPI: A tool kit includes compilers, MPI, Inspector and so on, optimized for Intel CPU.

NVIDIA CUDA 11.1 Toolkit: The most advanced platform necessity to compile and run Nvidia GPU programs.

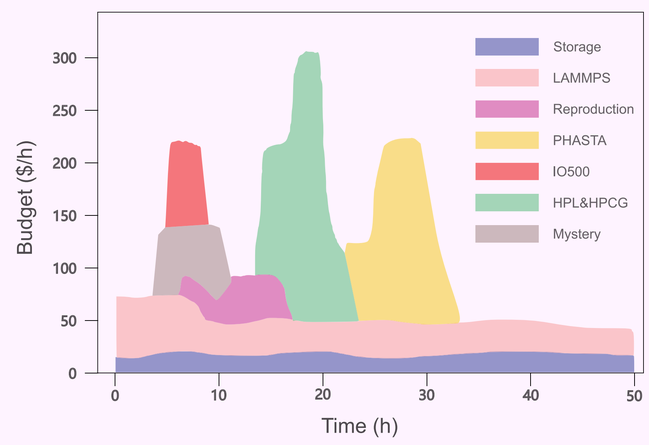
Spack: A package management tool designed to support multiple versions and configurations of software.

Slurm: A highly-scalable and fault-tolerant cluster management and job scheduling system.

CMake 3.24: An open-source, cross-platform family of tools designed to build, test and package software.

Budget Management

Budget Plan



- Node for management and storage will **fight until the end** to support other nodes to work well.
- LAMMPS will **last for the whole competition** to complete the computation.
- Make sure **enough resources** available to HPL&HPCG benchmark.
- Handle on the mystery application **first** to leave time for other applications and benchmarks.

Data Visualization

- Use Grafana to monitor the use of each VM
- Alert when exceeding budget
- Prevent VMs being idle
- Keep an eye on power

Optimization Strategy

Reproduction

- Comprehensively understand the improvement in 3Ps and the dataflow analysis in DaCe, especially SDFG IR-code mapping.
- Implement different Python frameworks(e.g. DaCe, Numba, CuPy) on several benchmarks like gemm, 3mm, 2mm for comparison.
- Evaluate the results on various architectures like CPU, GPU.
- Use automatic scripts for our experiments and analysis.

Mystery

- Deploy the suitable compiler and math libraries for best performance.
- Wisely balance the resource allocation for both GPU and CPU VM types.
- According to the experiment results, pick out the best SKU.
- Analyze the bottleneck(e.g. I/O performance, CPU affinity) of performance and optimize for better performance.

HPL&HPCG

- Use NVIDIA HPC-Benchmarks which is optimized for NVIDIA GPUs.
- Use Ampere NVIDIA GPUs(such as A100) to ensure the best performance.
- Tunning parameters to find the best option for specific architecture.
- Test different MPI implementations for optimal result.

IO500

- Use storage optimized SKUs with NVMe storage to maximize I/O performance.
- Use MadFS file system to support the test.
- Tune parameters according to cluster and file system configuration.

PHASTA

- Use Xeon Phi architecture with MIC technology to ensure best parallel performance.
- Try Intel compilers to better utilize vectorized instructions.
- Apply parallel computing tricks with OpenMP to tap the potential of Fortran and C codes.

LAMMPS

- Choose the best compilers according to the CPU and GPU architecture.
- Try different Accelerator Packages provided by LAMMPS to utilize GPU.
- Find a balance between the number of OpenMP threads and MPI workers for best performance.

Preparation

Regular Courses

- Researchers and engineers of the Big Data Center conduct training seminars and share their real-work experiences with all students on campus.
- Online courses such as "Introduction to High-Performance and Parallel Computing" and "Parallel and Distributed Computing"

Local Testing

- The SEU-HPC platform can give us powerful supercomputing clusters with 501 nodes for local testing.
- Our vendor will provide us financial support to do experiments on Azure machines.

Team Routines

- Regular brainstorms
- Maintain a team wiki to share HPC tricks.
- Read the latest papers and try to reproduce some crucial achievements to acquire up-to-date technologies.

Supports&Sponsors