SC21 Student Cluster Competition
Jinan University

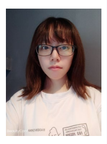
Introduction to the team

A. Team members

- ***Jiaming Li***, Team Leader, Grade-4 undergraduate student in Software Engineering, Winner of 2020 National Scholarship, Core member of ASC20-21 champion team (AI challenge). In VSCC, focus on reproducibility challenge and server deployment.

- ***Haolong Chen***, Grade-4 undergraduate student in Software Engineering, Interested in computer vision and supercomputer, focus on Quantum ESPRESSO and the optimization of HPCG in VSCC.

- ***Jing Hu***, Grade-3 undergraduate student in Information Security, Core member of ASC20-21 champion team (scientific computing), focus on Cardioid and the design of HPC cluster in VSCC.



- ***Zhiling Deng***, Grade-3 undergraduate student in Information Security, interested in algorithms, has participated in a series of collegiate programming contests, focus on Cardioid in VSCC.

- ***Junhong Cai***, Grade-2 undergraduate student in AI, GPA ranking #1 in his class, foucs on reproducibility challenge in VSCC.

- ***Renxing Chen***, Grade-2 undergraduate student in AI, focus on the optimization of HPL in VSCC.

B. Diversity of the team

The team is formed with Six Grade-2 to Grade-4 undergraduate students, and from Three different majors, including AI, Information Security and Software Engineering.

C. Team Advisor

Prof. Yang Guanghua

Full Professor, IET Fellow, Associate Dean of the School of Intelligent Systems Science and Engineering, Jinan University.

Adjunct Professor of the Shenzhen Institutes of Advanced Technology (SIAT) of the Chinese Academy of Science (CAS).

His research interests are in the general areas of IoT, big data and AI. In past more than ten years, Prof. Yang has worked across academia and industry. He has published 70+ papers/articles in top-referred journals and international conferences, led 20+ National and Provincial level R&D projects. Prior to returning academia, Prof. Yang was the founder of a startup company in the field of AIoT. He is now serving as the associate dean, coordinating the R&D of the school, and leading the major R&D activities in AI area.

About Jinan University

Jinan University (JNU), founded in 1906, is China's first overseas higher education institution, and is also a key comprehensive university listed in '211 Project'.

The school of Intelligent Systems and Engineering is the major organization in JNU specialized in HPC and AI. The school has a strong teaching and R&D team (leading by academicians of the Chinese Academy of Engineering), equipped with powerful HPC hardware and software infrastructure worth hundreds of millions dollars.

The school achieved considerable achievements in HPC and AI R&D in recent years. In the past five years, the school obtained more than 60 national and provincial-level projects, with total funding of 20M+ USD. 50+ patents and software copyrights have been granted, and 500+ high-level academic papers have been published.

JNU achieved remarkable achievements in world-class student supercomputing/cluster competitions, including Second Prize in ASC18, First Prize in ASC19, Champion of ASC20-21, 3rd Place Overall Winner of ISC21.

Software and Cloud Administration Approach

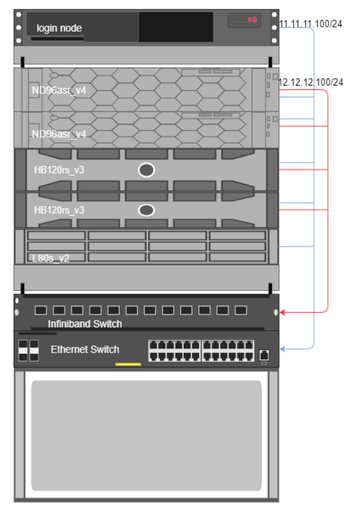
A. Cloud configuration and SKU's

- ***GPU nodes***: We analyzed and conducted preliminary test of the applications. Quantum ESPRESSO and Cardioid can run on GPU nodes. Meanwhile, LINPACK and HPCG Benchmark can get much better scores on GPU nodes.
- ***CPU nodes***: To better serve the Reproducibility Challenge and Mystery Application, powerful CPU nodes are also required.
- ***Storage***: An Optimized Storage Instance is selected to support the HPC applications and meet the challenge of IO500 Benchmark.

The table below shows the detailed information of our cloud configuration.

Node	CPU node	GPU node	Storage
Type	HB120rs_v3	ND96asr_v4	L80s_v2
Processor	AMD Epyc™ 7713	2nd-Gen. AMD Epyc™	AMD Epyc™ 7551
vCPU spec	120 cores 3.1 GHz	96 cores	80 cores 2.55GHz
Memory	448GB	900GB	640GB
Storage	1920 GB SSD	6000GB SSD	19.2TB NVMe
GPU Card	/	8*A100	/
RDMA Support	200Gb/s	8*200Gb/s	/
Cost/hour	3.96$	32.63$	7.488$
Number	2	2	1

B. Deployment and management strategy of cloud resources



- To enable the multiple nodes communicate with each other, all the nodes are connected to an ethernet switch.
- To enhance the performance, GPU/CPU Nodes are also connected to an IB switch.
- *Spack*, a multi-platform package manager, is selected for software installing and management.
- *Slurm Workload Manager* is selected for resource management.

C. Reasons for choosing this approach

The challenges we are facing in the competition are not limited to benchmarks but also to applications requiring powerful CPU as well as GPU computing power. Heterogeneous computing and storage resource need to be managed. Various kinds of applications also need to be managed and optimized effectively. In order to do so, *Spack* is selected to easy the software installing, configuration and management, while *Slurm* is employed to improve the resource utilization.

Competition Preparation and Application Optimization

A. Budgeting the cloud resources

- Budget estimation: $4000
- Cloud resources cost: $3872
  ((3.96*2+32.63*2+7.488)*48)

GPU and CPU nodes not fully utilized can be stopped and released to save budget.

More budget will be investigated to applications such as Quantum ESPRESSO and Cardioid for a better performance.

B. Strategies for running and optimizing the applications and preparation for the mystery application

- Quantum ESPRESSO and Cardioid: Compile the GPU version and run on our own cluster, then complete the running on multiple nodes. Meanwhile, make efforts to understand the flow of the applications and carry out optimization. Hotspots analysis is conducted with profiling tools such as Vtune. Further optimization can be done on time-consuming codes and data structures.
- LINPACK and HPCG Benchmark: Understand the configuration parameters through testing. Attempt to replace some of the libraries to achieve better results.
- Mystery application: Prepare a balanced hardware and software environment, make an optimization plan for both CPU and GPU applications.

C. Strengths of the chosen architecture and potential bottlenecks

- Strength: Experiment results demonstrated that under the budget limitation, the chosen architecture
  - Balance CPU and GPU computing power
  - Balance Computing Power and Storage Capacity/Performance.
- Potential bottlenecks: Due to the uncertainty of the working datasets, there is the possibility of insufficient or excess resources in the competition.