YOUTH

Cluster of stars

A group of skilled students helps Tsinghua University to retain its crown at major supercomputing competition despite the challenging change in format, **Li Yingxue** reports.

It was 6 am, and a team of six students from Beijing's Tsinghua University finally breathed a sigh of relief after an arduous 72 hours of competition.

While it might not have been physically challenging in the traditional sense, the nonstop testing, calculating and computer work can still take a huge toll.

Fortunately, it was all worth the sore necks, aching wrists and depleted mental resources as they were announced the winner of the 2020 Student Cluster Competition.

It was the third consecutive year the university had won in the same international competition.

In 2007, the US-based International Conference for High Performance Computing, Networking, Storage and Analysis launched the Student Cluster Competition to provide an immersive, high-performance computing experience to undergraduate and high school students around the world.

Due to the COVID-19 pandemic, the 2020 competition took place "in the cloud" in November.

The online competition attracted 19 teams of 114 students from different countries, including the Massachusetts Institute of Technology in the United States, the University of Warsaw in Poland, Friedrich-Alexander University in Germany, Singapore's Nanyang Technological University alongside China's Peking University and Tsinghua, among others.

Participants were required to design and build virtual "clusters" in the cloud and learn scientific applications, apply optimization techniques for their chosen cloud configurations to complete a set of benchmark tests and real-world scientific workloads against the clock.

A "cluster" is a collection of high-performance computing devices, linked together to utilize their combined resources to complete complex and high-volume data-based tasks quickly and accurately.
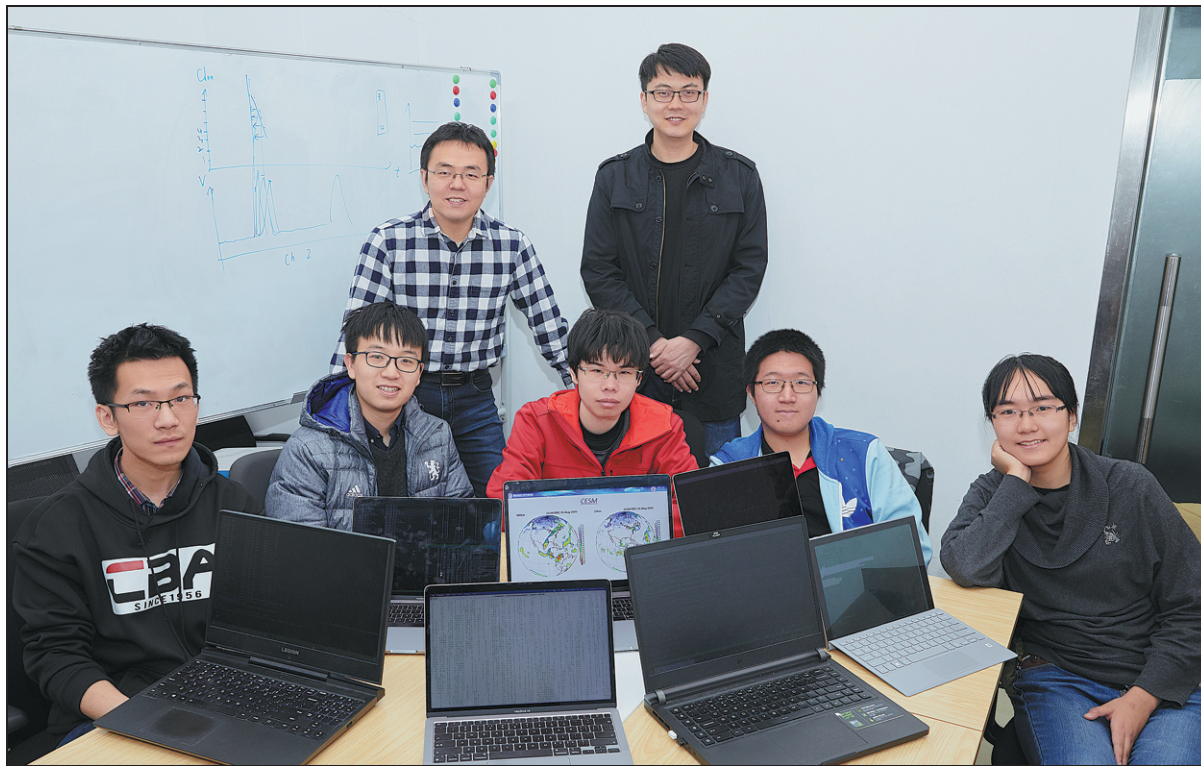
According to Han Wentao, the Tsinghua University team's tutor, the previous on-site competitions where all teams gathered together, bringing their own hardware, lasted 48 hours.

"Each team needed to assemble their own cluster on-site with a maximum power of 3,000 watts — it's like Formula 1, where you need to spend a lot of money on buying the best engines and parts," Han explains.

For the 2020 iteration of the contest, however, the rules were different — all preparations and computing were done in the cloud. Participating teams had a $3,200 initial budget to buy hardware and services online. Unbeknown to the competitors, however, 12 hours before the end of the competition, they would receive a windfall of $500 additional budget.

"This time, all teams were on the same starting line with a limited budget," Han says.

Another challenge in the virtual

Members of the Tsinghua University team pose with their tutor, Han Wentao (standing, left), as they prepare for the 2020 Student Cluster Competition. The university won the competition for the third consecutive year in November. LI PAI / FOR CHINA DAILY

> Each time I compete, I gain a little more knowledge about different applications and techniques, so it's a continuous process of learning."
>
> **Zhang Chen,** leader of the Tsinghua University team that participated in the 2020 Student Cluster Competition
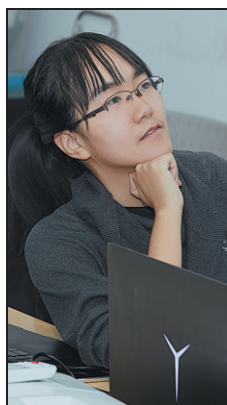
competition was the time difference. The competition started at 6 am in Beijing and the judges, who would be answering questions during the competition, were only online between 10 pm and 4 am.

Taking up tasks

The teams had to complete a number of tasks, including Linpack — or linear software package — benchmarks.

Linpack is basically using the computer to solve a linear equation with one million unknowns, Han says. Because many practical problems can be converted into this equation, so it is fundamental function that needs to be tested.

Han explains that the teams are scored on the speed at which they complete the task.

Another challenge was the Gromacs molecular dynamics application — something that is relevant to the COVID-19 pandemic. Molecular dynamics simulations offer accurate approximations of real molecular behavior, which is useful in different stages of drug development.

Song Zeyu was a member of the Tsinghua University team, who together with another teammate, was tasked with running that application.

The 20-year-old spent the first day learning about the topic and started to run the application on the second day of the competition. He and his teammate took turns to sleep so that one of them was always monitoring the application.

"When the competition was held

on-site, we got the chance to talk with the other teams, which was a great experience," recalls Zhang Chen, who joined Tsinghua's supercomputing team in 2018.

For the 2020 competition, Zhang was the team leader.

"Obviously, I have much more responsibility. I need to keep an eye on the progress of all the team members during the competition and make final decisions and some tough choices, including how to assign our resources and budget," says the 21-year-old.

During the competition, she would regularly gather the team together to discuss budgetary constraints and how to best distribute the resources.

"The budget is definitely not enough to finish running all of the application challenges, so we needed to decide how to assign the money," she says.

Zhang recalls on the second night, she needed to make the first big decision of the contest when they realized that a majority of their budget would have to be spent on just one of the challenges.

"On the third night, when we needed to decide how to use the extra $500 suddenly given to us, we shut down all devices to discuss it, which was a tense moment for us," Zhang recalls.

Zhang didn't get much sleep during the 72-hour competition — as the event wore on, her nights became more restless.

It was not in vain, though, as the team didn't only end up winning the

whole thing, but also achieved the highest score in Linpack benchmark.

Ability to calculate

Han notes that what supercomputing is trying to achieve is the highest speed of computational power and adds that it ideally suited to any subject that requires a massive amount of calculating ability.

Take weather forecasts as an example. To calculate future weather patterns, the map is separated into grid boxes and there are various data points within them that are taken into consideration, including temperature, humidity, wind direction and wind power.

There are two ways of making the forecast more accurate — one way is to make the grid boxes smaller and the other is to calculate with greater levels of iteration — that's when supercomputing comes to the fore — when massive amounts of calculation are required in an increasingly limited time.

Han began tutoring Tsinghua University's supercomputing team in 2016. He explains that, each year after the senior students graduate, they recruit new members.

Han says all team members are from the department of computer science and technology, and most of them have taken part in computer competitions during high school.

"There are usually five to six students in a team, and each of them has his or her own responsibilities — one or two members run each application, and in these 'sub-teams', as

well as running the tests, one member will compile a presentation and the other will write a report. So it's a multitasking job for each member," he explains.

"It's a competition that tests a comprehensive number of abilities."

After each competition, whether they win or lose, the team members will review their processes and discuss which areas can be improved upon.

"We are aiming to be the champion of each competition we enter, but we are not result-oriented. We hope our students can actually learn something from the competition, and learn what supercomputing is," Han says.

The competition is a good practice for the students, whether for scientific research or solving practical problems in the future, he adds.

"The competition is a microcosm of a modern high-performance computing center, which I hope inspires the students to pursue a career in supercomputing."

Song, a junior student, joined the team last year. He says before the competition he was unclear about how cluster computing actually worked. By preparing for the competition, he was able to gain access to advanced equipment and learn about how certain applications are run.

"It's a field that is doing something practical," he says.

Han says the development of supercomputing has come a long way — it started with the invention of the computer in 1946.

Supercomputing is relatively unpopular among science students, many of whom will choose other fields, such as artificial intelligence which is closer to people's daily lives.

Han is glad to see that through the competition, some students are drawn to pursuing their master's study, and even a career, in cluster computing.

Zhang says to prepare for each competition she needs to embark on additional study of related subjects to the applications she will be running. It allows her an insight into varied array of other fields.

"Each time I compete, I gain a little more knowledge about different applications and techniques, so it's a continuous process of learning," she says.

"The competition also points to the direction I want to take in my future research."

Expecting to finish her major with the highest distinction, and after scooping a 15,000 yuan ($2,300) scholarship prize from Tsinghua for her exemplary performance as a bachelor student, it is perhaps unsurprising that, upon graduating in June, Zhang is set to pursue a master's degree in supercomputing.

It does not take a cluster computing system to calculate that Han will be pleased to hear that.

Contact the writer at liyingxue@chinadaily.com.cn