ASIANSCIENTIST

Issue 09
January 2021

SUPERCOMPUTING ASIA

PEAK PERFORMANCE

HOW JAPAN'S FUGAKU BALANCED POWER AND EFFICIENCY FOR THE BEST OF BOTH WORLDS

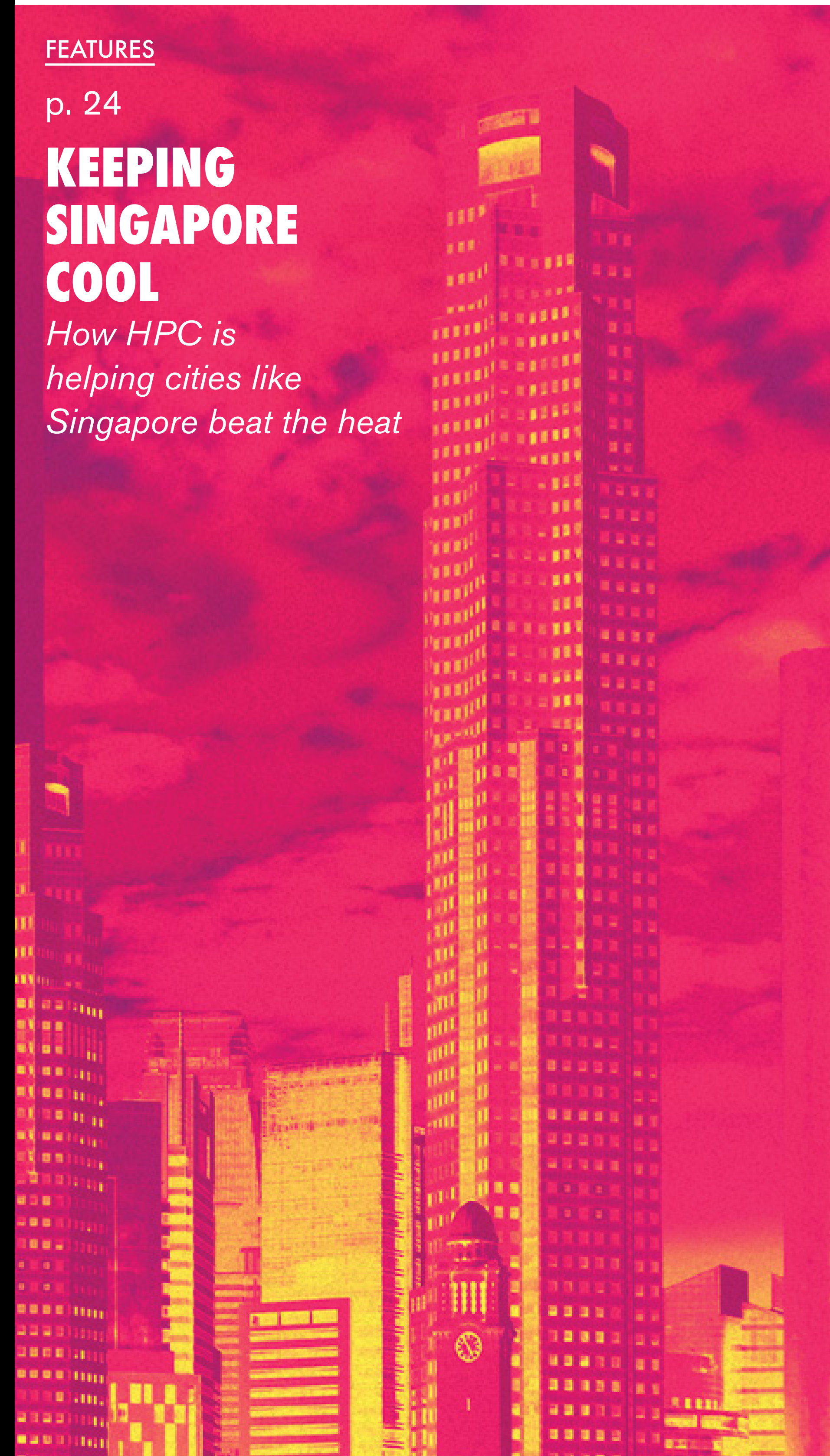
\+

KEEPING SINGAPORE COOL

CRUNCHING THE NUMBERS OF CLIMATE CHANGE

CONTENTS

Issue 09
January 2021



EDITOR'S NOTE

Fugaku may be the fastest supercomputer in the world but it was never designed simply to climb the rankings. Instead, the thousands of researchers working on it had a singular vision: the lofty Mount Fuji, a pinnacle of performance that would serve a broad base of users. In our cover story (*Peak Performance: Japan's Fugaku Supercomputer*, p. 16), we speak to Fugaku's chief architect, Professor Satoshi Matsuoka, who described it as being as easy to use as a Prius while nonetheless giving users the performance of a supercar like a Porsche.

Although COVID-related applications were understandably one of the first uses of the powerful machine, Fugaku's greatest impact may well be on the imminent threat of climate change. From forecasting the weather to simulating the entire Earth, scientists are using supercomputers to understand the impact of human activities on our planet—and what can be done to mitigate the damage (*Crunching the Numbers of Climate Change*, p. 10).

This is not without a touch of irony, since supercomputers themselves consume large amounts of electricity, with the most powerful ones requiring the equivalent a small city's supply to run. The power limit is a serious challenge designers of next generation exascale machines will have to confront, but one that is now being taken very seriously (*Striking a Balance Between Performance and Efficiency*, p. 38).

Rebecca Tan, Ph.D.
Editor-in-chief
Supercomputing Asia

/asianscientist · @asianscientist
Asian Scientist Magazine · asianscientist
AsianScientist · asianscientist

www.asianscientist.com

SUPERCOMPUTING ASIA

EDITORIAL ADVISORY COMMITTEE
Prof. Tan Tin Wee
Prof. Satoshi Matsuoka
Prof. John Gustafson
Yves Poppe

CEO & PUBLISHER
Dr. Juliana Chan

EDITOR-IN-CHIEF
Dr. Rebecca Tan

STAFF WRITER
Kamila Navarro
Sheryl Lee
Jill Arul

CONTRIBUTORS
Eugene Low
Tim Hornyak

ART DIRECTOR
Lam Oi Keat

SENIOR DESIGNER
Shelly Liew

SALES & MARKETING
Samantha Yeap
Sammi Ting
Jillene Kee

PUBLISHED BY
Wildtype Media Group Pte Ltd

DISTRIBUTED BY
Pansing Distribution Pte Ltd

HEAD OFFICE
Wildtype Media Group Pte Ltd
71 Ayer Rajah Crescent #06-26
Singapore 139951
hello@wildtype.media
Website: www.asianscientist.com/subscribe

WILDTYPE MEDIA

Combining savvy communication with technical rigor, Wildtype Media Group is Asia's leading STEM-focused media company, spanning digital, print, video, custom publishing and events.

Brands under Wildtype Media Group include the flagship *Asian Scientist Magazine* and *Supercomputing Asia*, award-winning titles available in print and online.

www.wildtype.media

To order reprints or e-prints, or request permission to republish *Supercomputing Asia* content, please contact hello@wildtype.media

Digital Dispatch

REVEALING THE SECRETS OF REMDESIVIR

Researchers from the University of North Texas are using supercomputers to simulate how remdesivir and other antiviral drugs block SARS-CoV-2 reproduction, with the aim of improving drug efficacy.

Using the Stampede2 and Frontera supercomputers at the Texas Advanced Computing Center, a team led by Professor Andres Cisneros at the University of North Texas modeled the chemical reactions taking place at two key proteins involved in coronavirus replication, namely NSP-12 and the main protease. Remdesivir, the first drug approved by the US Food and Drug Administration for use against the coronavirus, works by binding to and inhibiting both proteins, thereby blocking replication.

With a peak power of 38.8 petaFLOPS, Frontera is the fastest academic supercomputer and the eighth fastest in the world. Nonetheless, the team needed to save computational time to efficiently simulate their molecular systems made of hundreds of thousands of atoms. To do so, the researchers adopted a hybrid approach to better focus on interactions at the active site, with a more approximate molecular dynamics approach for the rest of the reactions. "Frontera, with not only compute power but the intercommunication between the nodes, allows us to run these quantum mechanics and molecular mechanics calculations with much higher speed and throughput," said Cisneros.

GOOGLE'S QUANTUM CHEMISTRY SIMULATIONS DOUBLES PREVIOUS RECORD

In the largest chemical simulation ever performed on a quantum computer to date, scientists from Google's AI quantum team used the company's 54-qubit Sycamore processor and an in-house quantum algorithm to model bond energies in chemical compounds. Their calculation was twice as large as previous chemistry calculations on quantum computers.

Accurately modeling the behavior of chemical bonds not only improves chemical synthesis methods, but also aids in the design of novel materials. However, current computers lack the exponential scaling required for such work. Hence, chemists are turning to quantum computers like Sycamore. The team's algorithm treats Sycamore like a neural network, dynamically adjusting parameters to handle noisy quantum logic and deal with errors during computation. This process is similar to how classical neural networks use optimization to manage imperfections in data.

"We validate that algorithms being developed for currently available quantum computers can achieve the precision required for experimental predictions, revealing pathways towards realistic simulations of quantum chemical systems," wrote Dr. Nicholas Rubin and Dr. Charles Neill, two of the team's researchers. Google has released its code on OpenFermion, its open source library for compiling and analyzing quantum algorithms.

MODELING CLIMATE CHANGE WITH EARTH'S DIGITAL TWIN

Planet Earth is about to get a digital twin in the form of a climate model that can simulate the atmosphere, ocean, ice, land and even human activity with unequaled precision. Called Destination Earth, the model developed by scientists from the European Union will run on supercomputers and be able to forecast extreme weather events days to years in advance.

Destination Earth will detail the planet's atmosphere in one-kilometer boxes, a scale far finer than existing climate models. Thanks to this high resolution, the platform can use more detailed real-time data and directly render even small-scale processes, increasing prediction accuracy. Predictions include convection, where hot air rises to form clouds and storms, and swirling eddies in the ocean that move heat and carbon.

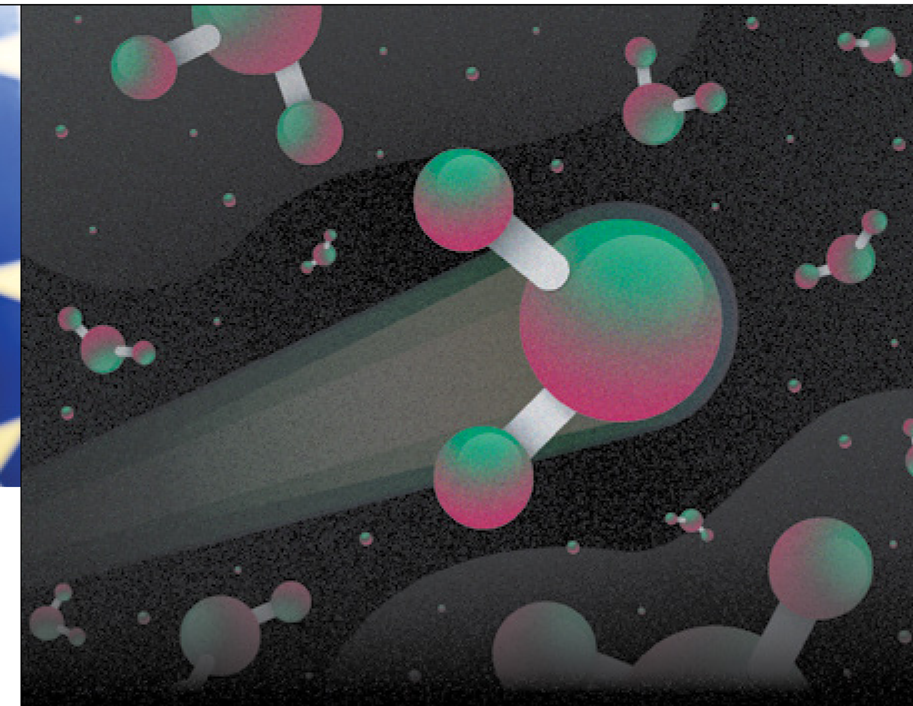
Beyond enhanced climate forecasting, Destination Earth is intended to help policymakers gauge the societal impacts of climate change. The model will be run on one of three pre-exascale supercomputers in the works under the EuroHPC Joint Undertaking: Finland's LUMI, Spain's MareNostrum 5 or Italy's Leonardo. The European Commission is planning to implement Destination Earth in 2021.

EUROPE BETS BIG ON SUPERCOMPUTING

The European Commission is planning to invest €8 billion (~US$9.47 billion) in its supercomputing industry over 13 years starting from 2021. The investment will be channeled through the EuroHPC Joint Undertaking, an entity that pools the resources of 32 European Union member states to develop supercomputing technologies and applications.

Currently in the works is a high-performance, low-power microprocessor codenamed 'Rhea.' Slated for release in 2021, Rhea is anticipated to play an important role in equipping European supercomputers with exascale capabilities. Apart from building exascale computers to reach supercomputing's next frontier, the funds will also be used to integrate quantum accelerators for hybrid machines.

Though the investment has yet to be formally approved by the Council of the European Union, the plan is set to develop European leadership in supercomputing. "Our proposal today will foster increased investment in supercomputing infrastructure, in view of its enormous potential to improve quality of life, boost industrial competitiveness, and advance science," said Ms. Margrethe Vestager, executive vice president of the European Commission's 'A Europe Fit for the Digital Age' initiative.



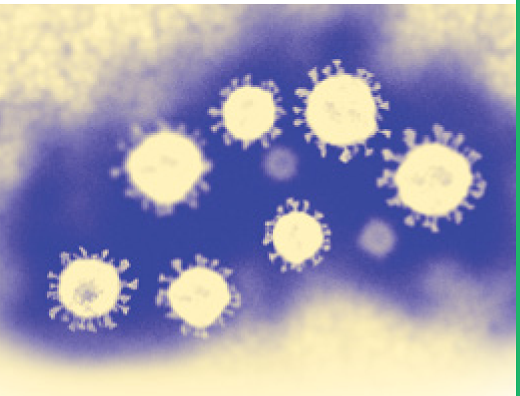
THE NOBEL PRIZE OF SUPERCOMPUTING GOES TO...

A team of nine researchers from Chinese and American institutions has won the 2020 Gordon Bell Prize for pushing the limits of *ab initio* molecular dynamics (AIMD) using machine learning. Compared with standard molecular dynamics, the AIMD approach allows for greater modeling precision, but is also more computationally intensive and has hence been conventionally limited to small-sized systems with a maximum of thousands of atoms.

Using a new machine learning-based protocol they developed, termed deep potential molecular dynamics (DPMD), the scientists were able to efficiently simulate a more than one nanosecond-long trajectory of over 100 million atoms per day. The team successfully scaled this proposed algorithm on US' Summit, the second fastest supercomputer in the world with a speed of 149 petaFLOPS.

The DPMD development opens the door to simulating unprecedented size and time scales with *ab initio* accuracy, the researchers said. "It also poses new challenges to the next-generation supercomputer for a better integration of machine learning and physical modeling," they wrote.

Digital Dispatch



VISUALIZING CORONAVIRUS SPREAD TO LIMIT IT

Japanese scientists modeling coronavirus contagion on the Fugaku supercomputer have found humidity to have a large impact on aerosol spread. The more humid the air, the lesser the amount of aerosolized particles leading to airborne transmission, according to the study by RIKEN and Kobe University.

Researchers used the world's most powerful supercomputer to model the emission and flow of virus-like particles from infected people in a variety of indoor environments. Simulations showed that air humidity that was lower than 30 percent resulted in more than twice the amount of aerosolized particles compared with humidity that was 60 percent or higher. The findings warn of heightened contagion risk in dry, indoor environments and suggest the use of humidifiers may limit infections during winter months.

The RIKEN research team led by Makoto Tsubokura has used Fugaku to model contagion in trains, classrooms and other confined spaces. "People's blind fear or unfounded confidence against the infection of COVID-19 is simply because it is invisible," Tsubokura told *Reuters*. With Fugaku's 442 petaFLOPS at their disposal, the researchers have been able to visualize—and now, hopefully limit—coronavirus spread.

SUPERCHARGING HIGHER EDUCATION

Students and staff at Singapore's Institute of Technical Education (ITE) will now have access to supercomputing resources from the National Supercomputing Center (NSCC) Singapore, following a Memorandum of Understanding signed between the two organizations in October 2020.

Staff will be able to leverage supercomputing power to experiment and innovate, while students can access it for their learning and projects. The technology will also boost ITE's artificial intelligence (AI) experiments in teaching and assessment. This includes innovative projects like analyzing in-class student videos to identify topics that are well-received and areas of lesson delivery that require improvement. Video analytics and AI-based assessments can also assess students as they wear smart glasses and perform timed practical tasks.

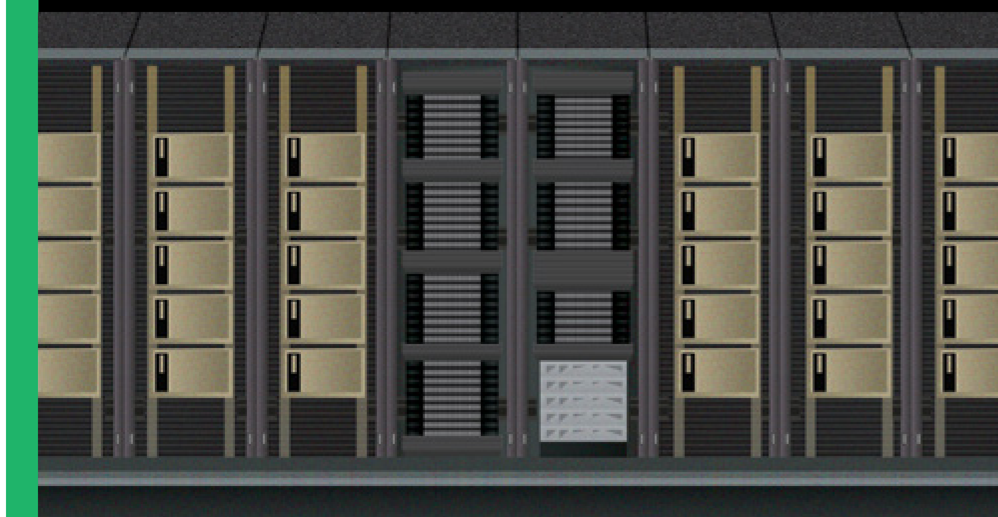
"High performance computing will play a key role in the development of smart nation innovations," said Associate Professor Tan Tin Wee, chief executive of NSCC. The sharing of supercomputing resources will help local researchers in creating such innovations, and give the next generation an early start and competitive edge in the supercomputing industry, he added.

ACCELERATING AI RESEARCH IN THAILAND

In a boost to Thailand's AI research infrastructure, CMKL University in Bangkok is setting up an AI computing cluster that will be the central node connecting to other research and university nodes across the country.

The cluster, which is modeled after NVIDIA's DGX POD, consists of six NVIDIA DGX A100 AI systems delivering 30 petaFLOPS of AI processing power and 2.5 petabytes of storage. CMKL will also set up an AI analytics platform to drive research in domains ranging from food and agriculture to healthcare and smart city innovations. Not only will the platform enable researchers across Thailand to easily store and manage their datasets, it will also facilitate data exchange and the running of models at lightning-fast speeds through cloud HPC infrastructure.

"This platform will accelerate AI work in various research and development fields to create substantial positive impacts on society. Examples are an increase in quality of life in cities through better management and logistics; better optimization of circular infrastructure, consumer insights, and biomedical research; and an increase in the quality and quantity of crop yields countrywide," said Dr. Akkarit Sangpetch, a program director at the university.



WHAT'S UP!

A VIRTUAL FIRST FOR SCA2021

SupercomputingAsia (SCA) will be back for its first-ever virtual conference from March 2–4 in 2021. First organized in 2018 by the National Supercomputing Center (NSCC) Singapore, the annual event routinely draws hundreds of delegates from more than 20 countries around the world.

SCA21 marks the much-anticipated return of Asia's premier high performance computing (HPC) conference, after the pandemic's onset in early 2020. This year's iteration will dive deep into the role of supercomputers in the fight against the coronavirus with the theme 'Supercomputing in the New Norm: Adapting to COVID-19 and Beyond.'

As always, participants can expect an exciting roster of illuminating keynotes from global HPC leaders in academia and industry, as well as numerous virtual networking opportunities. Save the date and see you online soon!

For more information, visit https://www.sc-asia.org

WHAT

SupercomputingAsia 2021

WHEN

MARCH 2-4, 2021

ISC HIGH PERFORMANCE 2021 GOES DIGITAL

The latest edition of ISC High Performance, the world's oldest and Europe's most important high performance computing (HPC) conference, will be redesigned as a digital conference to be held in the summer of 2021.

First held in 1986, the event brings together engineers, IT specialists, vendors, scientists, students and other members of the HPC community worldwide. The 2021 conference, otherwise known as ISC 2021 Digital, is expected to draw over 3,700 practitioners and 160 exhibitors. ISC 2021 Digital will revolve around five core topics critical to current developments and future advancements in HPC: system architecture, applications and algorithms, emerging technologies, parallel programming and machine learning.

Participants can look forward to insightful discussions on topics ranging from performance modeling and tuning for exascale systems to machine learning in the life sciences. ISC 2021 Digital will also present multiple opportunities for participants to learn about new products and technologies in the supercomputing sphere, exchange ideas with peers and make new connections. The full event program and registration details will be provided in January 2021.

For more information, visit https://www.isc-hpc.com

WHAT

ISC High Performance 2021

WHEN

SUMMER 2021

CRUNCHING THE NUMBERS OF CLIMATE CHANGE

Modeling our changing climate

From short-term weather forecasts to decades-long climate simulations, supercomputers are giving scientists a peek into our planet's future.

By **Sheryl Lee**

When much of the world began to implement lockdowns in March last year, shops shuttered, people stayed home and roads were emptied of traffic. Cities from Bangkok to Delhi collectively inhaled fresher air and basked under bluer skies. While you'd assume near-universal stay-at-home orders would significantly cut carbon emissions—especially those from transport—think again. Though emissions had declined an unprecedented 17 percent by April, scientists note that this marked change is likely only temporary. After all, even as economic activity virtually ground to a halt, greenhouse gases continued to be emitted by everything from household appliances to factories.

Individual behavioral change—whether it's commuting less or recycling more—will only have a small impact without a broader revamp of our energy system and industries. But how do we even quantify the environmental impacts of industrial activity? Or visualize the world as it warms so that we can better deal with it?

To find out, scientists have been collecting data 'snapshots' of our climate, measuring everything from temperature to air pressure and sea levels. Hidden within this deluge of data is a latent prophecy of a world decades from now—a future we can unlock with the help of powerful supercomputers that can rapidly process large amounts of data. Whether they are trying to predict the weather in sunny Singapore or understanding planetary trends, climate researchers need supercomputers to help them understand the sheer scale and complexity of this place we call home.

FITTING A PLANET IN A COMPUTER

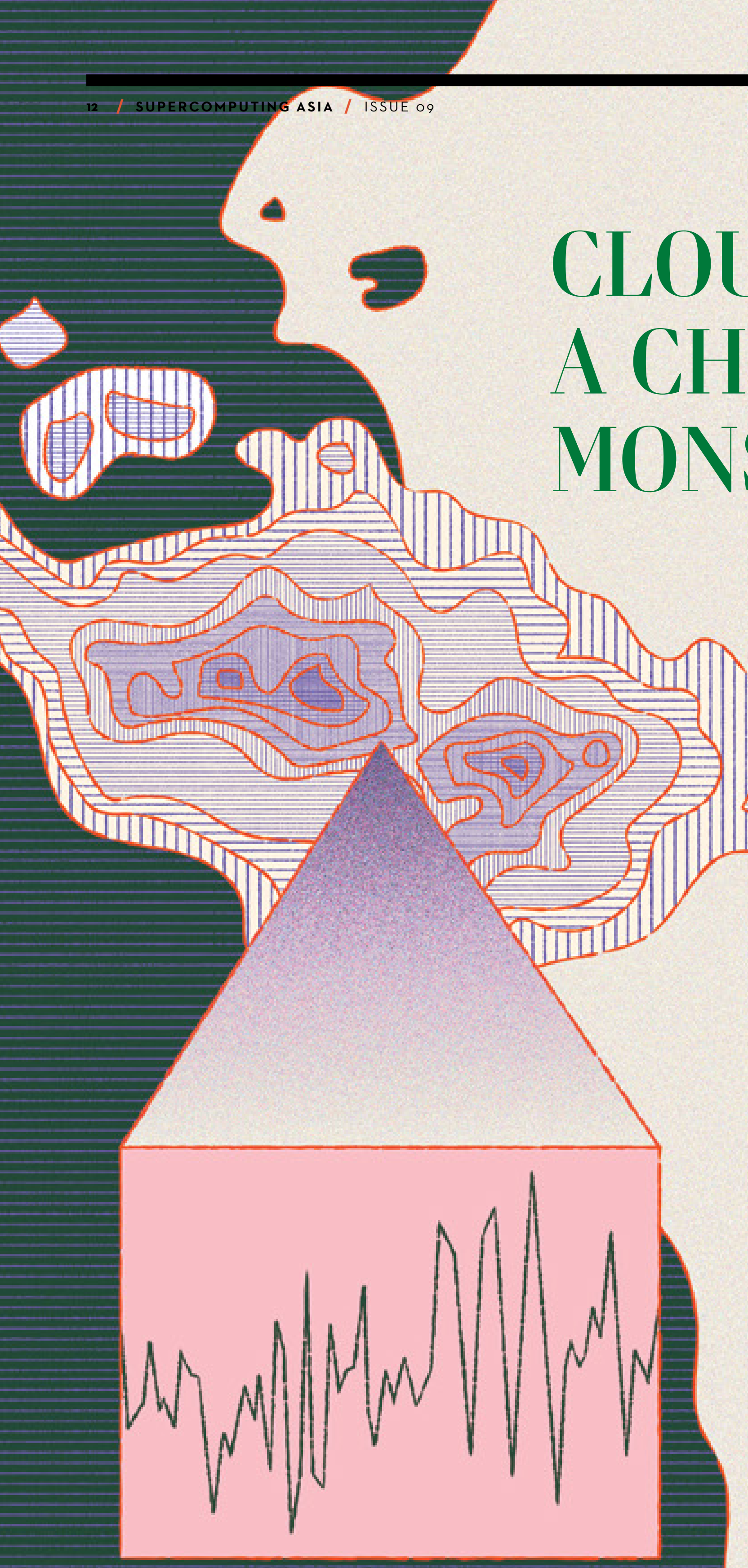
Decades ago, the most primitive climate models were only able to calculate surface temperature changes. Called energy balance models, they considered just one thing: the balance between energy entering the earth's atmosphere from the sun, and heat released back out to space. Today, their upgraded successors, known as Earth system models, can account for much more: heat transfer and flows of air and water between the atmosphere, oceans, land and ice; biogeochemical processes like the carbon and nitrogen cycle; even the impact of industrial activity.

One such model—the Energy Exascale Earth system model (E3SM) developed by the US Department of Energy—uses supercomputers to simulate the Earth's climate at different locations and predict its continuing evolution over time. Because the climate system is so large and complex, even the most sophisticated models cannot calculate wind, air pressure, temperature and humidity at every point on earth. Climate models therefore split the globe up into many three-dimensional boxes known as grid cells. Incredibly, the E3SM is able to capture our climate at a high resolution of 15–25 kilometers compared with hundreds of kilometers in the older models.

To properly account for all the biogeochemical processes, the E3SM also requires a vast network of specialized modules. Each module maps out an environment, from rainforest to river systems. Because each system and biogeochemical process impacts others in an intricate web, scientists have to contend with everything from the effect of plants on the water cycle, to how changes in surface temperature affect cloud formation.

Thanks to such detailed modeling, E3SM can predict near-term climate change effects: in particular, the formation and impact of the strongest storms threatening human life, property and coastal infrastructure. This helps provide advance warning for evacuation, visualize network outages for response efforts, and even guide investments in resilient infrastructure.

Illustrations by Shelly Liew/Supercomputing Asia

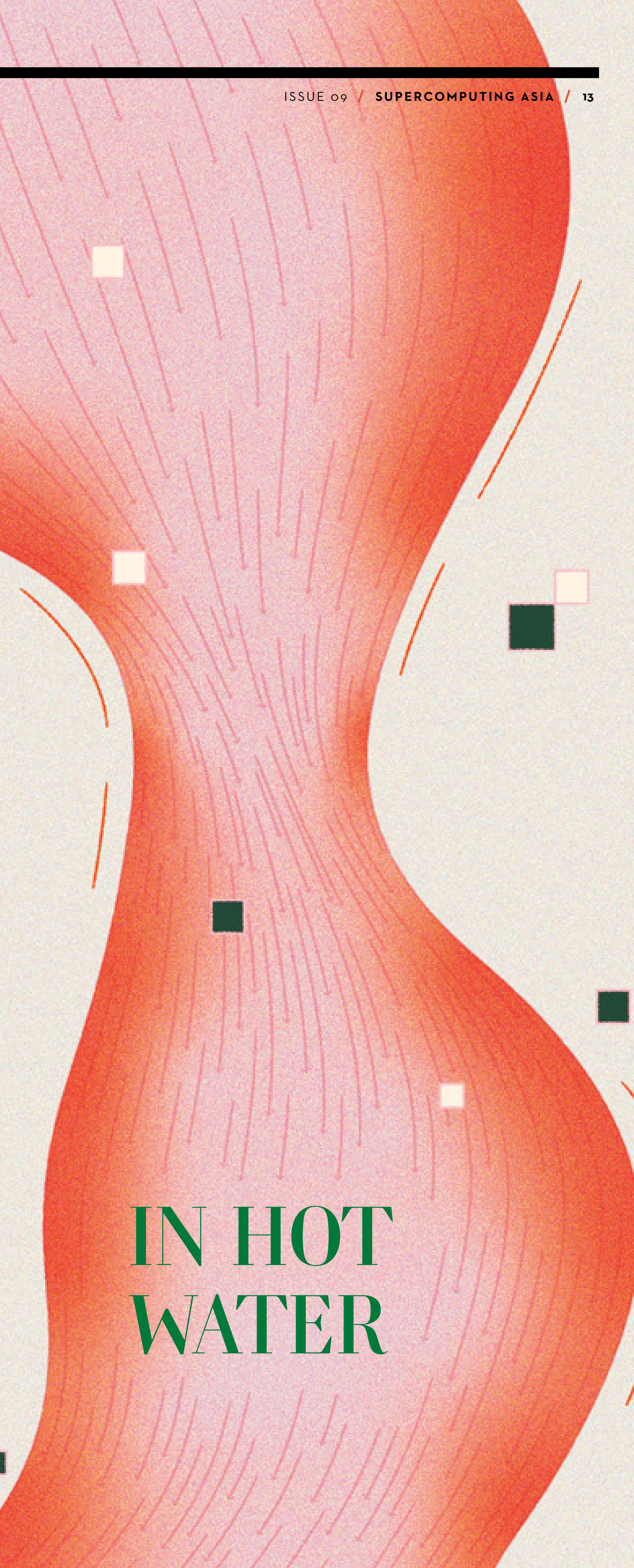


CLOUDY WITH A CHANCE OF MONSOONS

In India, where 42 percent of workers depend on agriculture for their livelihoods, supercomputers are giving farmers new hope for a better harvest. More than half of arable land in the country is dependent on rainfall, which means the monsoon season, lasting from June to September, has an outsized impact on the agriculture sector and other industries linked to it. As it stands, extreme weather events cost India US$9–10 billion annually. With the growing threat of climate change, productivity decreases in major crops could rise to as much as 40 percent by 2100 unless farmers adapt.

Enter Pratyush and Mihir—'rising sun' and 'sun' in Hindi—a supercomputing duo with a combined 6.8 petaFLOPS of power allowing the government to monitor monsoons and predict extreme weather events. The fourth most powerful weather research computer in the world, Pratyush can simulate world weather conditions at a high resolution of 12 kilometers, in 64 different layers of the atmosphere. For India, it can achieve a resolution as fine as three kilometers. As a result of this added computing power, the number of farmers who receive weather information through text messages is expected to increase from 24 million to 45 million.

In climate models, researchers use equations to represent the physical principles of our world. Weather forecasting is done by solving these equations. Pratyush has to crunch the numbers at every single imaginary point on the globe placed 12 square kilometers apart, then repeat this again for all the different layers in the earth's atmosphere. The calculations are then run 20 to 30 times while changing variables like temperature and humidity to ensure predictions are as accurate as possible. Clearly, monitoring monsoons is no mean feat.



IN HOT WATER

The holy Ganges river in India is covered with floating trash and ash from corpses cremated on its banks. Thousands of kilometers away in Malaysia, the Sungai Kim Kim river, tainted by chemical waste, poisons children with its noxious fumes. Such pollution is visible and visceral, but humans also directly pollute waterways in hidden and equally insidious ways—like discharging heated water that comes from thermal power plants into the environment. Over time, such thermal pollution tips the balance of our underwater ecosystems as fish and amphibians migrate to cooler regions and microbes flourish in warmer water, using up the dissolved oxygen.

Now, a team of researchers from the Chinese Academy of Sciences has come up with a model to quantify the impact such anthropogenic—that is, human-induced—heat emissions have on river temperatures worldwide. The scientists incorporated a large-scale scheme for computing river water temperature into a river routing model coupled to a global land surface model, allowing them to simulate river flows and show changes in water temperatures over the past 30 years.

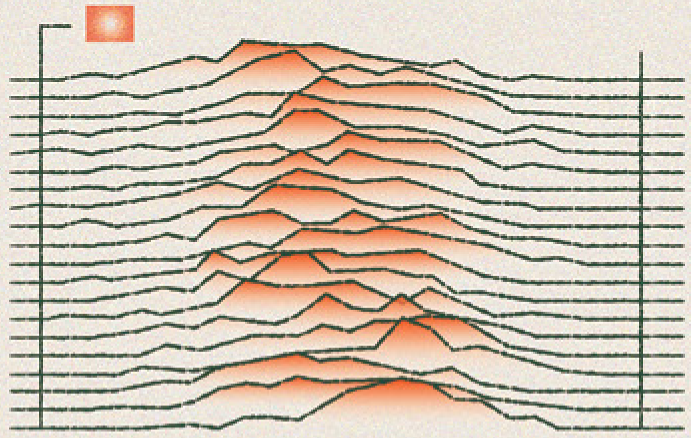
Running the model on the Tianhe-1 supercomputer, the study found that water temperatures in tropical rivers increased the most, about 0.5°C per decade from 1981 to 2010. In Asia, power plants alone increased local river temperatures by about 60 percent. "Climate change warmed river water, and anthropogenic heat emission further made the temperature increase," the researchers concluded. "The rivers, even climate, are gradually being dominated by humans."

LOOKING INTO THE CLIMATE CRYSTAL BALL

While global climate models like the E3SM are useful in simulating our planet, scientists usually downscale them to regional climate models when evaluating specific regions. Because regional models cover smaller areas, they can be run faster and at a higher resolution. They also allow researchers to input data on local topography and features like forests and lakes for more accurate climate prediction.

To understand the impact of a 2–3°C warming on East Asia's ecosystems, a team of South Korean scientists classified the region into distinct climate zones, such as tropical or polar. They then ran a regional climate model to investigate how each zone's climate might change with global warming. Using Supercomputer No. 4 from the Korean Meteorological Administration, the researchers were able to tweak a range of inputs including carbon concentration levels on earth, varying them in line with different estimates of expected emissions to derive multiple possible scenarios.

Peeking into a warming world decades from now, the team highlighted remarkable changes in climate in many of the zones surveyed. Most significant of all was the erosion of tundras in high altitude regions. At 2°C of warming, tundras in the Tibetan plateau would shrink by half. At 3°C, they would vanish completely. As the source of the largest rivers in Asia, the Tibetan Plateau and its glaciers play a crucial role in the water cycle of surrounding countries. When these glaciers shrink with climate change, rivers in Asia will reach 'peak water' in the next few decades before their flows decline, potentially leaving hundreds of millions thirsty. Thankfully, with advanced warnings from supercomputers, there's hope to prevent this yet.

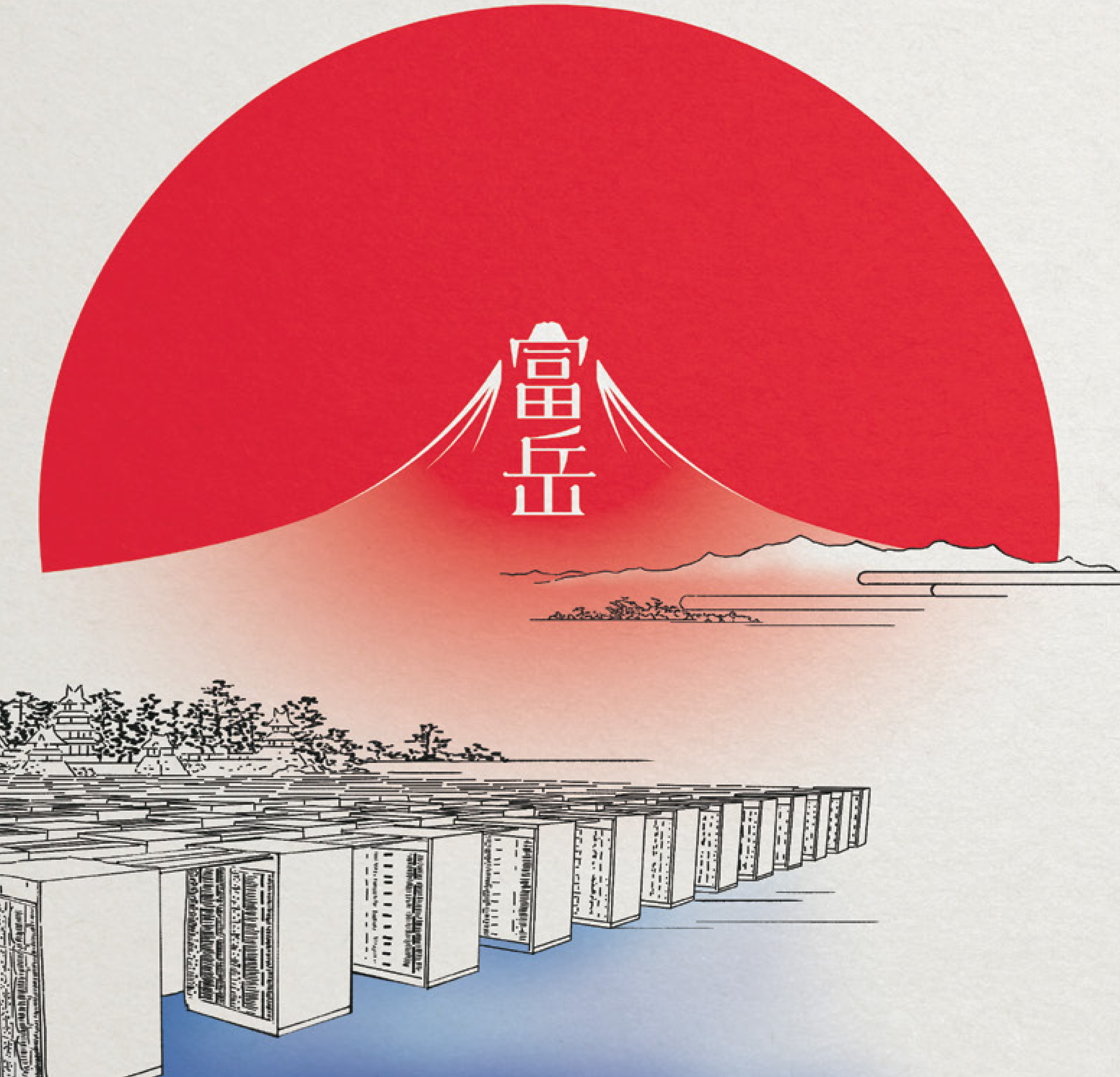


SIMULATIONS FOR SMALL SPACES

Measuring only 27 kilometers from north to south, Singapore is one of the smallest nations in Asia. Even if scientists were to try simulating its weather on some of the best global climate models, it would just show up as a blip in one of the grid cells. Thankfully, researchers from the Climate and Water Research cluster at the local Tropical Marine Science Institute have been able to downscale global and regional climate models to capture the whims of the weather in the city-state.

If you thought calculating a country's climate would be way less intensive than computing the whole world's, you're in for a surprise. To process the terabytes of data required for detailed and accurate weather prediction in such a small geographical area—and to do so quickly enough to issue four forecasts a day—the scientists have to rely on supercomputers. In particular, they run their islandwide weather simulations at a high resolution of 400 meters on the Köppen system at the National Supercomputing Center (NSCC) Singapore. With 1.2 petabytes of storage and 160 teraFLOPS of power, the system is specifically designed for climate and environment research.

Beyond short-term weather forecasts, the researchers also run simulations to investigate the effects of Singapore's weather and climate on its urban climate, as well as the seasonal haze phenomenon. They've even extended their simulations to the Southeast Asia region, modeling climate change at 8–10 kilometer resolutions on multi-decade time scales. Singapore may be small, but its supercomputing prowess is mighty.



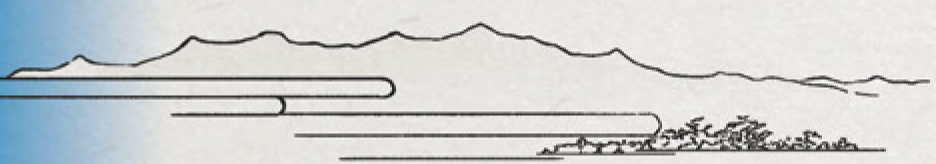
PEAK PERFORMANCE

How Japan's Fugaku balanced power and efficiency for the best of both worlds

Designed with general purpose users in mind, the Fugaku supercomputer has the performance and energy efficiency of a GPU-accelerated machine but remains easy to program.

By **Tim Hornyak**

When the novel coronavirus became a pandemic in spring 2020, Japan became something of a mystery to international observers. Why were relatively few people becoming infected and hospitalized, let alone dying? The debate witnessed all kinds of possible explanations, from the fact that Japanese routinely wear face masks to the continued use of the Bacillus Calmette-Guerin vaccine for tuberculosis. Few people, however, pointed to Japan's high-tech prowess, specifically its creation of a machine that would be harnessed to fight the pandemic: the Fugaku supercomputer.



Developed by Japan's state-backed RIKEN research center and Fujitsu, Fugaku is remarkable for several reasons. Not the least of which is its current status as an exascale machine that's the fastest supercomputer in the world—it has a High Performance Linpack result of 442 petaFLOPS, according to the November 2020 ranking by the TOP500 project. That's more than 442 quadrillion computations per second, which is about 2.8 times speedier than the US Oak Ridge National Laboratory's Summit machine, which Fugaku dethroned.

"Fujitsu's Fugaku installed at RIKEN is an impressive supercomputer that was designed for large-scale scientific computations," said Prof. Jack Dongarra, a TOP500 author who holds appointments at the University of Tennessee and Oak Ridge National Laboratory. "This shows in its taking the number one position for the TOP500, the HPCG, and the HPL-AI benchmarks."

In addition to those, Fugaku also scored tops in the Graph500 ranking, earning it a quadruple crown. But apart from returning Japan to the lead in the supercomputing race after almost a decade, Fugaku represents a novel chip architecture and an unprecedented deployment of high-performance computing to fight a once-in-a-century public health threat.

PETAFLOPS FOR THE PANDEMIC

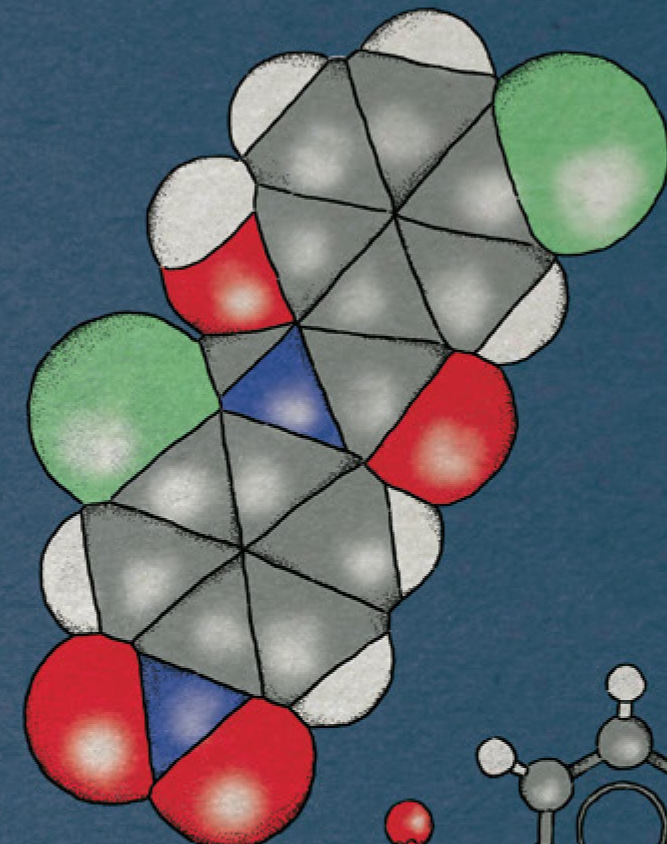
There's a stop on Kobe's Port Liner rail line called K Computer Mae, an homage to the last Japanese machine to summit the TOP500. But the automated trains have been running with plaques commemorating Fugaku's success, and at the adjacent RIKEN Center for Computational Science (R-CCS), the K's old red facade has been replaced with blue one emblazoned with an illustration of Mount Fuji. Given Fugaku's accolades, Japan's highest peak is an appropriate image for the machine—'Fugaku' is an alternate name for Mount Fuji, and the two names share the Chinese character 富, meaning wealth. But leaders the project say it wasn't built to win HPC races.

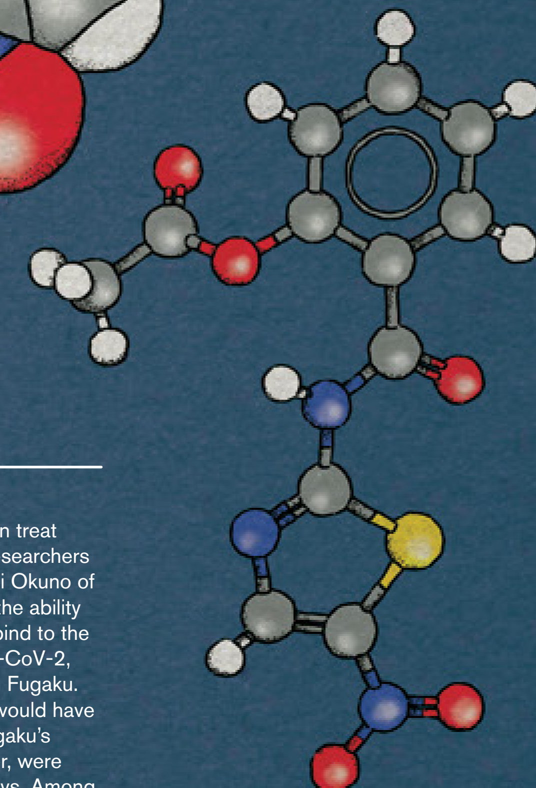
"Fugaku as a supercomputer has never been about ranking," Satoshi Matsuoka, director of R-CCS, told *Supercomputing Asia*. "Rather, our goal was basically to achieve application results, and respond especially to important and difficult societal goals and problems."

That aspiration was put to an early test by the novel coronavirus. In 2019, components of the K computer were being removed from R-CCS not long before COVID-19 cases began in December 2019, with full fledged operations scheduled to begin only in April 2021.

By spring, however, the coronavirus caused global supply-chain issues affecting the delivery of Fugaku components. Meanwhile, the center had received an appeal from the Japanese government asking whether the computer could be used in the fight against the virus as it was spreading quickly in Japan. In March and April, Fugaku's leaders quickly mobilized scientists to redeploy parts of Fugaku that were already up and running. The computer would be used for five projects: revealing the characteristics of the coronavirus; identifying therapeutic compounds; improving the diagnosis and treatment of COVID-19; insights into the spread of infections and socio-economic impact; and other projects that could help in the battle against the pandemic.

Within weeks, the first results were announced. In July 2020, researchers from RIKEN and Kyoto University said Fugaku had helped them narrow down a field of 2,128 drugs to dozens of potential therapeutics that might bind to and inactivate proteins that play a role in how the virus particles spread. Fugaku took ten days running





To identify drugs that can treat COVID-19, a team of researchers led by Professor Yasushi Okuno of Kyoto University tested the ability of over 2,000 drugs to bind to the main protease of SARS-CoV-2, using simulations run on Fugaku. The simulations, which would have taken over a year on Fugaku's predecessor K computer, were completed in just ten days. Among the potential drugs identified were niclosamide and nitazoxanide, a treatment for tapeworm infestations and a broad-spectrum antiparasitic respectively.

TofuD Interface

The TofuD interface has 20 lanes of high-speed serial signals of 18 Gbps transmission speed and interconnects up to 10 CPUs with a bandwidth of 6.8 GB/s.

PCIe Interface

The PCIe interface has 16 lanes of high-speed serial signals of 8 Gbps transmission speed and a bandwidth of 16 GB/s.

Core

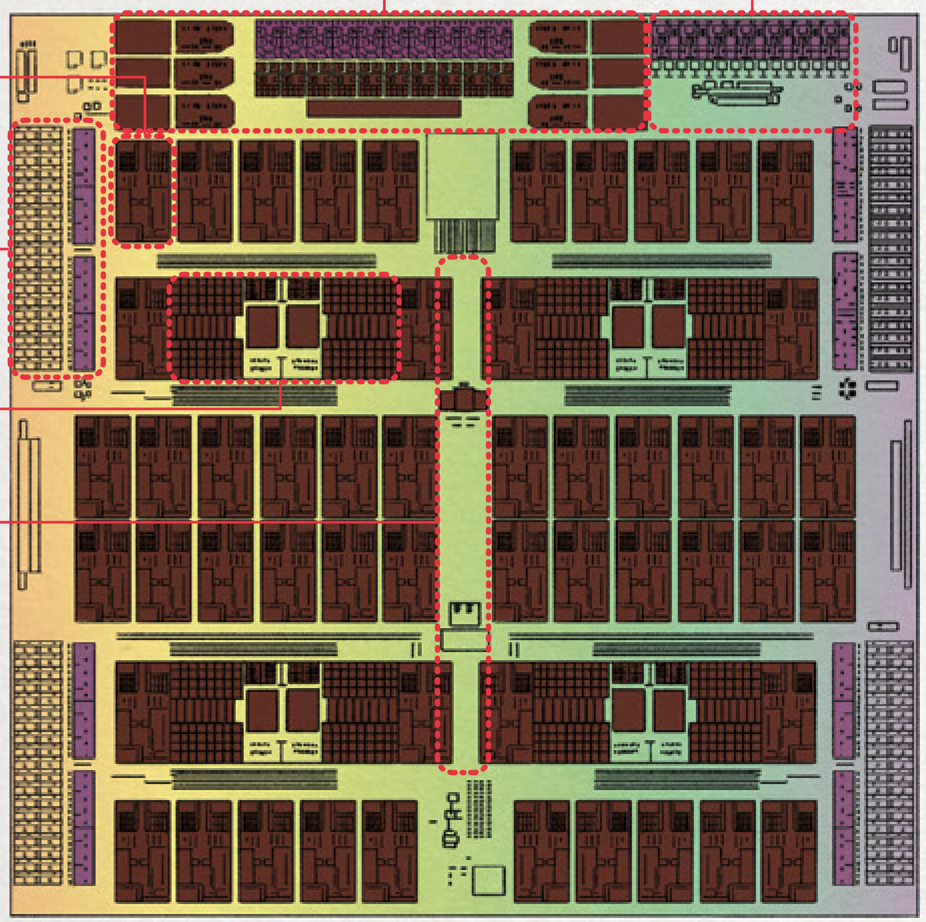
The processing is done on 48 compute cores aided by four assistant cores, giving each A64FX chip over 8.7 billion transistors.

HBM2 Interface

Each HBM2 has 32 GB of memory, giving the processor package a total of 1 TB/s of memory bandwidth.

Level 2 Cache

Ring Bus



Core Memory Group (CMG)

ARMED TO THE TEETH

The secret to Fugaku's success is the A64FX processor—or rather, 158,976 of them. While earlier generations of supercomputers used large numbers of x86 CPUs to achieve their performance, the TOP500 list has recently been dominated by hybrid architectures combining CPUs with GPUs to accelerate workloads. Fugaku, however, does neither. Instead, Fugaku bucks the trend with the ARM-based A64FX, which are CPUs that can perform like GPUs. Not only are the A64FX chips more powerful, they also consume considerably less energy than existing alternatives. The result is a supercomputer that is not only the most powerful in the world, but one that is also easy to work with and extremely energy efficient.

molecular simulations to find the candidates, some of which had not yet been considered as possible agents.

Apart from screening drug candidates, Fugaku has also helped scientists understand how the coronavirus can spread while airborne. A research team from RIKEN and Kobe University was preparing to use Fugaku to simulate vehicle fuel-injection systems when one member realized that it was basically the same mechanism as coughing. Working with biologists and other experts, the team ended up running what it described as the world's largest virus-droplet simulation ever conducted. Instead of atomized fuel, Fugaku was put to work modeling how the coronavirus can spread in trains, workspaces and schools.

They found that opening train windows, for instance, can triple ventilation and lower ambient particles. Another study looked at how particles emitted by infected people would behave in a variety of environmental conditions. They found that if air humidity is under 30 percent, there are more than twice the number of aerosolized particles compared to levels over 60 percent. Other findings revealed that the biggest risk to diners is people sitting next to, and not across from them, and that face masks are more effective than clear face shields.

"They got probably the best simulation in the world, because they were developing this extremely complicated but capable code that would not work too well on conventional machines because they're too slow," said Matsuoka, adding that some simulations that took a few days on Fugaku would have taken over a year on the K computer. "We had a very focused, very unified response to the pandemic. Everybody contributed according to their abilities and we moved quickly in the same direction."

> Most applications are limited by data movement rather than compute… Fugaku's objective is to break away from that.

BUILT TO MAKE A DIFFERENCE

It took ten years, about US$1 billion and the work of thousands of members of Japan's HPC community to realize Fugaku. RIKEN began working with Fujitsu in 2011 to design a processor that would be 100 times faster than applications on the K computer. The result is the A64FX processor, based on the ARM 8.2A instruction architecture set for supercomputers, with hardware barrier, sector cache and prefetch Fujitsu extensions. The A64FX has 48 calculation cores and either two or four assistant cores, with a theoretical peak performance of 3.3792 teraflops for double precision floating point calculations.

"It's like a chimera," said Matsuoka. "It has GPU-class speed but you can program it like a normal CPU, so it could run the same stuff as the CPU in your smartphone. Secondly, the power efficiency is also similar to that of a GPU for high-end applications. We were able to do that by coming up with new technologies that have not been tested on CPUs before."

Another notable aspect of the processor is that it attempts to buck the trend in which compute power has increased at the expense of the ability to move data around. Fugaku has solid performance across the board because the A64FX emphasizes data movement, Matsuoka said, adding that he expects other HPC researchers and developers to follow the same approach.

"Most applications are limited by data movement rather than compute," he said. "People have built machines while somewhat ignoring their utility and they get these artificial results which may not be telling the whole story with regards to the true performance of the chip or the machine across many applications. Fugaku's objective is to break away from that."

The processors are arranged in 158,976 nodes, each consisting of one

CPU and 32 GiB of memory, in 432 racks, and provide a peak performance in 2.2 GHz boost mode of 537 petaFLOPS for double precision (64 bit) and 1.07 exaFLOPS for single precision (32 bit).

"It's large scale, uses no accelerators, is very good for a number of applications having very fast, very good interconnect between the processors, has very high bandwidth, has quite a bit of fast memory on this machine," said Dongarra in a YouTube chat with *Primeur Magazine*. "We're looking forward to see how it performs on real scientific problems."

For Matsuoka, what distinguishes Fugaku from the K and other supercomputers is the fact that it's dedicated to being general purpose. He compares developing a very high-performance 'normal' supercomputer, as he describes it, to the extremely difficult challenge that Toyota Motor faced in the 1990s when it developed the Prius, the world's first mass-produced hybrid vehicle. While fuel-efficient hybrids are normal now, developing the Prius involved overcoming technological, production and economic challenges. In so doing, however, Toyota helped establish a new normal.

"In order to satisfy the goal of being able to cover a broad range of applications, Fugaku had to be absolutely standard. It's one thing to have this hardware, but people have to be able to program the software. It's all about the so-called software ecosystem, but at the same time as being normal, it has to be a supercomputer," said Matsuoka. "It's like a Prius but it's as fast as a supercar, so you can drive it to the supermarket but it can compete against any other supercar in the world. It is ease of use and performance at the same time. I think that's what makes Fugaku quite unique."

It is ease of use and performance at the same time. I think that's what makes Fugaku quite unique.

TACKLING THE BIGGEST QUESTIONS

Even as the pandemic continues, Fugaku has been used to further research in other areas. Scientists from Tokyo Medical and Dental University and Fujitsu Laboratories were able to estimate a network of regulatory relationships among genes in epithelial cancer cells. The work, which was based on analysis of data on 20,000 genes, would have taken several months on university supercomputers but was completed in less than a day on Fugaku.

The supercomputer will be used to crunch numbers related to other social and scientific issues of high priority such as drug discovery, personalized and preventive medicine, earthquake and tsunami simulations, big data weather and climate prediction, energy generation, transport and storage, materials science, next-generation design and manufacturing, and basic science research such as the fundamental laws and evolution of the universe.

"In terms of disasters, Japan gets a truckload every year, and supercomputer-derived results that have been used for disaster mitigation have already paid for the supercomputer in terms of money and lives," said Matsuoka. He's personally intrigued by the potential of Fugaku to create a virtualized smart city, including virtual sensors, IoT terminals and networks, as well as virtual environmental and weather, because it's difficult to scale experimental parts of a city that have been fitted with instruments. If the experiments yield positive results, it could be applied to real urban settings.

Fugaku could also be used to come up with new technologies or designs for next-generation supercomputers including potential successors to Fugaku and, further down the road, quantum computers. Finally, the supercomputer can be used to simulate cognitive activity in higher mammals including humans—in 2013, the K computer simulated one second of human brain activity.

"It's quite phenomenal that we can run a human brain-scale simulation, and it may lead to discoveries of brain phenomena that are not well understood," said Matsuoka. "We can also use it for artificial intelligence research and driving science with hypotheses, such as the next state of a molecular structure, and then using physics to validate them."

If Fugaku must live up to, and surpass, the research legacy of the K computer—whose achievements included simulating a human heart at the molecular level, a feat that contributed to some 1,300 research papers—then it's already off to a promising start. There are great expectations riding on this billion-dollar machine.

ACCESSING FUGAKU FROM SINGAPORE

Through a partnership between the RIKEN Center for Computational Science (R-CCS), Japan's Research Organization for Information Science and Technology (RIST) and the National Supercomputing Centre (NSCC) Singapore, the Fugaku system is now open to Singapore-based researchers interested to leverage the HPC resources of the world's premier supercomputer. Successful applicants will be able to directly access Fugaku's 442 petaFLOPS of computing power via a dedicated high-bandwidth optical fiber link at speeds of up to 100 Gbps.

In a nationwide call for HPC startup projects, NSCC is offering to assist researchers by reviewing their proposals for resources to use the Fugaku system. Alternatively, interested parties are free to submit their proposals directly to Japan's High Performance Computing Infrastructure (HPCI).

Who should apply?

- Researchers in fields such as genetics, biomedicine, advanced modeling and simulation, materials development, for projects that require advanced HPC resources
- Researchers working on projects at the forefront of HPC, with potential for national and global high impact outcomes

How do I apply?

- Download the application form from www.hpci-office.jp
- Refer to "A Guide for Proposal Writing" to fill up the form
- Submit the form to projects-admin@nscc.sg with the email title "Application: Call for Fugaku Startup Projects"
- Alternatively, submit the form directly through the HPCI Online Application System www.hpci-office.jp/entry/login/index
- R-CCS will assess, review and approve the applications for the usage of Fugaku, including whether an application is compliant to the export control regulations of the jurisdiction to which R-CCS belongs

Deadlines:

- For consultation with NSCC: February 28, 2021
- For direct application to HPCI: March 31, 2021**

**May be closed earlier at the discretion of HPCI once all available resources are allocated

Project period:

Up to six months from the end of March 2021

For clarifications and queries, please contact NSCC at **projects-admin@nscc.sg**

KEEPING SINGAPORE COOL

How HPC is helping cities like Singapore beat the heat

With global warming at our doorstep, researchers in Singapore are leveraging supercomputers to tackle the complex challenge of urban heat.

By **Kamila Navarro**

Two years ago, the Intergovernmental Panel for Climate Change issued a stark warning: limit global warming to 1.5°C by 2030 or risk cataclysmic consequences. Should we surpass this limit to 2°C, large swaths of the world will become uninhabitable—and not just for humans. Though seemingly insignificant, a half-degree difference could expose a third of the world's population to extreme heat waves and wipe out nearly all coral reefs.

While Singapore has so far been spared the worst of climate change's effects, it's only a matter of time. With an increase of 0.25°C per decade, the island is heating up twice as fast as other low-latitude regions. By the end of the century, daily temperatures of 35–37°C could become the norm in an already-humid setting. Under such conditions, our body's ability to regulate temperature is compromised, resulting in potentially life-threatening illnesses like heat exhaustion and even heat stroke. As extreme heat reduces quality of life, it also threatens Singapore's overall productivity and attractiveness as an investment destination.

But Singapore isn't the only city that's getting warmer. Despite covering only three percent of the Earth's surface, cities emit more than 60 percent of total greenhouse gases and consume about 75 percent of global energy. With less than ten years left for the world to collectively beat the heat limits warned by the IPCC, scientists are harnessing supercomputers to design cooler, more liveable cities and turn the tide of climate change.

UNDERSTANDING URBAN HEAT

Singapore's gleaming cityscape may be instantly recognizable, but its rapid urbanization also comes at a cost. As shade-providing greenery is replaced by infrastructure that absorb or produce heat, the island's urban centers are getting warmer compared to surrounding rural areas. This phenomenon, called the urban heat island (UHI) effect, is part of a vicious cycle. Faced with a warming city, Singapore's residents are resorting to air conditioning and private transport—two activities that, in turn, generate heat.

Greenery, infrastructure, air-conditioner use and transport are just a few of the many factors that influence urban heat in cities. Given the complex interrelationships between these elements and the veritable deluge of data, supercomputers have become essential tool for understanding urban heat and fighting climate change as a whole. "Scientists use models to understand these interactions and project future changes," noted Professor Dale Barker, director of the Centre for Climate Research Singapore (CCRS) of the Meteorological Service Singapore under the National Environment Agency.

"Weather and climate modeling is numerically intensive work, for which scientists have relied on computers for over half a century," he added. At the CCRS, for instance, Barker shared that they leverage an in-house Cray XC40 supercomputer called Athena with a peak capacity of 212 teraFLOPS, 6,336 cores and 3.6 GB RAM of memory per core. Through Athena, CCRS produces numerical weather forecasts of temperature, wind, rainfall and other meteorological variables through their SINGV system. To deliver their next set of regional climate predictions, CCRS will also be partnering with the National Supercomputing Center (NSCC) Singapore to tap into up to 3.2 petaFLOPS of capacity.

VIRTUAL CITIES, REAL IMPACT

As Singapore rapidly digitalizes, addressing urban heat through computational means has become quite literally a hot topic in recent years. Rising up to the challenge is a team from the Agency for Science, Technology and Research (A*STAR) and the Housing & Development Board (HDB). In 2019, the researchers won the President's Technology Award for developing an advanced modeling tool powered by high performance computing known as the Integrated Environmental Modeler (IEM).

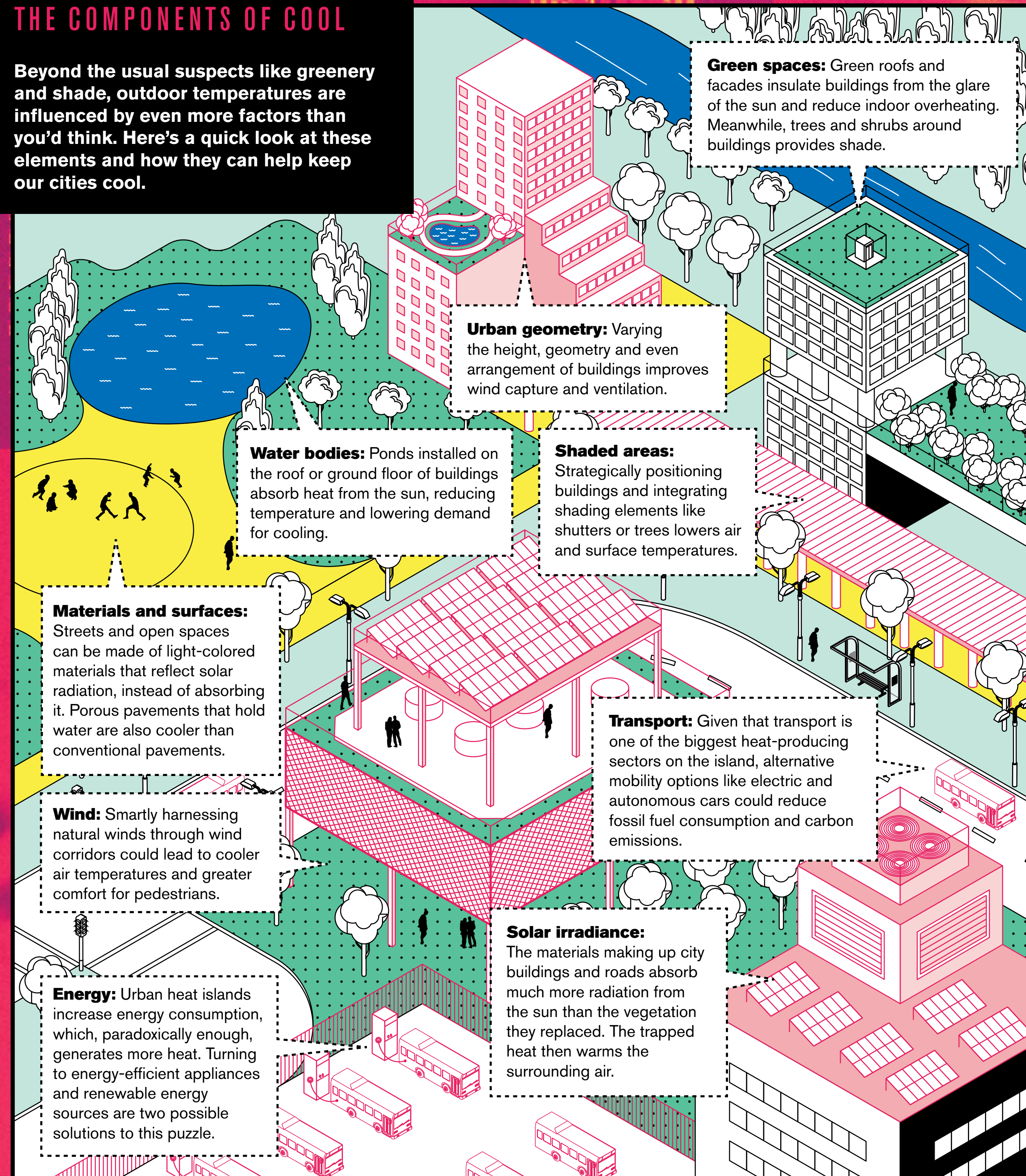
"The IEM integrates the urban planning and design process with environmental simulation," explained Dr. Poh Hee Joo, a senior scientist at A*STAR's Institute of High Performance Computing and one of the project's lead researchers. "It allows users to predict the interrelationships and combined impact of solar, wind, temperature, noise and other environmental factors in an urban setting."

Not only can the IEM model the combined effects of the environmental factors, but it can also simulate their individual impacts. Effects of urban elements like buildings and roads as well as natural features such as vegetation are captured on the platform. This makes the IEM the first truly integrated tool of its kind, as other commercially available modelers typically assess only a single environmental factor at a time.

To develop the IEM, Poh and his colleagues first painstakingly transformed three-dimensional (3D) geometric data of Singapore into a highly realistic simulation of the city. At the same time, 43 solar-powered sensor nodes were deployed in Punggol and Singapore's eastern half to collect environmental data on wind, temperature and solar irradiation. Data from

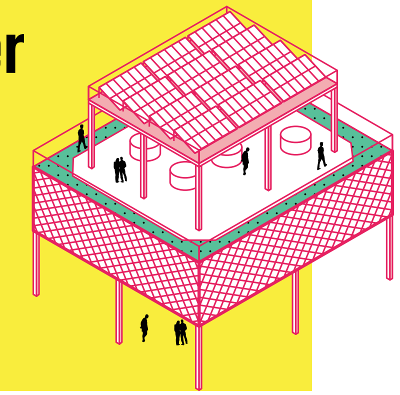
THE COMPONENTS OF COOL

Beyond the usual suspects like greenery and shade, outdoor temperatures are influenced by even more factors than you'd think. Here's a quick look at these elements and how they can help keep our cities cool.



Illustrations by Lam Oi Keat/*Supercomputing Asia*

"The Integrated Environmental Modeler will reduce the overall carbon footprint for new housing estates to achieve greater sustainability and liveability."



Dr. Poh Hee Joo
Senior Scientist
Institute of High Performance Computing
Agency for Science, Technology and Research

these sensors were then transmitted wirelessly to the research team, allowing them to validate the models' results. By allowing for easy visualization of the effect of various factors on the urban environment, users could first refine and test their designs computationally on the platform, reducing the risk of costly physical trial-and-error.

Pushing the boundaries of supercomputing, the team performed the first-ever 3D air flow simulation that depicts all of Singapore's buildings on the IEM. This simulation, performed at a ten-meter horizontal resolution, was completed in only five days using 6,000 processors of NSCC's ASPIRE 1 supercomputer. According to the researchers, having an island-wide wind map could help identify wind corridors, allowing urban developers to plan and design cooler buildings that maximize wind flow, promote natural ventilation as well as reduce light from the sun. "Therefore, the IEM will reduce the overall carbon footprint for new housing estates to achieve greater sustainability and liveability," concluded Poh.

A DIGITAL TWIN FOR A DIGITAL CITY

Poh and his team aren't the only ones tackling the urban heat challenge. As quite literally indicated by its name, the project Cooling Singapore, led by Professor Gerhard Schmitt of the Singapore-ETH Center (SEC), is dedicated to identifying cooling strategies and designing climate-responsive guidelines for urban environments. Since 2017, the project has brought together multi-disciplinary teams from partners like the SEC, National University of Singapore, TUM CREATE and Singapore-MIT Alliance for Research and Technology.

While the first phase of Cooling Singapore concentrated on selecting suitable tools to assess UHIs, the recently-kickstarted second phase will now put these tools into action by building an operational digital twin of Singapore. The platform, also known as Digital Urban Climate Twin or DUCT, will allow researchers and decision makers to explore various scenarios that contribute to urban heat.

Using the platform, users can simulate the UHI effect by calculating the impact of urban elements like buildings and roads on air temperature, then comparing these results with a scenario where all urban areas are replaced with vegetation. "You can also modify DUCT to run simulations for specific scenarios, like district cooling, [the presence] of solar panels or electric vehicles," explained Dr. Heiko Aydt, project leader at Cooling Singapore, in a webinar organized by NSCC in September 2020.

According to Aydt, DUCT is meant to be a decentralized, modular platform with multiple components—a federation of models. These include micro- to macro-scale models of the climate from the meteorological service as well as computational models of UHI factors like land surface, traffic and building energy. The platform will encompass outputs from A*STAR's IEM, which is used by the Cooling Singapore team as a microscale urban climate model. DUCT also provides risk and impact models of the urban climate on the economy, environment and health. All this advanced modeling entails crunching millions of data points at the same time—a job supercomputers are certainly well-suited for.

Though Cooling Singapore's second phase commenced only in September 2020, they've already implemented a few prototypes of the DUCT with in-house resources at the SEC. "In this case, we used an internal computing cluster with a rather modest 72 CPUs and a bit short of 200 gigabytes of RAM," shared Aydt. Eventually, the research team hopes to deploy DUCT across several organizations using multiple machines. For this, the Cooling Singapore team is looking to soon tap onto NSCC's computational resources to make DUCT's simulations available on demand to researchers and policymakers alike.

"We want to have a solution that is intuitive and useful for the end-users," concluded Aydt. "We plan to work closely with the NSCC and others to ensure that high performance computing resources can be tapped on in a seamless way."

THE SUPERCOMPUTER COOLING CONUNDRUM

While it's clear that supercomputers can help address Singapore's cooling woes, ironically, they're also part of the problem. With extreme data crunching comes extreme heat, and supercomputers can consume as much power as a small city. Though you'd think that most of the available power goes towards computing, a large portion actually goes into cooling the facilities and racks.

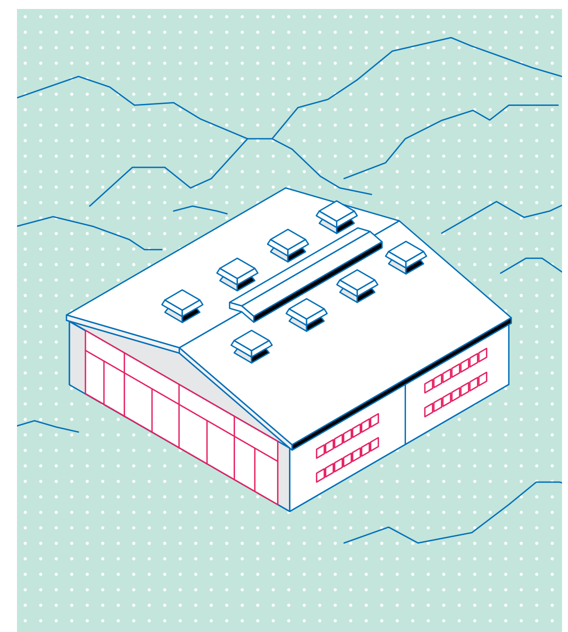
"I remember when I first saw Japan's K computer—once the world's fastest supercomputer—and the gigantic cooling plant right next to it," recalled Schmitt, speaking in the same NSCC-organized webinar. "I knew then that these supercomputer centers would create significant cooling loads for their surroundings."

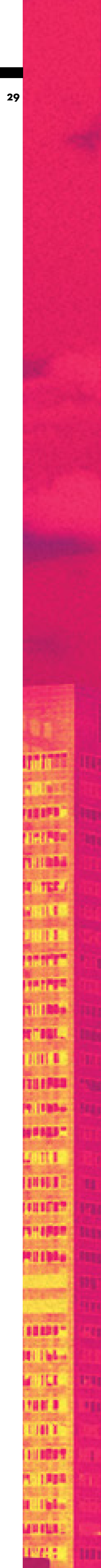
With more researchers turning to supercomputers to study climate change and its impacts, the annual electricity bills of their organizations have also skyrocketed to millions of dollars each year. "This is equivalent to tens of thousands of tons of carbon emissions per year, if fossil fuels are used in electricity production," explained Barker. For all the might of supercomputers, ensuring energy-efficiency is a major hurdle that needs to be overcome. "It's a challenge not only from a financial perspective, but also because climate science needs to lead the way in sustainable green computing," he added.

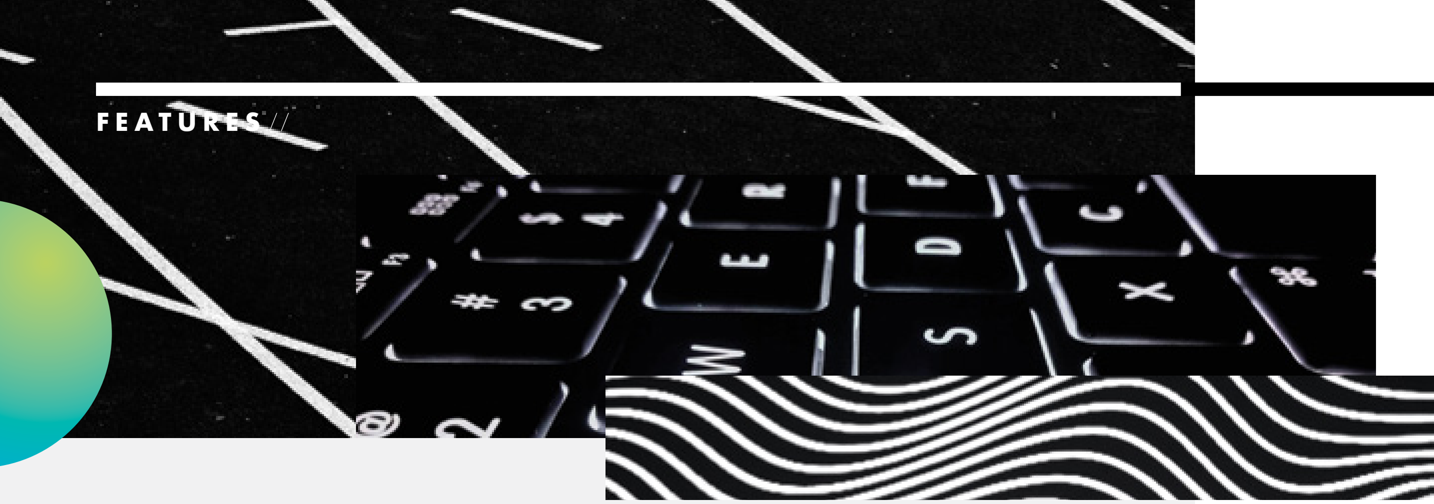
Accordingly, a few countries have tapped onto renewable energy to solve this supercomputer conundrum. A good example is Iceland, which has since become an international home for climate and weather computing due to its abundant supply of renewable energy and its naturally cold, temperate climate. Combined, these two factors reduce the amount of energy required to cool supercomputers in Iceland.

Beyond alternative energy sources, newer technologies are also being developed to reduce the overall electricity bill. Some examples include GPU-based processors, artificial intelligence-based systems for managing load and even recycling waste heat to warm offices. "Green computing in terms of FLOPS per watt is also increasingly part of the criteria for assessing bids when tendering for new climate supercomputers," said Barker.

The demand for supercomputing resources in climate science research is unlikely to ease in the upcoming decades. "More complexity, more local detail and even more scenarios mean that the supply will likely never satisfy the climate scientists' insatiable desire for computing resources," explained Barker. Still, supercomputers could play a crucial role in keeping Singapore cool—so long as we find an efficient way to keep them cool as well. S







OF VIRTUAL COMPETITIONS AND REAL RELATIONSHIPS

/ / / / /

>> Although the Student Cluster Competition went fully virtual in 2020, what remained unchanged was the importance of teamwork and community.

By **Sheryl Lee** and **Jill Arul**

From Monday to Friday, Liu Siyuan is a system administrator at a trading firm. His is the domain of mammoth machines that move data at lightning speeds. From software updates to hardware maintenance, he spends his days taking care of their every need.

But Liu is no ordinary system administrator. As captain of the Nanyang Technological University, Singapore (NTU) team that won the Student Cluster Competition (SCC) in 2017, he was once keeper of bigger beasts. For a few precious hours each weekend, he relives this experience as he trains a new team of juniors to take up the mantle. "When I first started, I didn't have a lot of useful materials and had to learn a lot of things the hard way," Liu recounted. "I want my juniors to get a good start."

The first SCC was held in 2007, but it wasn't until almost a decade later that NTU started sending teams to take part. By then, the competition had grown to involve more than a dozen teams from countries all over the world, including China, the US and Europe. In the school's second year participating, the team of six undergraduates led by Liu won the top prize.

FOR THE LOVE OF SUPERCOMPUTING

A Chinese national, Liu moved to Singapore in 2013 to study computer science at NTU. There, he found a home in the high performance computing (HPC) community, which was then a motley crew of just ten students. Though small, the informal club was dedicated to proving itself at competitions: team members would meet weekly during school term, and spend more than half their holidays preparing together. "A lot of my good friends now are from the team," Liu remarked. "I think we built up this relationship because we spent so much time together. Most of my university life I spent with this club."

Dr. Francis Lee, who revitalized NTU's informal HPC club when he took over as club advisor in 2016, said that the school's HPC community has been growing since then. "With our success in 2017, more people got to know about us and more people have joined," he said. "That's how our ecosystem works; we have enthusiastic students who want to learn, and seniors who are willing to come back to share."

Today, the club is a diverse community where students of different nationalities have come together to bond over a shared love of supercomputing. The latest SCC team, Supernova, is a mixed bunch hailing from Indonesia, Malaysia, China and Vietnam. "Even though we are from different countries, we all speak English," said team leader Li Shenggui, a fourth-year computer science student from China. "We're all from NTU and we're used to the way NTU students work together. We share a common culture."

> WHAT'S COOL ABOUT HPC IS THAT WE'RE ALWAYS STRIVING FOR BETTER PERFORMANCE.

For 2020's competition, which was held in November, Li partnered Aurelio Jethro Prahara, an Indonesian student in his second year at NTU with a background in data science and artificial intelligence. Together, they tackled the reproducibility application—one of four in the competition—which required reading a research paper on computed tomography reconstruction and replicating its findings under different settings and datasets.

EVERY CLOUD HAS ITS SILVER LINING

While the applications were similar to past years' challenges, the 2020 competition was the first to be conducted virtually. In the past, students had three months to build a cluster and ship it to a convention center, with cloud computing playing a minor role. This time, with the competition held online, all teams migrated completely to cloud-based HPC with only a month to prepare. Previous years' power caps were also replaced by budget limits in Microsoft Azure cloud—constraints the team's cloud manager Pang Jin Hui carefully navigated.

"It's a trade-off between performance and price," explained Pang, a third-year computer science student from Malaysia. "If you spend more, your performance will be very good. But you have a limited budget."

Though the learning curve was steep, cloud computing also had its perks. "In some sense, it's very easy because we don't need to do a lot of configuration," Pang mused. In previous years, existing libraries in the hardware and outdated software made it difficult to install new libraries bug-free. The cloud cluster, though, always started out empty, with relatively new versions of both hardware and software. And power could be ramped up easily where needed, allowing users to harness supercomputing power without investing in costly hardware.

"Supercomputing is going to be the norm pretty soon. But it's still got a long way to go," said Lee. "A lot of people don't know how to optimize; they think that as long as it runs, it's okay. I hope some of these students will go into the field like their seniors have, change that environment, and contribute to it."

ONE COMMUNITY, MANY NODES

With changes to the competition format and the intense time crunch, Liu's guidance has become even more valuable this year. On a particular Saturday afternoon before the competition, the team gathers in a Zoom room for yet another training session. All muted with video off, only their names appear on the black screen: Aurelio Jethro Prahara, Pang Jin Hui, Tan Jia Qing, Long Do Xuan, Dong Yunxing and Li Shenggui.

These meetings are not like the ones Liu remembers—there isn't much room for quality time spent together in-person during a pandemic—but the spirit and camaraderie of the club endure. After a 20-minute-long lesson on cloud software, the team settles down to discuss strategy, peppering their professor and seniors with questions on how much time and power they need to allocate to each section. As the afternoon wears on, the discussion meanders from budgeting to benchmark performance before veering off to analyzing specific parameters. It's dark when they finally disperse.

After two months of intense preparation, D-day arrived on November 11. Racing against 18 other teams working across six different time zones, team Supernova gave their all over the next 72 hours, completing the benchmarks, applications and the mystery application announced only at the start of the competition. Despite their best efforts, the team eventually lost out to Tsinghua University, which dominated the competition by coming out tops for all three benchmarks and most of the applications.

A slightly disappointing result perhaps, but Liu knows there'll be a next year, and he knows he'll be around. HPC is much more than winning competitions, he said. "What's cool about HPC is that we're always striving for better performance," Liu explained, his enthusiasm growing with each word. "That's what we do each competition, and that's what HPC centers are doing all over the world. They build very complex machines and try to accelerate all kinds of scientific discovery. Getting as much performance as we can with what we have—that's what I really enjoy."

Ultimately, Liu concluded, HPC is about a journey to constantly outdo yourself in the face of constraints. Along the way, you impart knowledge to newcomers and build community. And as your cluster of nodes grows, so does your collective power.

THE SHOW MUST GO ON

Highlights from a completely virtual SC20

SC20 saw the high performance computing community coming together from across the globe for ten days of talks, exhibitions and competitions—all from the comfort of their own homes.

By **Jill Arul**

From left to right: Rommie Amaro; Alessandro Vespignani; Ilkay Altintas; and Rick Stevens.

For the last 30 years, the supercomputing community across the globe has gathered annually for the International Conference for High Performance Computing, Networking, Storage and Analysis, better known as SuperComputing or simply SC. For the first time since its inception, the 2020 edition of SC was held online, spanning ten days from November 9–19.

Though the students, researchers and industry professionals were not in the same physical space this year, moving SC online provided more opportunities for people to participate in other ways. Jam-packed with talks, panels, exhibition booths and workshops as with previous years, over 600 hours of content is now available online for anyone who wishes to keep up with the latest in the world of high performance computing (HPC).

Centered around the theme "Everywhere We Are," SC20 celebrated the collaborative spirit of HPC and underscored the power of that collaboration when faced with challenges like COVID-19.

COVID-DRIVEN COLLABORATION

Plagues, epidemics and pandemics have always been watersheds throughout history and COVID-19 is no exception. Instead of sparking revolution like the third bubonic plague did in India 165 years ago, or changing global healthcare infrastructure as SARS did in 2003, COVID-19 has served as a catalyst pushing us even further into the digital age and cementing our reliance on cloud, telecommunications and supercomputing technologies.

In fact, supercomputing has played a prominent role in accelerating our understanding of the new disease, said panelists speaking at the "More Than HPC" plenary session on COVID-19. From biological simulation to epidemiological modeling, supercomputers have been central to planning our defenses against both this and future pandemics, said Professors Rommie Amaro and Alessandro Vespignani of the University of California, San Diego, and Northeastern University respectively.

Amaro, for instance, is building an all-atoms model of the SARS-CoV-2 virus on the Frontera supercomputer, using its immense power to visualize the impact of glycans, a carbohydrate shield around viral proteins that helps them evade the human immune system. "There is a whole community of researchers that is actively investigating the molecular pieces of the virus and we've committed to not only sharing our methods but also our data and models so that we can collectively move forward as quickly as possible," she said.

In particular, Amaro highlighted how her work greatly benefitted from the COVID-19 HPC consortium, a collaboration between 37 organizations from the US as well as the Korea Institute of Science and Technology Information and the RIKEN Center for Computational Science in Japan to combine computing resources, services and skills.

But Vespignani, an expert in epidemiological modeling and prediction, offered a less optimistic counter point—this collaboration might be temporary. "Everybody wants to contribute because we are now in war time, everybody wants to do something relevant for the effort," he explained. "In war time it is easier to have a call to arms, but when we enter peace time it may not be so."

Stressing the importance of preparation, Vespignani emphasized that it is crucial to prepare now for future health disasters while the population are on their toes. This spirit of cooperation is integral to building resilient and vigilant institutions that will last throughout periods of peace and are ready to protect against future disasters, he said.

> **"In war time it is easier to have a call to arms, but when we enter peace time it may not be so."**
>
> Professor Alessandro Vespignani
> *Northeastern University*

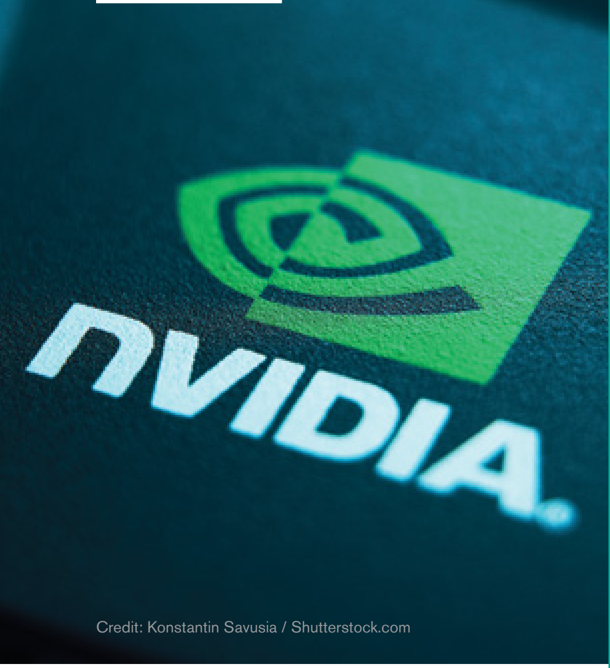
LOOKING TO THE FUTURE

However, if the participation rates of the annual student cluster competition are anything to go by, the future looks bright. Featuring a record 19 teams fighting for top honors, the competition saw participants adapt to a new format, running their applications not on custom hardware but on the cloud. This year, the Tsinghua University team came out tops, reaching a peak performance of 300 teraFLOPS.

Clinching another coveted prize was a nine-member team led by Jia Weile of the University of California, Berkley, which was awarded the 2020 Gordon Bell Prize for their molecular dynamics study of over 100 million atoms. A special award—the Gordon Bell Special Prize for HPC-Based COVID-19 Research—was also given out to a 12-member team including Amaro, for their work on using AI to illuminate SARS-CoV-2 spike dynamics.

Whether the topic was on COVID-19 research or something else, a common thread throughout the conference was the power of HPC when put in the hands of a collaborative community. Though being in the physical presence of the HPC community was sorely missed, work being done across the same community is helping to ensure that a physical SC21—and a better world—is a possibility in the year ahead.

Business Bytes



Credit: Konstantin Savusia / Shutterstock.com

AN ARM (AND A LEG): NVIDIA'S US$40B ACQUISITION

US-based chipmaker NVIDIA Corp. will be acquiring chip design firm Arm Ltd. for US$40 billion in what would be the semiconductor industry's largest-ever deal, if approved. The powerful team-up brings together NVIDIA's artificial intelligence (AI) capabilities and Arm's ubiquitous chip technology to accelerate innovation and expand into new markets.

Previously owned by Softbank Group Corp., Arm's revenue primarily comes from licensing its chip technology and selling processor designs. According to NVIDIA CEO Mr. Jensen Huang, a key priority will be investing in Arm's efforts to design chips for data-center computing. While the company has carved out a niche in supplying AI graphic processing units (GPUs) to companies like Google and Facebook, ARM-based central processing units (CPUs) could help it to gain some of Intel's 90 percent CPU market share.

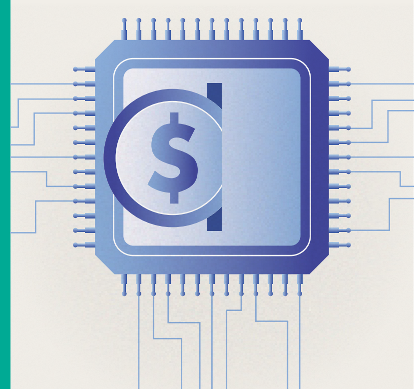
"Uniting NVIDIA's AI computing capabilities with the vast ecosystem of Arm's CPU, we can advance computing from the cloud, smartphones, PCs, self-driving cars and robotics, to edge Internet of Things, and expand AI computing to every corner of the globe," said Huang.

ADVANCED MICRO DEVICES TO BUY RIVAL CHIPMAKER XILINX

Semiconductor designer Advanced Micro Devices Inc. (AMD) is buying fellow chipmaker Xilinx Inc. for US$35 billion in stock, representing the firm's largest acquisition ever. The deal will add a wider range of products to AMD's offerings to take on industry leader Intel Corp in the data center chip market.

Known for inventing a user-configurable integrated circuit called the field-programmable gate array, Xilinx's products are used in various industries ranging from consumer electronics to research and development. By acquiring Xilinx, AMD is set to offer the industry's strongest portfolio of high performance processor technologies for applications in cloud, edge as well as end devices. Together, the combined company will explore opportunities in high-growth segments like data centers and defense among others.

"Our acquisition of Xilinx marks the next leg in our journey to establish AMD as the industry's HPC leader and partner of choice for the largest and most important technology companies in the world," said AMD President and CEO Dr. Lisa Su. While the acquisition still needs to be approved by shareholders and regulators, including authorities in China, it is expected to close by the end of 2021.



CONFIDENTIAL COMPUTING IN THE CLOUD

IT giant International Business Machines Corp. (IBM) and semiconductor firm Advanced Micro Devices Inc. (AMD) will work together to leverage open-source frameworks and drive confidential computing in hybrid cloud and high performance computing (HPC). The collaboration aims to enhance security and AI products for both companies, the firms said when they announced their joint development agreement in November 2020.

Confidential computing is a technology that allows sensitive data to be encrypted even as a virtual machine is processing it. This helps prevent hackers from accessing private information even if they were to break in.

Currently, cybersecurity is the top barrier to cloud adoption in companies, according to IBM data. Combining confidential computing with hybrid cloud would be the first step towards enabling regulated industries such as finance, healthcare and insurance to benefit from cloud technology.

"The commitment of AMD to technological innovation aligns with our mission to develop and accelerate the adoption of the hybrid cloud to help connect, secure and power our digital world," said Mr. Dario Gil, director of IBM Research.

MEET MN-3, THE WORLD'S MOST ENERGY-EFFICIENT SUPERCOMPUTER

The MN-3 supercomputer has been ranked by the Green500 as the world's most energy-efficient supercomputer for 2020. Built by Japanese artificial intelligence start-up Preferred Networks Inc. (PFN) and Kobe University, the system is powered by MN-Core™, a highly-efficient custom processor developed for deep learning applications.

Launched in May 2020, MN-3 delivers a whopping 26 gigaFLOPS per watt, for a total performance of 1.62 LINPACK petaFLOPS. This means that the supercomputer performs 26 billion calculations for every watt of power consumed in a single second. This achievement is 15 percent more efficient than the 2018 Green500 record of 18 gigaFLOPS per watt, previously held by RIKEN.

MN-3's milestone was partly achieved by densely integrating multiple MN-Core dies onto each board. Despite already topping the Green500, there's still more work to do on MN-3. PFN plans to further increase the supercomputer's energy efficiency by improving installation methods, cooling and MN-Core™ specific middleware.

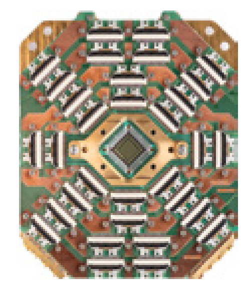
HELPING BUSINESSES TAKE ADVANTAGE OF QUANTUM COMPUTING

Canadian quantum computing company D-Wave Systems Inc. has announced the availability of its new quantum computer for use by businesses. Known as Advantage™, the system features 5,000 qubits and a 15-way qubit connectivity. It will be made available to customers through a quantum cloud service.

Advantage™ is powered by annealer technology, which means that its qubits are cooled during the execution of an algorithm. This allows for the qubit's value to be passively changed. In contrast, Google and IBM's quantum computers use a gate-model approach, in which qubits are modified as an algorithm is executed.

The system is already being used by several customers for a wide range of use cases. Global multinationals like Accenture and Volkswagen are using Advantage to develop and optimize business applications. Meanwhile, synthetic biology company Menten AI is harnessing the system to design new proteins.

"This launch continues what we've known at D-Wave for a long time: it's not about hype, it's about scaling, and delivering systems that provide real business value on real business applications," said Mr. Alan Baratz, CEO of D-Wave.



The D-Wave Advantage chip holder and 5,000 qubit chip.

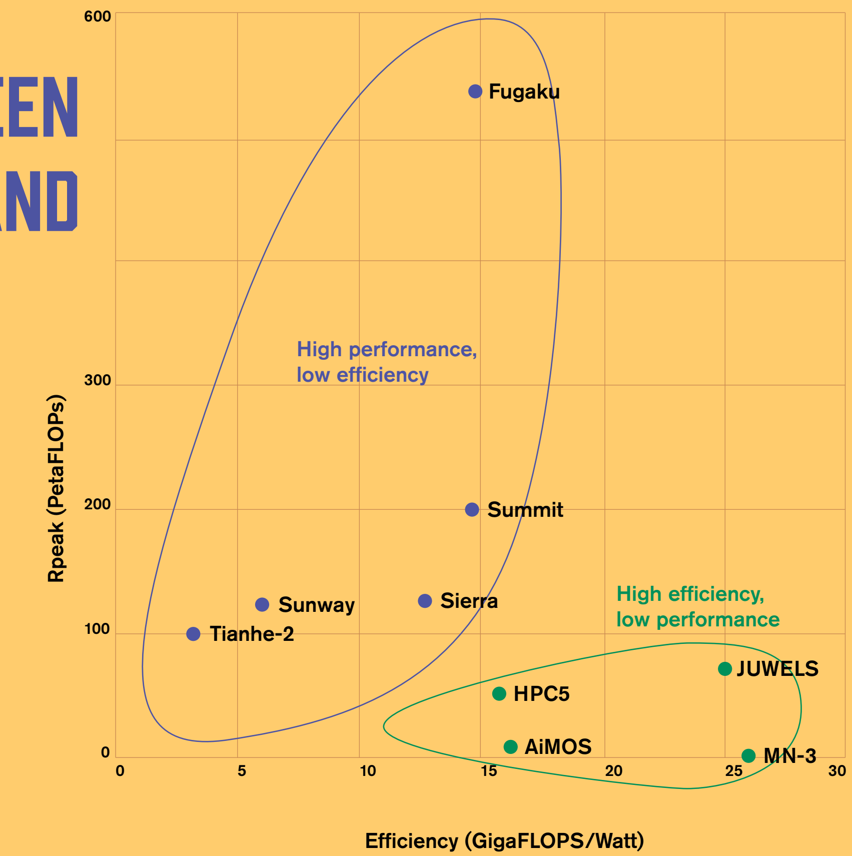
Credit: D-Wave Systems Inc.

Super Snapshot

STRIKING A BALANCE BETWEEN PERFORMANCE AND EFFICIENCY

THE TRADE-OFF BETWEEN PERFORMANCE AND POWER CONSUMPTION

Machines on the TOP500 (blue) pack in the FLOPS, but typically do so at the expense of energy efficiency. Conversely, the top supercomputers on the Green500 (green) list are very efficient, but are typically lower powered machines.



FLIPPING THE SCRIPT

In recent years, researchers have been looking for innovative ways to develop energy efficient supercomputers that will not sacrifice speed. Here are some of the most promising technologies that are moving supercomputing towards a greener future.

Directing data movement

Roughly 28–40 percent of total energy cost is spent moving data and much of that energy is wasted or leaked. Prioritizing the movement of data relative to computing power can help to reduce power leakage for more efficient computing.

Keeping cool

About half of the energy costs of running supercomputers goes into cooling. Oil immersion, used by the Green500 topping TSUBAME-KFC, is one innovative approach to lowering the cost of keeping supercomputers cool.

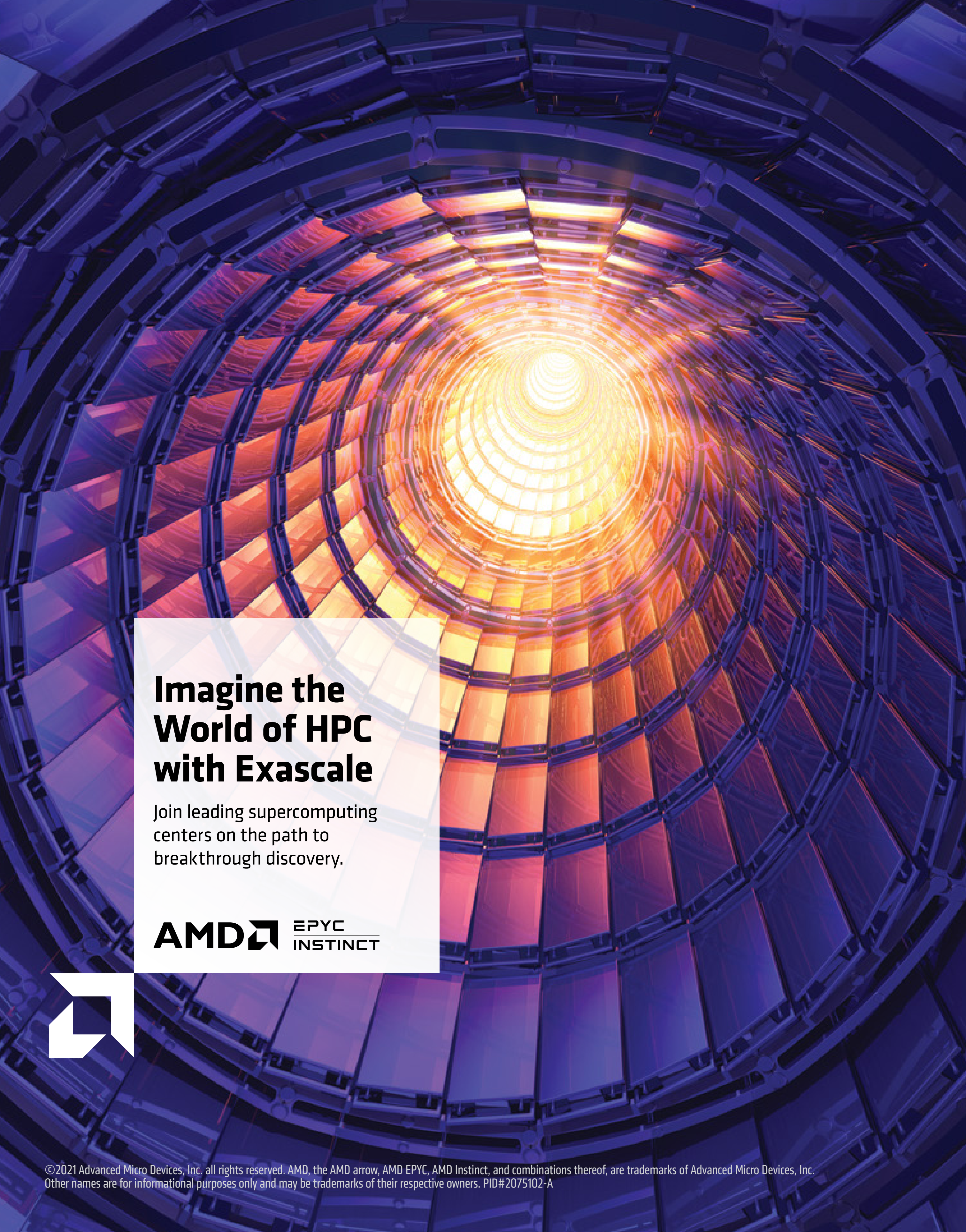
Adjusting power priorities

By adjusting the power supply depending on computing load and thermal requirements, power management systems like the one jointly developed by ABB and HPC can reduce energy usage by up to 30 percent.

Renewable energy

Europe's Large Unified Modern Infrastructure (LUMI) will be completely carbon neutral—running on renewable hydroelectric energy and harnessing waste heat to provide heat to the local town.

1 Kestor, G., Gioiosa, R., Kerbyson, D. J., & Hoisie, A. (2013). *Quantifying the energy cost of data movement in scientific applications. 2013 IEEE International Symposium on Workload Characterization* (IISWC). doi:10.1109/iiswc.2013.6704670


Imagine the World of HPC with Exascale
Join leading supercomputing centers on the path to breakthrough discovery.
AMD
EPYC
INSTINCT