**SICOS BW GmbH**

High-performance computing and data analytics to combat coronavirus

In an epidemic, data exchange and the use of innovative technologies are essential to be able to act effectively and quickly against the infection. High-performance computers and data analytics make a valuable contribution to this. SICOS BW is a promising solution centre that bridges the gap from numerical simulation, big data and AI to companies in the medical technology sector.

In the case of the coronavirus disease COVID-19, there are several starting points for new technologies: it is possible to unravel the structure of the new SARS-CoV-2 coronavirus with the help of simulation and supercomputers and thus develop appropriate therapeutic approaches. Data and technologies can also be used to improve public safety.

Opening the door to high performance systems
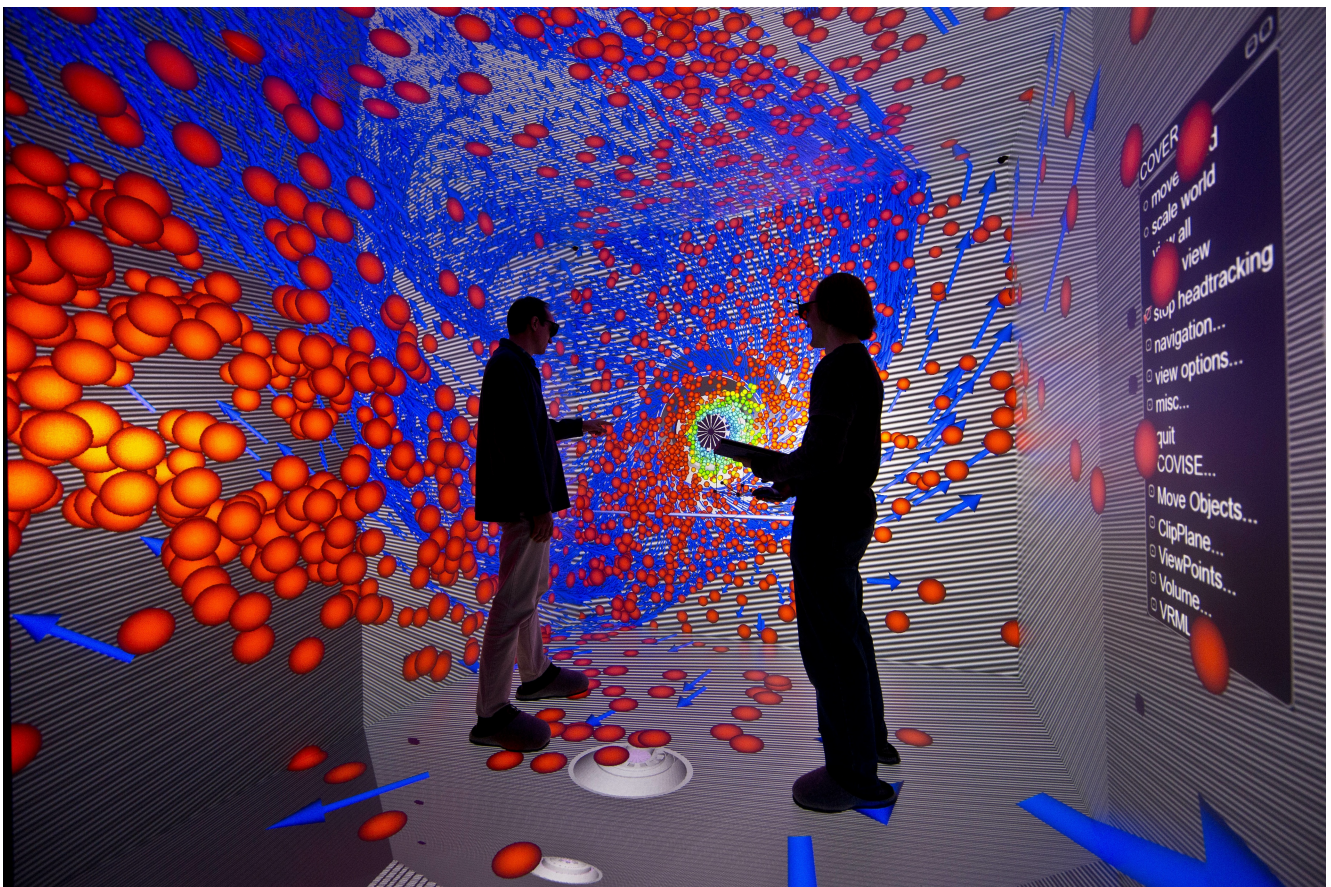
The HLRS (High-Performance Computing Centre) uses the latest virtual reality technology in its 3D CAVE to analyse the complex calculations. © HLRS

"We want to make it easier for companies, institutes and software manufacturers to be able to access high-performance computing and data analytics," explains Andreas Wierse, head of SICOS BW GmbH. "Companies need to continuously achieve top performance in order to process new data effectively and remain competitive. We act as an interface for our clients between science and industry. SICOS BW is a GmbH, a limited liability company, but is not commercially aligned. The financial

support that it receives from its two shareholders, the Karlsruhe Institute of Technology (KIT) and the University of Stuttgart, as well as funding from the Baden-Württemberg Ministry of Science, Research and the Arts is therefore essential to ensure that smaller companies also have good access to big data, smart data or are able to perform simulations," says Wiese. With their two data centres, the Steinbuch Centre for Computing (SCC) and the High-Performance Computing Centre Stuttgart (HLRS), the KIT and the University of Stuttgart provide the technology for all kinds of conceivable applications in the fields of mechanical engineering, climate simulation or materials sciences.

Large-scale calculations - targeting corona docking proteins

High-performance computing is already being used innovatively in biotechnology, for example in computer-aided drug design. Wiese, a mathematician with a PhD in engineering, knows exactly what industry requires and what is possible in science: "High-performance computing is used extensively in the field of numerical simulation. We offer the associated IT infrastructure and are looking for methodological and technological tools to effectively manage the large quantities of data."

It is a lengthy process between the discovery of a new active ingredient to production. However, speed is what is currently required. A high processing performance of many petaflops - i.e. a number with 15 zeros – can, for example, be applied in the search for the point of entry of the new coronavirus into the host cell. Flops, floating point operations per second, are a measure of computer performance indicating how many arithmetic operations can be performed per second.

The coronavirus uses the spike protein to attach to the ACE2 receptor on the host cell surface and enter human cells. The representation of the atomic structure of the spike protein provides a lot of data that has to be recorded and processed precisely. After all, it could be the key to effective antiviral medicines. The virus is very sneaky: although its surface structure initially appears "inconspicuous", the conformation of the spike protein changes shortly before the virus enters the cell, in order to create a shape that fits into the cell receptor[1,2]. This is where specific simulation models can provide precise and valuable support. The specific cleavage site of the spike protein, which has not yet been found in related viruses, can also be included in the calculations[3].

Once the attack surface has been exposed in detail, calculations can begin: high-performance computers can be used to simulate how the new coronavirus behaves under the influence of thousands of active ingredients. The world's fastest supercomputer, Summit (OLCF-4), which has 10,000 IBM processors, is currently working on this at the Oak Ridge National Laboratory[4] in the United States. It is being assisted by the Sierra supercomputer at the Lawrence Livermore National Laboratory in California. It took just eight days to analyse around 8,000 drug candidates. 77 of these candidates are seen as promising in that they may weaken or completely prevent the virus' docking process. This analysis would have taken months with a conventional computer, and longer still with laboratory tests. The data from the 77 substances are now being supplemented with current research data on the virus, tested again and the most promising candidates are then being tested in the laboratory.

Europe's fastest computer is ready for use

The HLRS works in cooperation with supercomputing facilities worldwide to perform these analyses. Folding@home[5] is a distributed computing project focused on the dynamics of proteins implicated in a variety of diseases and brings together many scientists around the world. What contribution can SICOS make here? "Our computing centres support complex calculations with their computing resources, and our university partners provide their analytics expertise and infrastructure." Wiese has over 25 years of expertise in simulation and visualisation, and is supported in his work by his eight staff as well as other companies. Hawk, the HLRS supercomputer flagship, has been optimised for technical simulation applications. In combination with other high-performance data analysis and artificial intelligence platforms, Europe's fastest computer (26 petaflops) supports new work processes, simulation, data generation and analysis, in combination with deep learning.

Efficient countermeasures through pattern recognition and contact tracing apps

Hawk, the new supercomputer at the High-Performance Computing Centre (HLRS) in Stuttgart has over 720,000 cores and is almost four times faster than its predecessor.
© HLRS

It is important to fully identify and interpret the infection process in order to be able to deal effectively and quickly with the pandemic. How does the virus spread? Which disease courses are likely in which patient groups? Data analytics and machine learning can be used to specifically identify patterns. Typical lung infection patterns of people who have been infected can also be identified more quickly.

Until vaccines have been developed and approved, we have to live with certain precautions in order to avoid a renewed wave of infections. A contact tracing app is one such precautionary measure. Tracing apps work with Bluetooth technology. A new identification number is created every few minutes and sent to the surrounding area. If two smartphones with this app are less than two metres apart over a period of 15 minutes, the two anonymous identification numbers are saved on both phones. If users of a tracing app subsequently show symptoms of the disease, they can report this via the app, which replaces, and is quicker than, standard health authority procedures. The app can be used to send a message to people with whom the infected individual has been in contact, requesting them to get tested. The identity and location of the individuals concerned will not be recorded. For data protection reasons, the German government has agreed to decentralised storage of these data, i.e. the data will only be stored on users' smartphones. Other technologies such as thermal cameras or sensors in the area of the Internet of things (IoT) can also help in the early detection and containment of further outbreaks of the infection.

There are therefore many ways in which parameters that can be used for the simulations can accumulate. Supercomputers can also be used to simulate analyses, such as how infections can spread on board aircraft. It's worth noting here that SICOS BW also offers many simulations and calculations for the aerospace industry.

There are many areas of application where HCP can be used. In the current corona pandemic, however, accelerated access to HPC resources enables data from complex calculations and forecasts as well as simulation research for developing medicines and vaccines to be obtained quickly with the aim of containing the infection process, now and in the future. Because one thing is clear: we will continue to face many epidemics. These technologies can be a crucial guide for fast, targeted countermeasures.

References

[1] Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface

(https://chemrxiv.org/articles/Repurposing_Therapeutics_for_the_Wuhan_Coronavirus_nCov-2019_Supercomputer-Based_Docking_to_the_Viral_S_Protein_and_Human_ACE2_Interface/11871402)

[2] Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation: https://science.sciencemag.org/content/367/6483/1260.full

[3] https://www.dpz.eu/de/infothek/wissen/coronaviren.html

[4] https://www.olcf.ornl.gov/2020/03/05/ornl-team-enlists-worlds-fastest-supercomputer-to-combat-the-coronavirus/

[5] https://foldingathome.org/2020/02/27/foldinghome-takes-up-the-fight-against-covid-19-2019-ncov/

Article

18-Jun-2020
Simone Giesler
© BIOPRO Baden-Württemberg GmbH

Further information

SICOS BW GmbH
Dr. Ing. Andreas Wierse, CEO
Phone: +49 (0)711 217 2828-0

▸ SICOS BW GmbH

The article is part of the following dossiers

Data mining: new opportunities for medicine and public health

Telemedicine: advances in high-tech healthcare

Big data - the big promise of the new digitised world

Artificial intelligence in medicine: assistance for human senses

The quantum revolution in the healthcare industry

simulation | University of Stuttgart | KIT | data mining | drug discovery | big data | artificial intelligence

innovation | SARS-CoV-2/COVID-19