HPC in South Africa

The Centre for High Performance Computing delivers leading-edge compute and storage resources to the African research community.

Customer profile

CHPC
CENTRE FOR HIGH PERFORMANCE COMPUTING

Scientific Research | South Africa

"This is a national facility that keeps growing. We provide resources to researchers in academia and industry, and we are working to provide the same to government departments that can benefit from HPC."

Happy Sithole
CHPC Director

Business needs

The Centre for High Performance Computing needs massive compute, storage and networking systems to support thousands of researchers in South Africa and points beyond.

Business results

- Putting the power of supercomputing into the hands of scientists and engineers
- Accelerating discovery and innovation across the African continent
- Addressing critical issues — from COVID-19 to climate change
- Making compute and storage resources readily accessible via a private cloud

Solutions at a glance

- Dell EMC PowerEdge servers with NVIDIA® GPUs
- NVIDIA InfiniBand® networking and high-speed Ethernet networking
- Dell EMC storage with the Lustre parallel file system
- OpenStack® private cloud

Proving HPC resources and expertise

The Centre for High Performance Computing in South Africa (CHCP) provides HPC resources and domain-specific support for large-scale science projects to public and private sector users in the South African research community. In addition, the Centre extends its resources and expertise to other African countries, including those in the Southern African Development Community and those that participate in the Square Kilometer Array project, an international effort to build the world's largest radio telescope.

A service-oriented organization, CHPC helps research teams optimize their applications for top performance on the HPC systems they use for computational fluid dynamics, material science, chemistry, climate modeling and other data-intensive work. In these undertakings, CHCP works to build HPC expertise among a community of users that extends across the African continent.

"This is a national facility that keeps growing," says Dr. Happy Sithole, CHPC director and Centre manager for National Integrated Cyber Infrastructure in South Africa. "We provide resources to researchers in academia and industry, and we are working to provide the same to government departments that can benefit from HPC."

HPC at the speed of a Cheetah

In all its work, CHPC seeks to ensure that the researchers it serves have access to the latest and greatest systems for high performance computing. That was the case in 2016 when the Centre introduced a peta-scale supercomputer with approximately 33,000 compute cores. The machine, delivered by Dell Technologies, debuted as the fastest supercomputer in Africa and the 121st fastest computer in the world, per TOP500 rankings.

Named Lengau, for the Setswana word for Cheetah, this groundbreaking system went into production with more than 1,000 Dell EMC PowerEdge servers and 5 petabytes of Dell EMC storage. And in 2018, CHPC expanded its compute resources to include a GPU cluster consisting of 30 state-of-the-art NVIDIA V100 GPUs.

Networking is a very big piece of the HPC puzzle for CHPC, because fast application performance depends on fast I/O speeds. To meet this need, the Lengau system uses InfiniBand networking from NVIDIA, along with high-speed Ethernet networking.

"Our big cluster is based on InfiniBand," Sithole says. "This is very critical because we always evaluate our applications on how they perform based on the various network architectures. We have seen that our applications are sensitive to access to storage, so we always optimize the design of our cluster on the fabric we deploy. That's one of the things we focus on first when it comes to optimization."

A cloud connection

After joining the Square Kilometer Array project, CHPC began development work on a converged cloud and HPC data center infrastructure to support the automated orchestration of compute and storage resources and make HPC systems accessible to more users. This private cloud was based on OpenStack and OpenStack CEPH storage software. And then the COVID-19 pandemic hit.

"Our big cluster is based on InfiniBand. This is very critical because we always evaluate our applications on how they perform based on the various network architectures."

Happy Sithole
CHPC Director

> “Today, we are providing services to many researchers via the OpenStack cloud. And those services have been mainly supporting the country during the pandemic.”
>
> **Happy Sithole**
>
> CHPC Director

Just days after the CHPC OpenStack Production Cloud was put into service, South Africa went into lockdown and demand for cloud-based computing resources soared. The new cloud was soon overwhelmed, prompting CHPC to focus on a major expansion of the environment. For this upgrade, CHPC worked with Dell Technologies and to deploy additional compute and storage resources to create a larger and more robust OpenStack Production Cloud.

With this expansion, the cloud environment added 15 new compute nodes using Intel-based Dell EMC PowerEdge servers, 26 new storage nodes, 60 terabytes of hot data storage using SSD DC drives and 480 terabytes of additional storage.

“Today, we are providing services to many researchers via the OpenStack cloud,” Sithole says. “And those services have been mainly supporting the country during the pandemic. The cloud services have been very helpful during this time. The research community is sharing a great deal of data sharing via the cloud.”

In another benefit, the expanded CHPC OpenStack Production Cloud gives users access to a rich menu of tools for data-intensive research, including Intel artificial intelligence technologies and machine learning libraries. Sithole notes that these tools are especially useful to researchers in the medical field.

Diverse use cases

The value offered by CHPC comes to light most clearly in the diverse research projects the center supports. The center currently supports 26 universities, more than 900 research programs and 2,000-plus researchers who use the Lengau system, which runs close to 100 percent of the time.

Workloads running on the system are as diverse as the range of scientific disciplines. Researchers use Lengau for climate modeling and long-term weather forecasting, engineering and computational fluid dynamics, material sciences and energy storage systems, bioinformatics and precision medicine, and precision agriculture, among other research domains.

In addition, the Lengau supercomputing has emerged as a powerful tool for use in DNA sequencing, molecular modeling and vaccine development associated with the COVID-19 pandemic.

“This COVID-19 research has been very helpful in looking at the various variants through mutations,” Sithole says. “We managed to track down different mutation mechanisms and show how the variants react to various vaccines. This research has yielded very important information that guided, in terms of decision making, which vaccines should come to use.”

The Centre for High Performance Computing in South Africa is among the ranks of the Dell Technologies HPC & AI Centers of Excellence. These distinguished centers, found throughout the world, provide thought leadership, test new technologies and share best practices with the community.

Another area of research that makes heavy use of CHPC's resources is climate modeling. Researchers working in this field are exploring new weather models that shed light on the specific climatic conditions and weather events in Africa and the effects of climate change such as agricultural production.

"In climate modeling, we have to simulate various scenarios," Sithole says. "And this is where things like machine learning and artificial intelligence come in to play in making predictions. This is one of the big research projects that we are doing and looking at — how we can use machine learning and artificial intelligence to scan through various combinations for climate modeling."

Ultimately, for all of the research conducted with the assistance of CHPC, high performance computing is an essential tool. And that's why the CHPC team works continuously to keep its systems at the leading edge of technology — from high-performance compute and storage systems from Dell Technologies to GPU acceleration and lightning-fast networking from NVIDIA.