Enabling early detection and prevention of genetic diseases

Center of Excellence for Medical Genomics

To advance genetic, screening, diagnosis, and precision medicine, CEMG expanded its genome sequencing infrastructure with Lenovo ThinkSystem SR950 V3 servers powered by 4th Gen Intel® Xeon® Scalable processors.





Customer background

Who is CEMG?

The Center of Excellence for Medical Genomics (CEMG) is part of the Faculty of Medicine at Chulalongkorn University—Thailand's oldest and most prestigious university. Established in 2009, CEMG provides comprehensive clinical genomic services for patients with genetic diseases, creates new knowledge and generates innovations in genomics and precision medicine, and provides training in these fields. Its mission is to enable equitable access to high-quality medical diagnoses and effective therapies with fewer side effects, helping to prevent disease and promote better public health.





The challenge

Genomics, the study of an organism's complete DNA sequence, has revolutionized healthcare in recent decades. Decoding an individual's genetic information enables faster and more accurate diagnosis and more personalized treatments, and can be used to predict and help prevent future conditions.

CEMG recently launched one of the world's first projects to perform long-read whole genome sequencing (WGS) in newborns. The initiative screens for 113 genetic diseases, with the goal of early prevention and treatment, opening a new frontier in precision medicine and improving lifelong health outcomes.

The challenge

"Why wait until a person is sick to check their genetic information? We can screen newborns and prevent symptoms before they appear," says Prof. Dr. Vorasuk Shotelersuk, Director of the CEMG and Deputy Dean for Research at the Faculty of Medicine, Chulalongkorn University. "Because our genes stay the same, we can detect disease risk from birth and intervene early."

The newborn genome sequencing project uses long-read WGS to determine if a child is at risk of serious genetic diseases. If abnormalities are found, children receive early, targeted prevention and treatment. Around 300 participants are expected in the first year, which will require CEMG to analyze huge volumes of data.

The challenge

"In recent years, genomic data has seen exponential growth," says Piyakrit Wongboonchai, a bioinformatician at CEMG. "Previously, whole-genome files were approximately 100 GB. Now, with the advent of long-read sequencing technologies, whole-genome datasets have increased ten-fold. This means that more computational power is required to process larger, more complex data."

CEMG looked to scale out its Lenovo high-performance computing (HPC) infrastructure to increase both compute performance and storage capacity.



"Genetic diseases can be detected presymptomatically, with mutations present from day zero. Symptoms may appear at birth, age 2, or even age 40. The project's goal is to detect them early to prevent symptoms."

Prof. Dr. Vorasuk Shotelersuk

Director, Center of Excellence for Medical Genomics

3 The solution

Infrastructure expansion

After consulting with Lenovo, CEMG decided to scale out its existing HPC cluster with nine Lenovo ThinkSystem SR950 V3 servers, each powered by eight 4th Gen Intel® Xeon® Scalable processors.

As well as meeting the organization's requirements for high performance, memory capacity, and bandwidth, each Lenovo ThinkSystem SR950 V3 server features 16 PCIe slots—critical for future expansion. "Scalability is crucial. With a well-designed system, the genome sequencing project could potentially expand nationwide in the future," notes Dr. Wanna Chetruengchai, Head of Bioinformatics at CEMG.

Hardware

Lenovo ThinkSystem SR950 V3
Lenovo ThinkSystem SR850 V3
Lenovo ThinkSystem SR670 V2
Lenovo ThinkSystem SR650 V3
4th Gen Intel® Xeon® Scalable
processors

Software

<u>Lenovo XClarity Administrator</u> IBM Storage Scale

Services

Lenovo HPC Services – Cluster Design, Configuration, and Installation

3 The solution

Leveraging IBM Storage Scale, the Lenovo ThinkSystem servers provide a software-defined storage solution, supporting genome data growth with more than 10.4 PiB capacity and the ability to scale out without limit.

The solution

Expert advice and support

CEMG worked closely with Lenovo during the planning stage to design an expansion cluster capable of handling very large genomic datasets.

Dr. Chetruengchai recalls: "We routinely transfer terabytes of genomic data from our laboratory to our HPC infrastructure. To support this, we needed a high-bandwidth InfiniBand network to minimize transfer times and prevent I/O bottlenecking. Because we operate multiple sequencing services, we needed a highly parallelized system to handle these tasks. The experts from Lenovo took all these requirements on board and designed a solution perfectly optimized for genomic services and research."

Piyakrit Wongboonchai adds: "The 4th Gen Intel® Xeon® Scalable processors deliver excellent performance and efficiency. Several algorithms in our computational pipeline are optimized for Intel® processors, making software compatibility another key factor in our decision."



"Lenovo were by our side at every step of the process, from delivery to implementation. They provided lots of valuable advice and support, and designed a solution that met all our technical requirements."

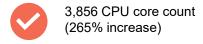
Dr. Wanna Chetruengchai

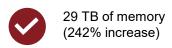
Head of Bioinformatics, Center of Excellence for Medical Genomics

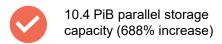
The results

The expansion cluster gives CEMG significantly more compute power and more storage capacity to support the genome sequencing project, enabling users to run larger datasets faster and more efficiently.

"The fat node, high CPU core count, and large memory capacity of the Lenovo ThinkSystem SR950 V3 servers give us a major performance boost," confirms Dr. Chetruengchai. "The Lenovo ThinkSystem D4390 provides a high-capacity, parallelized storage system suitable for large genomic datasets and enables fast, stable data transfers."







The results

Improving health through genomics

Scaling out its HPC infrastructure with Lenovo and Intel® technology gives CEMG the resources it needs to ramp up the newborn sequencing project. The organization hopes to expand the project nationwide, giving all Thai children access to their genetic blueprint from birth.

Prof. Dr. Shotelersuk comments: "Knowing one's genome will be the foundation of future medicine—not only for prevention and treatment, but eventually for gene editing and gene therapy. For example, knowing a high risk for colorectal cancer enables prevention and reduces mortality. If an individual is at risk for breast cancer, early preventive measures can be taken."



"One day, I believe every child will be born with their genome sequenced."

Prof. Dr. Vorasuk Shotelersuk

Director, Center of Excellence for Medical Genomics

Why Lenovo?

A long-time user of Lenovo HPC technology, CEMG engaged the Lenovo team for helping scaling out its existing infrastructure.

"Our experience working with Lenovo has always been very positive," says Dr. Chetruengchai. "For this expansion, the Lenovo team showed great flexibility in configuring the hardware to match our projected workloads and budget."

Looking ahead, CEMG is planning to expand the HPC cluster once again to support upcoming Al workloads. "We expect Lenovo to remain a close partner for many years to come," concludes Dr. Chetruengchai.

How do you improve the health of a nation?

Backed by Lenovo and Intel® technology, CEMG is screening newborns for genetic diseases for better detection, prevention, and treatment. That's the power of Lenovo with Intel Inside®.

Explore Lenovo HPC Solutions

