KDRI Uses New SGI UV 1000 for Genome Sequencing and Analysis

System Helps Researchers Successfully Shorten Time for Analysis

Kazusa DNA Research Institute (KDRI), located in Chiba, Japan, has approximately 60 researchers among a total employee count of approximately 170, as of April 2012. Founded in 1991, KDRI recently purchased an SGI UV 1000 coherent shared memory high performance computing server.

System Configured to Meet Their Needs

The SGI UV platform, though based on standard hardware, can be configured to meet the individual needs of the hundreds of organizations that have purchased the platform since its launch in 2008, in areas as diverse as science, research, weather, manufacturing and more.

In June 2011, KDRI purchased a UV 1000 with 216 cores of Intel® Xeon® processor 7500 series at 2.66GHz. The single rack system is made up of 18 blades and features 4.2TB of coherent shared memory.

UV 1000 in Genetics Research

KDRI uses the system for genome sequencing, gene function analysis, DNA polymorphism analysis, and comparative genomics analysis of plant, human and microbe data from a Next Generation Sequencer (NGS). The following applications are in primary use:

- Homology Search: BLAST, FASTA

- Domain Search: InterProScan, HMMER

- Assembler for NGS data: MIRA, WGS assembler, Abyss

- Mapper for NGS data: Bowtie2, MAQ, BWA

According to Dr. Hideki Hirakawa, Chief Scientist, Laboratory of Applied Plant Genomics at KDRI, "In order to assemble and map the massive amounts of read sequences generated by Next Generation Sequencers like Illumina GA and Roche GS, sequence data needs to be stored as a hash table in memory. We needed a system with a very large amount of memory for this purpose. The SGI UV 1000 gives us an unmatched memory size of 4TB, and multiple computing cores at the same time."

He continues, "We considered other computers with maximum memory size of 2TB and 64 CPU cores, but they were not enough for our computing needs. We even considered a virtual SMP system (vSMP), but we were not comfortable betting on vSMP technology for our performance requirement. SGI UV 1000 with NUMAlink® 5 that provides high speed inter-node communication, enabling access to a large pool of shared memory resource, was the best fit for our needs."

NUMAlink 5 is the fifth generation of the SGI-developed system interconnect technology, capable of handling 15GB/s of peak bandwidth through two 7.5GB/s unidirectional links.

sgi

"We have successfully shortened the time for homology searches and domain searches for genome sequencing, as well as data assembling and mapping of data generated by Next Generation Sequencers," states Dr. Hirakawa. "One example of such a successful project was our jatropha genome sequencing. Jatropha is known as the raw material for the creation of biodiesel. We have run BLAST searches on our UV 1000 against the Non-Redundant Protein database maintained by the National Center for Biotechnology Information. The excellent results of this project are being published through *Plant Biotechnology Journal*, and the database is available for public access at www.kazusa.or.jp/jatropha.

System Replaced Older Machines

Dr. Hirakawa says, "Our new UV 1000 replaced a Sun Fire X4600 and an HP DL580. The UV enables us to perform very large-scale assembly, mapping, homology searches and domain searches with the ease of system administration with which we are familiar, since it uses regular Linux® server software. We have successfully shortened these calculations from weeks to days, which significantly accelerates our research projects."

"When we think about the best reason for our purchase of the SGI UV 1000, we cite the increase in speed of sequencing jobs such as homology search and domain search, as well as acceleration of assembling and mapping for Next Generation Sequencer data," Dr. Hirakawa concludes.

About KDRI

KDRI was founded in 1991 and has since been engaged in large-scale sequencing and analysis of the genomes of plants, human and related bacteria. Based on these results, research has been carried out on various genes functioning in plant and human cells.

Please visit www.kazusa.or.jp/e/index.html