The Largest Shared Memory Machine in the State

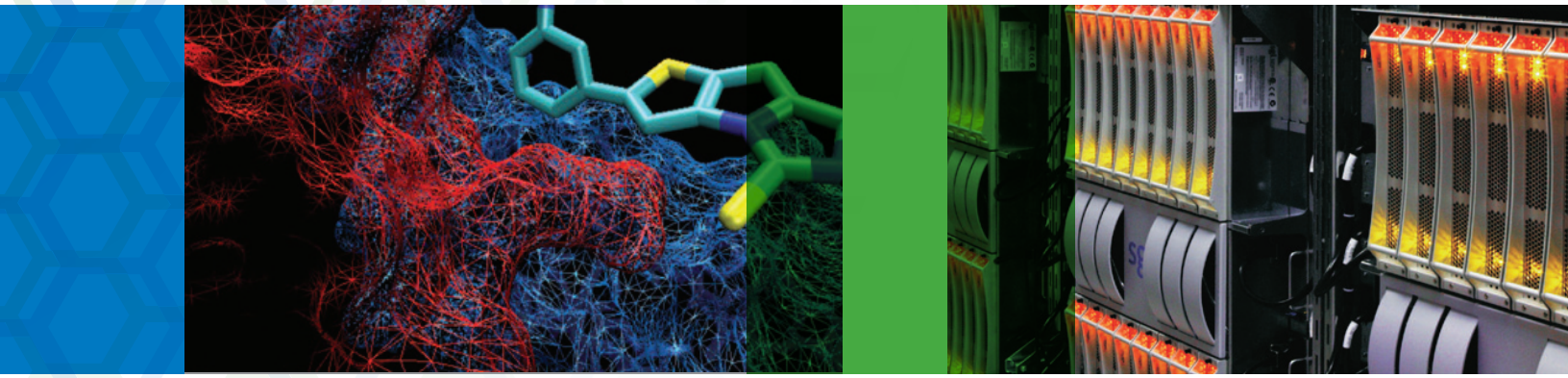
Dozens of Research Groups Accessing System for Interesting Scientific Studies

Key Facts

Organization:
Center for Scientific and Academic Services of Catalonia

Locations:
Barcelona, Spain

Industry:
Higher Education & Research



Users at the Center for Scientific and Academic Services of Catalonia (CESCA) can now access the largest shared memory supercomputer in the Spanish state. The system, named Pirineus, is an SGI UV 1000 with a total of 1,344 processor cores, 6.14TB of memory and a peak performance (Rpeak) of 14.30Tflops/sec. Pirineus replaced an SGI Altix® 3700 that was installed in 2005 and had 128 processors and 384GB of memory. Because of its architecture and large memory resources, the system will allow users to face new research challenges. With installation of the new SGI UV 1000, the peak performance of CESCA resources has jumped from 5.45Tflops/sec to 19.75 Tflops/sec.

At the time of purchase in 2010, several international centers already had SGI UV servers in operation. For example, the North German Supercomputer Alliance, HLRN, has two SGI UV 1000 systems with a total of 4,352 processor cores and 18TB of memory; the University of Tennessee (USA) has an SGI UV 1000 with 1,280 cores and 1TB of memory; and in Spain, the University of Valencia has a system with 26 six-core processors and 832GB of memory.

With Pirineus in operation, the growing demand for calculation power required by research groups in various knowledge areas to which CESCA provides service should be satisfied. CESCA resources had previously registered usage levels of more than 90%, hampering the execution of parallel tasks of more than eight processors and increasing the waiting time for projects with large memory and disk requirements. The new hardware, which provides 968,000 computational hours (CH) per month, helps to solve capacity issues and allows new projects to come on board.

System Configuration Ensures Large Shared Memory and High Throughput Performance

In UV architecture, because of the large bandwidth and low latency in internal communication, and because all subsystems are shared in an efficient manner, the level of parallelism in applications is not limited to the number of processor nodes, and no memory restrictions occur. The UV shared memory system architecture is possible because the Intel® Xeon® processors considerably increase the performance and efficiency while providing great scalability.

The CPUs are grouped into a single system image made up of 1,344 cores. The UV 1000, initially made up of six independent rack units (IRUs) in groups of 256 processors, can be enlarged to up to eight IRUs, creating a single shared memory system.

The SGI UV has a basic blade with a pair of RAID1 disks for the operating system and utilities, as well as PCIe slots for inserting communications cards, GPUs, etc. The product has a blade-type mechanical-electrical architecture and is made up of four 42U cabinets that are powered by a three-phase electrical power network. Each compute blade has a pair of processors and up to 16 DIMMs of RAM. There are five cabinets, which include an SGI InfiniteStorage™ 4100 disk cabinet with 112 Fibre Channel (FC) 1TB, 7,200 RPM disks designed to meet the scratch space requirements of applications that are run on the system. This cabinet is powered by a single phase electrical power network.

Communication with the disk cabinet is through two FC ports at 8Gbps and a 10 Gigabit Ethernet port for the CESCA data network. There is a special node accessible via the web for managing the system, which can also be done via a 1U console unit located inside one of the cabinets.

The internal network is one of high speed and low latency — at only 140ns — allowing all of the processors to work while sharing the same memory space, which could reach tenths of a petabyte. The new SGI UV operates with the SUSE® Linux® Enterprise Server (SLES) distribution, as at the time of installation was the only version of the Linux operating system that could offer a kernel scalable up to 2,048 cores. The standard SLES distribution guarantees compatibility with any other application running on it.

Pirineus Technical Specifications

SGI UV offers a range of multiprocessor configurations with a shared memory design, providing a flexible and competitive working environment for the applications that are characteristic of CESCA users. In comparison to a classic cluster, its architecture allows the Center's users to work with any desired node size without causing any typical cluster application (e.g., programmed with MPI) to stop working. The CESCA configuration includes:

- SGI UV 1000 chassis
- 224 Intel® Xeon® processor X7542, 2.66GHz, 64-bit, six-core, totalling 1,344 cores
- 6.14TB memory
- 112TB hard disk capacity
- 14.3Tflops/sec.

Benchmark Performance of the SGI UV 1000

The new UV arrived on September 10, 2010, and in October, after its configuration and performance testing, start-up benchmarks were performed.

Benchmarks evaluate the performance of the new hardware, as well as its scalability for parallel tasks. Although similar benchmarks were run during the tender process, it was necessary for the run time of these tasks to be longer in order to represent the real burden of work at the Center.

Two categories of benchmarks are typically performed — general system benchmarks and application benchmarks. The SGI UV 1000 particularly stands out in the system benchmark tests designed to evaluate the performance of the hardware at full load. To carry out this assessment, the degradation suffered by a benchmark task (UBJJN2) is calculated when the machine has all cores occupied by two sets of tasks with different types of memory and hard disk requirements. It was discovered that the degradation of the task was only 11.9% and 7.7% when the node is made up of an elevated number of cores (1,344). These results can be compared to those obtained on a Bull Novascale cluster node, where the degradation was of a similar order (7.1% and 7.6%), bearing in mind that it is a node with only eight cores, compared to the 1,344 cores of the new UV.

Application benchmarks were selected in collaboration with a selection of users who run the biggest workloads at CESCA. Applications from various academic areas such as materials sciences, biomedical and life sciences, and environmental sciences were included. In order to define the benchmarks, different types of methodologies and bases were used, using three of the most widely used programs at CESCA — Gaussian, Gamess and NWChem. Gaussian was chosen mainly because it is the most used code at the Center, totalling 75% of available computational hours. Furthermore, it exemplifies calculations with an OpenMP parallelization model.

In the work proposed by Dr. Carlos Alemán of the Technical University of Catalonia, with the UPCCA2 benchmark, the calculation of olygomer is representative of the field of organic semiconductors. Containing the geometry of the compound, a dication to the electronic triplet state, optimized to level UB3LYP/6-31G (d), it is used to calculate the specific energies with various DFT methods. From a chemical point of view, these calculations are interesting for knowing how the difference in the dication of energy changes between the triplet and doublet state with the number of chemical units that make up the olygomer.

Another Gaussian project has been proposed by Dr Mariona Sodupe of the Autonomous University of Barcelona. In this case, the project UABMS2, which uses a numerical method, the Coupled Cluster (CC), was originally formulated in the field of nuclear physics, but then in the 1960s was reformulated in the field of quantum chemistry to accurately deal with the electronic correlation in atoms and molecules. These types of CC calculations, though very accurate, are used only in systems of a few atoms because of the large demand on memory and hard disk. Today, however, thanks to the increase in resources available in the new SGI UV 1000, calculations with tens and even hundreds of atoms can now be tackled in a reasonable time.

The Gamess code was also used, which has a particular implementation in the parallelism with MPI libraries and is one of the programs commonly used throughout the world in the field of theoretical and computational chemistry. In this case, two input files were used — CSICSO2 and UBSO2 — by Dr. Santiago Olivella of the Spanish National Research Council and Dr. Albert Solé from the University of Barcelona, for a research project in the field of atmospheric chemistry using the CASSCF method. These benchmarks are quantum mechanic calculations that use this variational method to improve the description of the wave function of the conventional SCF calculations, including more than one electronic configuration. These types of calculations are necessary when the described system includes energy degenerate systems or open layers, such as in the case of radical peroxides.

The scalability for parallel tasks was evaluated using the UBJJN2 (Gaussian09), UBSO2 (Gamess 2009) and UBER2 (NWChem 5.1.1) benchmarks, together with an ADF 2009.01 task. In this case, multiples of six have been used (and not two as had been done until now) as the Intel® Xeon® processor X7542 consists of six cores. As a result, Gaussian shows good behavior in parallel up to 36 processors, although the software run with MPI shows the best scalability.

As for the comparison of hardware performance at the Center, the best results are obtained with those that implement the Intel® Xeon® processor X5500 series (new nodes on the older system named Prades) and Intel® Xeon® processor X7500 series (on Pirineus), despite having a yield point lower than the Intel® Xeon® processor E5400 series (original Prades nodes). Particularly, with the new SGI UV 1000 (Pirineus), an improvement of 33% was obtained when compared to the old SGI Altix 3700 (named Obacs).

35 Research Groups Put New SGI UV 1000 to the Test

Once the new system was installed, in November the researchers who had requested hours to perform tests have verified its performance on applications such as Gaussian09, VASP, NWChem, Gromacs, NAM and others, with tasks of up to 512 processors. A total of 1,680,000 computational hours were distributed among 35 research groups from various universities and other institutions throughout Catalonia.

One of the first groups of researchers to use the new UV 1000 during the test period was that led by Juan J. Novoa of the Institute of Theoretical and Computational Chemistry, University of Barcelona. The object was to determine the phase transition mechanism in the magnetic bistability crystal of TTTA. This crystal is one of the prototypes of molecular bistable materials. The TTTA crystal shows two phases — a diamagnetic phase at low temperatures and a paramagnetic phase at high temperatures. The group ran two tasks simultaneously on the UV — one to evaluate the transition of the structure from low temperature to high temperature and another to evaluate the inverse transition. A total of 256 cores, 100GB of memory and 50GB of hard drive space were used with the CPMD code, based on the implementation of the Functional Density Theory used by plane waves.

A group led by Agustí Lledós of Autonomous University of Barcelona also used the UV for the computational study of the origin of enantioselectivity in the Suzuki-Miyaura reaction. One of the discoverers of this known coupling reaction, Akira Suzuki, was recently awarded the Nobel Prize in Chemistry. The group hopes to perform a computational study of the reaction mechanism of this first enantioselective version of the Suzuki reaction to determine its origin, which may be the key to designing new, more efficient catalysts for the asymmetric version of this reaction. A total of 64 cores, 100GB of memory and 300GB of hard disk were used with the Gaussian09 software.

The research group of Carles Bo from the Catalan Institute of Chemical Investigation has used the UV to determine the molecular structure of the uranium U28 nanocapsule. The task consisted of calculating the frequencies of the atoms found in the U28 nanocapsule, a very demanding computational operation, especially in regards to memory resources. In this operation, 256 processors and 1TB of memory were used with the ADF software.

Another group was that of Eliseo Ruiz of the University of Barcelona, whose mission is the study of the variation of the magnetic properties of unimolecular magnets deposited on metallic surfaces, with possible applications in electronic devices, generally for the storage of information at a molecular level. The calculations were carried out with 256 processors and 40GB of memory, using the Siesta and NWChem software.

The research project in which Glòria Rodríguez of the Center for Materials Physics, Spanish National Research Council and Polytechnic University of Valencia, participated carried out simulations of the processes of magnetization and spin wave excitation in arrays of particles in a magnetic vortex state. These arrays are good candidates for forming superstructures with adjustable properties that can be used for the transmission of electromagnetic signals without an electric current, as the transmission of the signal is performed by the spin waves traveling through the material. A total of 32 processors, 30GB of memory and 300GB of hard disk were used, along with the Object Oriented Micro Magnetic Framework software.

Domingo Giménez from the University of Murcia used the UV to optimize the dense linear algebra routines that are used in the solving of all manner of scientific problems. These have real applications in medicine and radio frequency components. In this case, 256 processors and 25GB of memory plus hard disk space were used with the BLAS and LAPACK libraries and an OpenMP parallelization model.

About the Center for Scientific and Academic Services of Catalonia

At the Center for Scientific and Academic Services of Catalonia, we help researchers to advance in their work. We give them access to infrastructures which support research, and we offer them platforms for searching and disseminating the scientific activity generated. Please visit cesca.es on the web.

Translated and edited with permission from TERAFLOP, magazine for the Supercomputing Center of Catalonia, December, 2010.

Global Sales and Support: sgi.com/global