

CHAPTER 2

THE EARTH SIMULATOR

Jack Dongarra

The Earth Simulator (ES) is a high-end general-purpose parallel computer focused on global environment change problems. The goal for sustained performance of the Earth Simulator was set to 1,000 times higher than that of the most frequently used supercomputers around 1996 in the climate research field. On a series of real climate applications the Earth Simulator provides a 50-fold computational performance advantage. It represents maximum-capability computing applied to a targeted attempt at a scientific breakthrough in a specific area. It is the result of a focused, long-term, top-down Japanese design effort. Access to the ES is only possible on-site in Japan and through research collaborations with Japanese scientists. In November 2004, two American supercomputers regained the lead in the Top500 list, but until then the Earth Simulator certainly allowed higher resolution and quicker response times in its applications than those available to American scientists.

BACKGROUND

In July 1996, as part of the Global Change Prediction Plan, the promotion of research and development for the Earth Simulator plan was reported to the Science Technology Agency, based on the report titled "For Realization of the Global Change Prediction" made by the Aero-Electronics Technology Committee. In April 1997, the budget for the development of the Earth Simulator was authorized to be allocated to the National Space Development Agency of Japan (NASDA) and Power Reactor and Nuclear Fuel Development Corporation (PNC). The Earth Simulator Research and Development Center was established, with Dr. Hajime Miyoshi assigned as the director of the center. Discussions were made regarding the Earth Simulator Project, at meetings on the Earth Simulator under the Computer Science Technology Promotion Council (Chairman: Prof. Taro Matsuno of Hokkaido University), which was held six times from March to July in 1997. With a report on "Promotion of Earth Simulator Project," specific proposals were put forward to the Science and Technology Agency.

The conceptual system design of the Earth Simulator proposed by NEC Corporation was selected by bidding. Japan Atomic Energy Research Institute (JAERI) had joined the project for PNC.

Under the Computer Science Technology Promotion Council, the Earth Simulator Advisory Committee was instituted with seven members with profound knowledge in the area (Chairman: Prof. Yoshio Oyanagi of Tokyo University). From June to July in 1998, five meetings were held on the basic design. As a result, on the 24th of August it was confirmed with "The Evaluation Report for the Basic Design of the Earth Simulator." In February 1999, the Japan Marine Science and Technology Center (JAMSTEC) joined in the Earth Simulator development project and decided to build the Earth Simulator facility in the Kanazawa ward in Yokohama, which had been the industrial experiment station of Kanagawa prefecture. Manufacturing the Earth Simulator began in March 2000, under NASDA, JAERI, and JAMSTEC. After completion, the entire operation and management of the Earth Simulator was decided to be solely charged by JAMSTEC. At the end of February in 2002, all 640 processor nodes (PNs) started up operation for initial check-up. The Earth

Simulator Research and Development Center verified and gained the sustained performance by AFES (an Atmospheric General Circulation Model for ES), recording 7.2 Tflop/s with 160 PNs, more than 1.44 times faster than the target performance, 5 Tflop/s. The Earth Simulator Center (ESC) with Director-General Dr. Tetsuya Sato began the actual operation in March 2002. On May 2, 2002 the Earth Simulator achieved sustained performance of 26.58 Tflop/s by using the AFES. The highest performance record of 35.86 Tflop/s was achieved with the Linpack Benchmark on the next day. ES was honored with the top position on the Top500 list in June 2002.

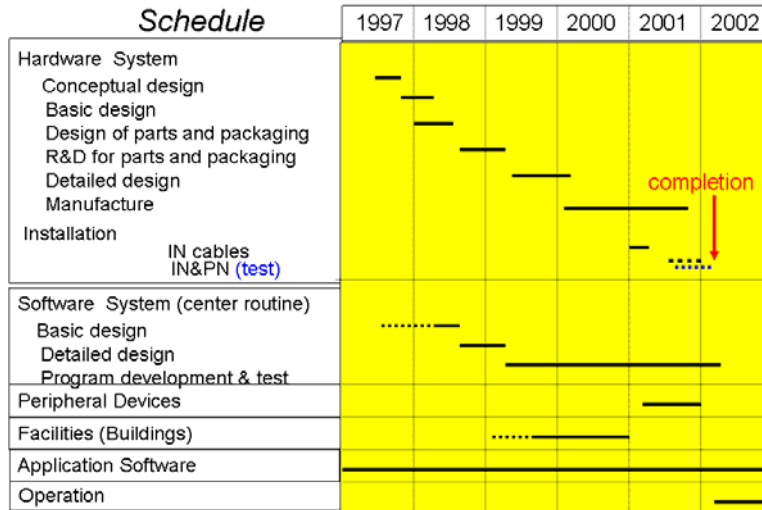


Figure 2.1. Earth Simulator Development Schedule (Courtesy JAMSTEC)

The ES was conceived, developed, and implemented by Hajime Miyoshi. Dr. Hajime Miyoshi was born in Tokyo in 1932. He graduated from the Mathematics Department of the Faculty of Science and Engineering of Waseda University in 1956. He joined the Science and Technology Agency (STA) in 1958 and later moved to the National Aerospace Laboratory (NAL) and became the Director of its Computer Center. He joined STA's Research Institute for Science & Technology (RIST) in April 1995 and became the Director of the Earth Simulator Research and Development Center in April 1997. Hajime Miyoshi is regarded as the Seymour Cray of Japan. Unlike his peers, he seldom attended conferences or gave public speeches. However, he was well known within the HEC community in Japan for his involvement in the development of the first supercomputer in Japan, the 22 Mflop/s Fujitsu FACOM230-75AU, which was installed at the National Aerospace Laboratory (NAL) in August 1977. He stayed at the forefront of HEC development in Japan. He led the development of the Numerical Wind Tunnel (NWT) at NAL. This won a Gordon Bell prize at the annual IEEE Supercomputing Conference in 1994. In 1997 he took up his post as the director of the Earth Simulator Research & Development Center (ESRDC) and led the development of the 40 Tflop/s Earth Simulator. Hajime Miyoshi had profound influence over the Japanese vendors and his work contributed much to the success of today's Japanese HEC industry. With the success of the FACOM230-75AU, Fujitsu developed their first commercial supercomputer, the VP100, which was installed at Nagoya University in the fall of 1983. The development of the Numerical Wind Tunnel helped Fujitsu to develop their VPP500 supercomputers, which were announced in 1992.

Dr. Miyoshi was appointed as Deputy Director-General in 1992 and resigned from NAL in March 1993 as he turned 60, the enforced retirement age for government employees. After he left the aerospace engineering community, he targeted the next supercomputer. He foresaw Computational Fluid Dynamics (CFD) technology achieving a matured status, and he believed CFD application could easily be done at the Numerical Wind Tunnel and following commercially based machines. He understood that global weather and ocean circulation simulation lacked the power of even the NWT. These areas were becoming more important as the global environmental issue deepened in importance. So his final target was to



Figure 2.2. Dr. Miyoshi

develop the Earth Simulator, which would serve as powerful computational engine for global environmental simulation.

Prior to the ES, global circulation simulations were made using a 100km grid width, though without coupling ocean and atmospheric models. To get quantitatively good predictions for the evaluation of environmental effects may require at least 10 km grid width or 10 times finer meshes in x, y and z directions and interactive simulation. Thus a supercomputer 1000x faster and larger than a 1995 conventional supercomputer might be required. Miyoshi investigated whether such a machine could be built in the early 2000s. His conclusion was that it could be realized if several thousand of the most advanced vector supercomputers of approximately 10 Gflop/s speed were clustered using a very high-speed network. He forecasted that extremely high-density LSI integration technology, high-speed memory and packaging technology into small-size, high-speed network (crossbar) technology, as well as an efficient operating system and Fortran compiler all could be developed within the next several years. He thought only a strong initiative project with government financial support could realize this kind of machine.

In 1997 he organized the ES Research & Development Center, which was financially supported by JAMSTEC, JAERI and NASDA, all of them belonging to the former STA, which has now been merged into the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of the Japanese government. He became the head of the ES R&D Center.

The machine was completed in February, 2002 and presently the entire system is working as an end user service. He supervised the development of NWT Fortran as the leader of NWT project and organized HPF (High Performance Fortran) Japan Extension Forum, which is used on the ES. He knew that a high-level vector/parallel language is critical for such a supercomputer.

ARCHITECTURE

The Earth Simulator is a highly parallel vector supercomputer system consisting of 640 processor nodes and an interconnection network. Each processor node is a shared memory parallel vector supercomputer, in which eight arithmetic vector processors (AP) are tightly connected to a main memory system. The vector processor has a theoretical peak of 8 Gflop/s, and main memory of 16 GB. The total system consists of 5,120 vector processors. This gives the system a theoretical peak of 40 Tflop/s and 10 TB of memory capacity.

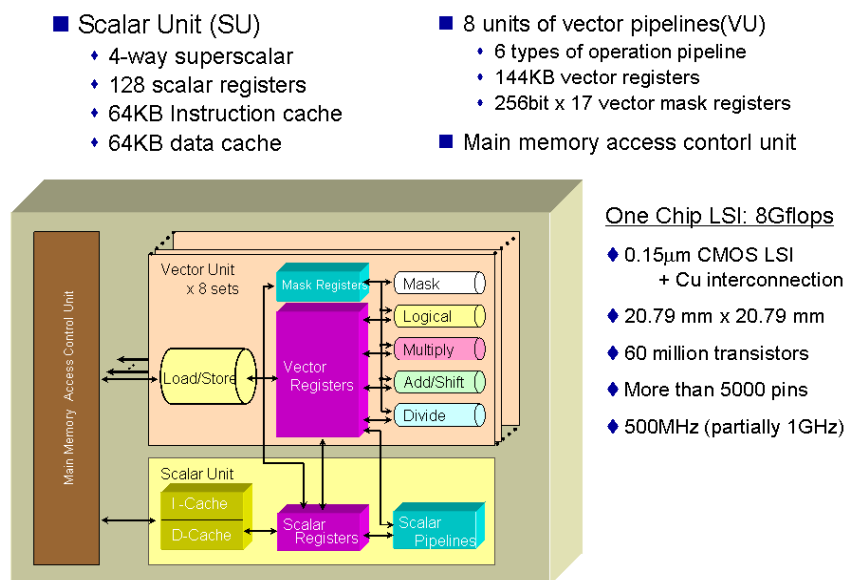


Figure 2.3. Arithmetic Processor Configuration (Courtesy JAMSTEC)

- Peak performance/AP : 8Gflops
- Peak performance/PN : 64Gflops
- Main memory/PN : 16GB
- Total number of APs : 5120
- Total number of PNs : 640
- Total peak performance: 40Tflops
- Total main memory : 10TB

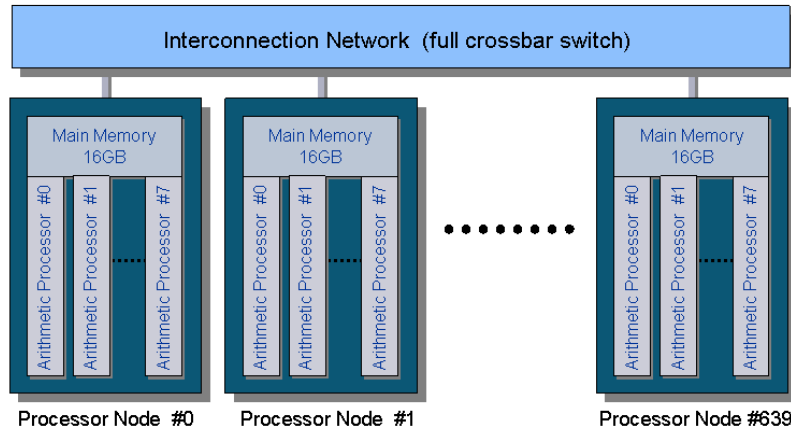


Figure 2.4. Configuration of the Earth Simulator (Courtesy JAMSTEC)

The interconnection network is a 640 by 640 non-blocking crossbar switch that connects the 640 processor nodes. The interconnection bandwidth between every two nodes is 12.3 GB/s in each direction. The aggregated switching capacity of the interconnection network is 7.87 TB/s.

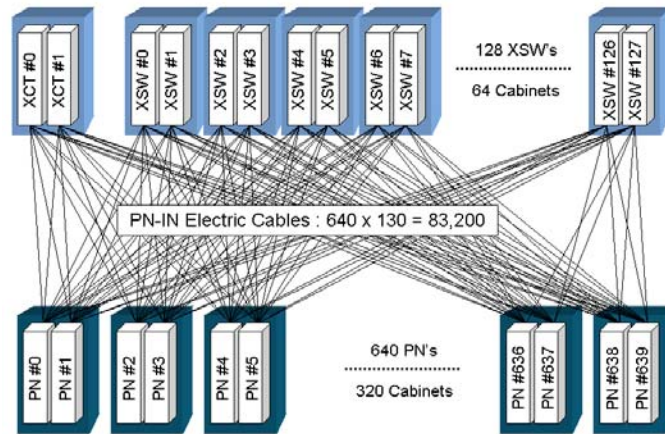


Figure 2.5. Connection between Cabinets (Courtesy JAMSTEC)

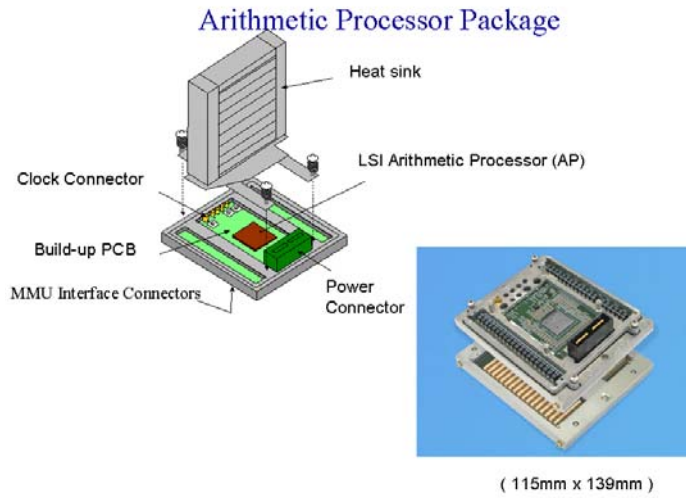


Figure 2.6. Arithmetic Processor Package (Courtesy JAMSTEC)

The vector processor is a single chip processor. The LSI uses 0.15 μm CMOS (Complementary Metal-Oxide Semiconductor) technology with copper interconnections. The vector pipeline unit operates at 1 GHz while the other components operate at 500 MHz. The processor is air-cooled and dissipates at approximately 140 Watts.

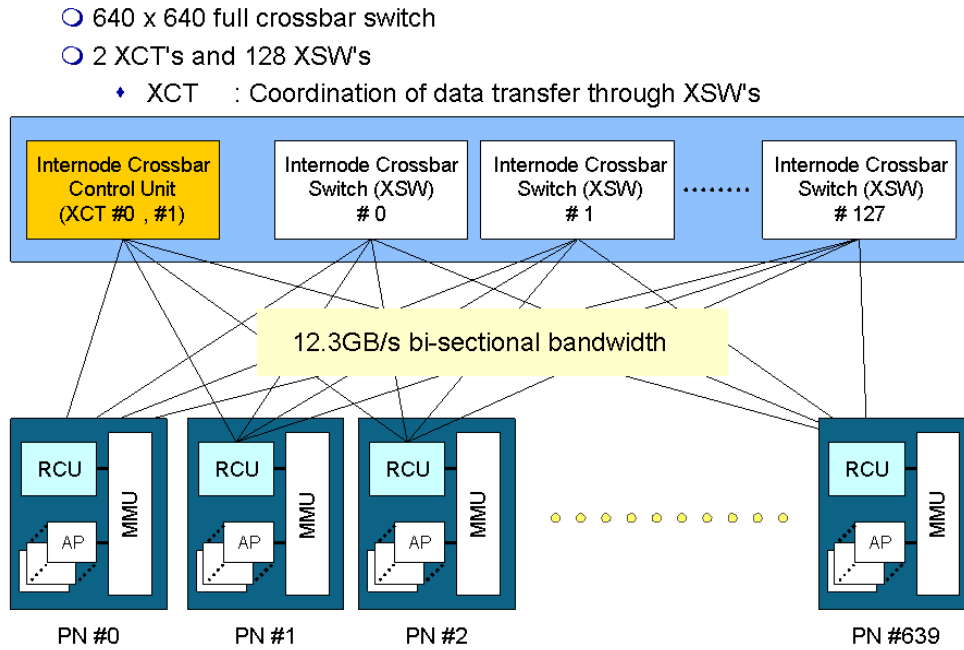


Figure 2.7. Interconnection Network (IN) (Courtesy JAMSTEC)

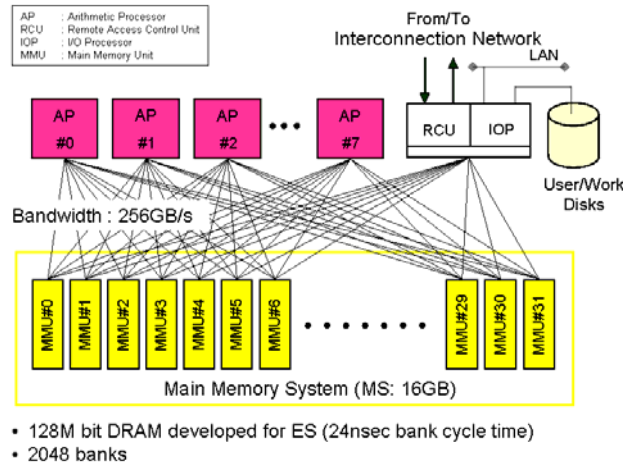


Figure 2.8. Processor Node Configuration (Courtesy JAMSTEC)

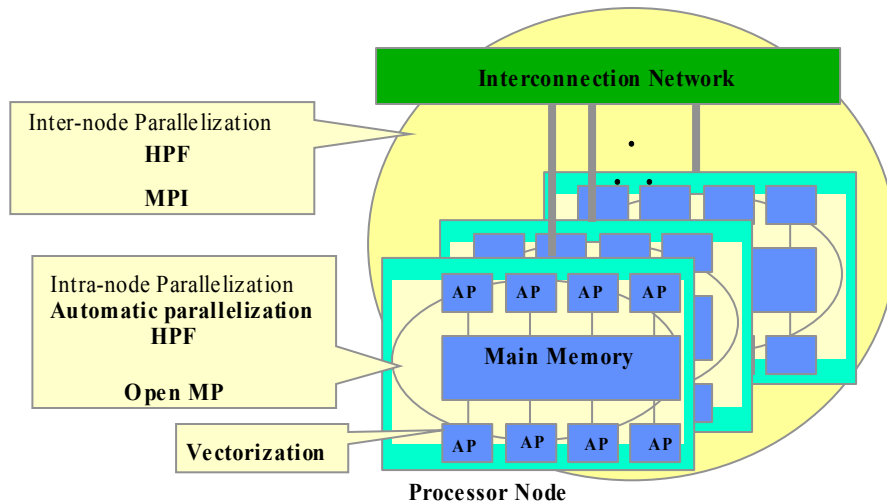


Figure 2.9. Vectorization and Parallelization (Courtesy JAMSTEC)

SOFTWARE

Applications

One of the strengths of the Earth Simulator project is its focus on a particular important class of problems. An attempt to understand large-scale phenomena on the earth such as global warming and El Niño events by numerical simulations is a very significant and challenging effort, in addition to the promotion of computational science and engineering. The Earth Simulator project had been started by the STA in 1997 aiming to understand and predict global environment change on Earth. The ESRDC was a joint team established by NASDA, JAERI, and JAMSTEC.

Between July 2002 and March 2003, there were only five whole system failures. Twice the problems arose with the Interconnection Network Nodes and three times with the job scheduling operation software. This is minimal trouble considering the Earth Simulator consists of 640 processor nodes and it has an enormous number of parts: 5120 vector processors, 1 million memory chips, 20,000 memory controller chips, 160,000 serial-parallel translate chips and so on. The storage system consists of RAID5 disk arrays, which helps to save user data from disk failures. Some small errors happen every day. However, practically every error is

recovered by using a correcting system or by retrying the hardware and software. The diagnosis reports are checked and preventive maintenance is done on nodes in which some recoverable errors have happened.

One of the most important problems is how to store and move massive amounts of user data. This problem has two sides. One is how to provide processing nodes with data from permanent storage systems. Another is how to retrieve data from the Earth Simulator. The ES has such a high performance level that the simulation data itself proves enormous in size. It has a tape cartridge system, which provides petabyte-class storage. However, the accessibility of the tape system is not so convenient which creates problems for users and wastes system resources. A hierarchical storage system, consisting of disks and tapes, was implemented at the ES in December 2003, which has improved this situation considerably.

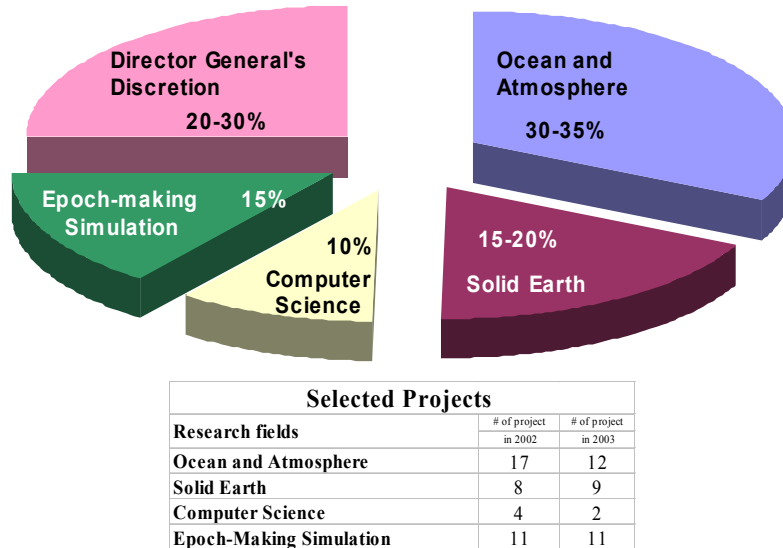


Figure 2.10. Allocation of Computer Resources (Courtesy JAMSTEC)

USAGE

For an application to use a large number of processors on the ES a number of conditions must be satisfied.

- Number of PNs for a job must be less than or equal to 10
- This can be extended to 512 by the application, based on other conditions being satisfied.
- Wall clock time for a job is less than or equal to 12 hours
- Number of PNs \times wall clock time is less than 1536 PN*hours
- The number of PNs can be expanded if the vectorization ratio is greater than or equal to 95% and the parallelization efficiency is greater than 50%

These conditions will insure that the machine achieves relatively high efficiencies.

○ Vectorization ratio $\geq 95\%$

(vector operation ratio may be used)

○ Parallelization efficiency $\geq 50\%$

*If a program needs T_1 hours with 1 node,
and T_n hours with n nodes*

$$\text{Parallelization efficiency} = (T_1/T_n)/n$$

Parallelization ratio should be more than 99.9% for a program using more than 128 PN's to keep the parallelization efficiency as 50%, if the parallelization ratio is not affected by the number of PN used for the computation.

Figure 2.11. Condition to extend the PN number for a JOB (Courtesy JAMSTEC)

REFERENCES

- Annual Report of the Earth Simulator Center, Outline of the Earth Simulator Project. 2003.
<<http://www.es.jamstec.go.jp/esc/images/annualreport2003/pdf/outline/outline.pdf>> Last accessed February 23, 2005.
- Habata, S., M. Yokokawa, K. Shigemune. 2003. The Earth Simulator System. *NEC Res. & Develop.*, Vol. 44, No. 1,
<<http://www.owl.net.rice.edu/~elec526/handouts/papers/earth-sim-nec.pdf>> Last accessed February 23, 2005.
- JAMSTEC, Earth Simulator Hardware. <<http://www.es.jamstec.go.jp/esc/eng/ES/hardware.html>> Last accessed February 25, 2005.
- Sumi, A. 2003. The Earth Simulator and Its Impact for Numerical Modeling.
<<http://www.tokyo.rist.or.jp/rist/workshop/rome/ex-abstract/sumi.pdf>> Last accessed February 25, 2005.