

**Earth Simulator - An Analysis****Shubham Verma**

Department of Computer Science & Engineering
Noida Institute of Engineering & Technology, Greater Noida, India

Ravi Gupta

Department of Computer Science & Engineering
Noida Institute of Engineering & Technology, Greater Noida, India

Rajesh Yadav

Department of Computer Science & Engineering
Noida Institute of Engineering & Technology, Greater Noida, India

Abstract:

This document contains concurrent and parallel computing seminar work at Lappeenranta University of technology; the topic focuses on super computer technologies, specifically Earth Simulator. The Earth Simulator (ES) is a high-speed super computer technology that was started as a research project in 1997 to better understand and predict the climate change and global environment at large. The Earth simulator is a large scale super computer which encompass of 640 processor elements bonded by a single stage full crossbar network.

Every node element contains 8 vector processors, which uses one single shared memory of capacity 16GB. This massive super computer brought about a high scaled performance of 40Tflops and 10TB memory capacity.

The goal of the Earth Simulator project is to simulate atmospheric and oceanic phenomena on Earth. There are five research projects, each of which concentrates on a specific area of Earth research.

Key words: Supercomputer, HPC (High Performance Computing), Parallel, Shared Memory, Vector Processor, Crossbar Network.

I. Introduction

1.1. Supercomputer

Supercomputers are the type of computer which works at a faster rate and they are referred to as the fast computers. As these computers are quite fast so they are quite expensive to buy and even they are operated by those people who are particularly specialized in the usage of such computers. This type of computer is also meant for several calculations as it can solve all the problems at a higher speed.

The computer is mainly used for the mathematical and numerical functions as the supercomputers are very strong in calculations so they provide answer to each and every number that had been entered there. The main function of supercomputers is to provide each and everything to the person at the faster rate most of the time that is spent by the supercomputers is in the execution of the programs so that they can be opened at a faster rate. This is one of the most beneficial features which make a huge difference between the supercomputers and the other types of computers that had been invented so far. Supercomputers are very much efficient in solving all the problems that are complex and are scientific in nature. First supercomputer (Cray-1) was made in 1976 by the US engineer Roger Cray (1925-1996).



Figure 1: IBM Blu-Gen supercomputer

1.2. Introduction To Earth Simulator

The Earth Simulator is a special purpose machine, made by NEC with the same type of vector technology as is available on the SX-6. The decision of NEC to base the design entirely on vector processor indicates a new trend towards supercomputer design. Vector processors are

already around for some time but no manufacturer took the risk of implementing them full scale in supercomputers.

This is a different approach than the Americans are taking, who base the large machines on off-the shelf technology. A clear advantage of the Japanese approach is that they "only" need 5120 processors for the Earth Simulator. The Ascii White at the US Lawrence Livermore National Laboratory and made by IBM, already needs 8192 processors for 7 Tflop/s. The importance of this machine can be found in the type of research environment. The issues of environmental changes have become more and more important during the past decade where pollution became serious problem threatening climate and our world we live in.



Figure 2: The Earth Simulator

And the climate is exactly where the ESS is going to support research. ES was replaced by the Earth Simulator 2 (ES2) in March 2009. ES2 is an NEC SX-9/E system, and has a quarter as many nodes each of 12.8 times the performance (3.2x clock speed, four times the processing resource per node), for a peak performance of 131 TFLOPS. With a delivered LINPACK performance of 122.4 TFLOPS, ES2 was the most efficient supercomputer in the world at that point. In November 2010, NEC announced that ES2 topped the Global FFT, one of the measures of the HPC Challenge Awards, with the performance number of 11.876 TFLOPS.

2. Hardware Configuration Of Earth Simulator

The Earth Simulator is massively parallel vector system with shared memory architecture, which is constituted of 640 processor nodes, bonded by (640 X 640) single stage crossbar switches. each node holds 8 vector processor with maximum performance of 8 Gflops, and shared memory

of capacity 16GB, remote access control unit which links and controls data path and data transmission, also Input and output (I/O) processor. Any single processor has vector operation unit, a four way superscalar operation unit, and main memory access control unit.

The whole hardware system is divided into three parts:-

- Interconnection network
- Network Queuing System
- Mass Data Processing System (MDPS)

2.1. Interconnection Network

Usually in multiprocessing systems, a different level of communications and cooperation occurs for solving a given problem, such as by sending message or by sharing memory. Parallel systems must be strongly attached if there are numerous processors interacting through shared memory. Parallel processing requires the use of capable system interconnection for speedy communication among multiprocessors, shared memory, I/O, and peripheral devices. The machine's speed is determined by three key factors, the memory bandwidth, the interconnection network topology, and an interconnection network's dynamic behavior.

Switched network such as crossbar network makes available dynamic interconnection among processor and main memory. A crossbar network offers the higher bandwidth and optimal interconnection. A dedicated connection path between two nodes is being setup by cross point switch, the switch can be set on or off dynamically upon a demand.

The Earth simulator interconnection network is a very enormous network of 640x640 non-blocking crossbar switches that connects the 640 nodes. The interconnection bandwidth between every pair of nodes is about 12.3 GB/s full-duplex. The entire switching capacity of the interconnection networks is approximately 7.87 TB/s.

Remote control unit (RCU) is connected to the crossbar switches and manages inter-node data transfer at 12.3 GB/s full duplex transfer rate for both sending and receiving data, hence the whole bandwidth of the network among nodes is about 8 TB/s.

The earth simulators crossbar network consist of two units as can be seen in the figure below; Internode crossbar control unit (XCT), which is responsible for coordinating of switching operations.

Another one is the inter-node crossbar switch (XSW) which is a real data path.

2.2. Network Queuing System

Since ES uses batch job system, the Earth nodes are classified as clusters of two types, S-clusters and L-clusters. One cluster from the 40 clusters of ES (cluster 0 nodes from 0-15) is dedicated for S-cluster operations, which provides interactive features for end users such as debugging, and compiling. Within S-cluster node include special nodes for job execution. The rest of the 39 clusters (625 nodes) are positioned in L-cluster to provide specific batch environment to execute massively parallel computing tasks employing fast number of processor nodes. To perform these massive and huge tasks the MDPS utilizes mechanism called job scheduling.

2.3. Mass Data Processing System

While numerous large-scale distributed parallel programs are carried out on the Earth simulator (ES), I/O processing for multiple data such as user files and work files has become the bottleneck of numerical simulations. Handling huge volume of data generated by the ES is extremely difficult and creates significant problems.

Therefore, in 2003 it was started on network queuing II and Mass data processing system (MDPS) to overcome these problems, and to improve the system utilization and maintainability as well.

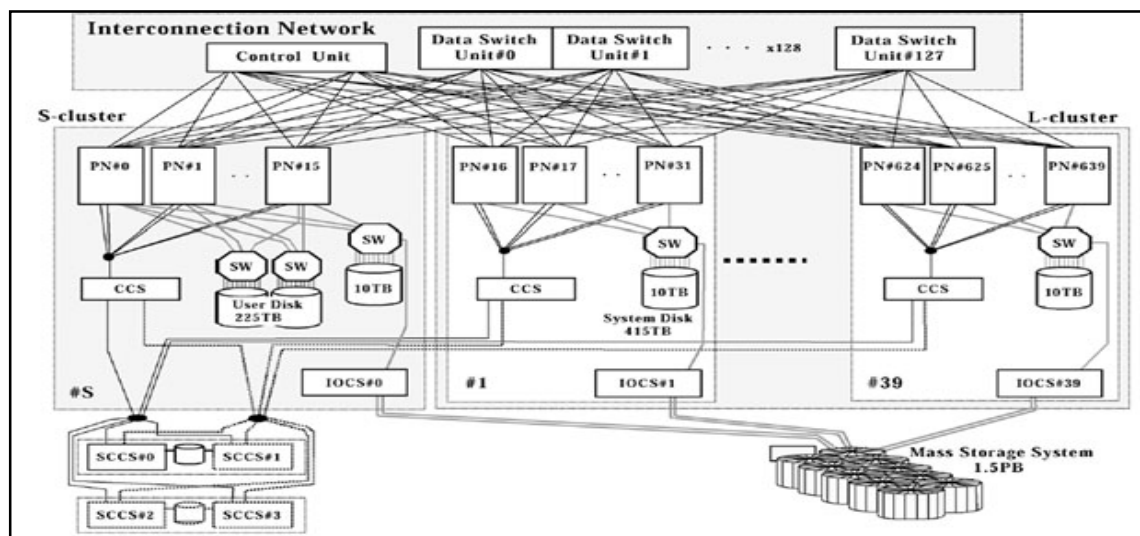


Figure 3: Hardware configuration of ES

3. Software Configuration Of Earth Simulator

All of the software available on the ES system are designed and developed so that users can fully and readily exploit the outstanding performance of the world's largest computer. Below are the description of software technologies used in the operating system, Job Scheduling and the programming environment of ES:

3.1. Operating System

The operating system running on ES is developed for NEC's SX Series supercomputers. SUPER-UX is an operating system that takes the function from BSD and SVR4.2MP as an operating system based on

UNIX System V, and strengthens the function necessary for the super computer in addition.

3.2. Programming Environment

Parallel programming plays essential role in large scale a scientific computation, parallelization techniques often employed is task level program partitioning. This technique entails data communication between different partitioned tasks, reducing the communication time between these tasks would have significant impact on the performance the program's execution time.

So according to the above properties Earth Simulator have three compilers:

- FORTAN 90
- C++
- C

3.3. Job Scheduling

ES is basically a batch-job system. Network Queuing System II (NQSII) is introduced to manage the batch job.

The Earth nodes are classified as clusters of two types, S-clusters and L-clusters. One cluster from the 40 clusters of ES (cluster 0 nodes from 0-15) is dedicated for S-cluster operations, which provides interactive features for end users such debugging, and compiling. Within S-cluster node include special nodes for job execution. The rest of the 39 clusters (625nodes) are positioned in L-cluster to provide specific batch environment to execute massively parallel

computing tasks employing fast number of processor nodes. To perform these massive and huge tasks the MDPS utilizes mechanism called job scheduling.

A process node in the earth simulator is restricted to access to user disk, thus user files are moved from user disks in the S-cluster to work disks in L-cluster before the job is carried out.

This method of moving user files from user disks to work disk for job execution in the L-cluster is called “Stage –in”, while when moving results computed in the Earth simulator to user disk for users is called “ Stage-out”.

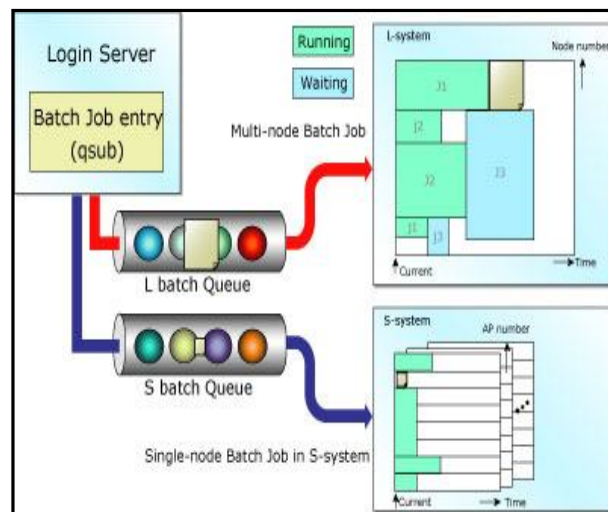


Figure 4: Job Scheduling

4.Application Of Earth Simulator

As mention earlier, the Earth Simulator was established to monitor the Earth’s environmental changes to prevent natural disasters such as earth quake, hurricanes, heat weaves, or extensive and lengthy heavy rains.

The Earth Simulator has been employed for many research areas, focusing on specifically the Earth research. In this section we will describe few of these application and give a brief description of 2008 Earth Simulator project.

The Atmosphere and Ocean Simulation project was one of the earliest researches implemented by a research group, this research intended to develop a better ways of enhancing weather forecasts and attempt to earlier prediction upcoming disasters. In this research the Earth Simulator simulates a portion of the earth and sea water’s surface and their impact on weather were studied.

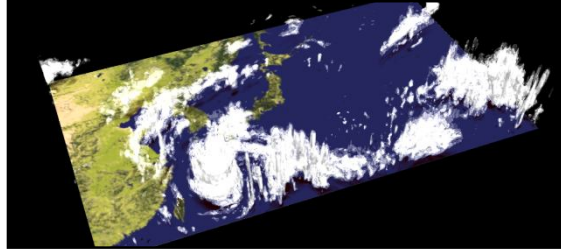


Figure 5: The predictability of the atmosphere

The research exploration made by Multi-scale simulation research group intended to examine and develop non-hydrostatic coupled ocean-atmosphere simulation code with high computational performance. The code enables us to simulate physical process on the earth and it also provides mechanism to comprehend of both weather and climate system. This is holistic simulation discovers the complex interdependency among micro-and –macro scale process.

5.Advantages & Disadvantages

There are two big advantages to performing a simulation rather than actually building the design and testing it-

- The biggest of these advantages is money. Designing, building, testing, redesigning, rebuilding, retesting for anything can be an expensive project. Simulations take the building/rebuilding phase out of the loop by using the model already created in the design phase. Most of the time the simulation testing is cheaper and faster than performing the multiple tests of the design each time. Considering the typical university budget cheaper is usually a very good thing. In the case of an electric thruster the test must be run inside of a vacuum tank. Vacuum tanks are very expensive to buy, run, and maintain. One of the main tests of an electric thruster is the lifetime test, which means that the thruster is running pretty much constantly inside of the vacuum tank for 10,000+ hours. This is pouring money down a drain compared to the price of the simulation.
- The second biggest advantage of a simulation is the level of detail that you can get from a simulation. A simulation can give you results that are not experimentally measurable with our current level of technology. Results such as

surface interactions on an atomic level, flow at the exit of a micro electric thruster, or molecular flow inside of a star are not measurable by any current devices. A simulation can give these results when problems such as it's too small to measure, the probe is too big and is skewing the results, and any instrument would turn to a gas at those temperatures come into the conversation. You can set the simulation to run for as many time steps you desire and at any level of detail you desire the only restrictions are your imagination, your programming skills, and your CPU.

There are two big disadvantages to performing a simulation as well-

- The first of these disadvantages is simulation errors. Any incorrect key stroke has the potential to alter the results of the simulation and give you the wrong results. Also usually we are programming using theories of the way things work not laws and theories are not often 100% correct. Provided that you can get your simulation to give you accurate results you must first run a base line to prove that it works. In order for the simulation to be accepted in the general community you have to take experimental results and simulate them. If the two data sets compare, then any simulation you do of your own design will have some credibility.
- The other large disadvantage is the fact that it is a simulation. Many people do not consider what they do engineering unless they can see, hear, feel, and taste the project. If you are designing a light saber a typical engineer needs to be able to hold the light saber in their hand in order to consider the project worth his or her time. If you are capable of moving your craft into the virtual world of simulations you are no longer restricted by little things like reality. If you want to design a light saber in the virtual world it is not a problem, but in reality that is another matter altogether. The virtual world is difficult to get use to the first time you use it for design, but after that the sky isn't even your limit.

6. Conclusion

The Earth Simulator has been ranking for many years on the top of the list of supercomputer presented in top500 list, until other super computers such as IBM runner, Blue-Gene, Columbia NASA came along to take over its place in the ranking.

This paper work presents the architectural overview of ES, network topology used in the ES and interconnection between processor nodes. Similarly this seminar introduces ES data storage, work disks, user disks and mechanism to handle its job management. This seminar work also covers software aspect of the ES and its programming environment including various libraries used in the ES. Finally the application and research domain of the ES is presented.

7.Reference

1. Top 500.org - <http://www.top500.org/>
2. Intel Microprocessor Quick Reference Guide -
<http://www.intel.com/pressroom/kits/quickrefyr.htm>
3. The Architecture of Earth Simulator
<http://www.cmpe.boun.edu.tr/courses/cmpe511/fall2004/Gulfem%20Isiklar%20-%20Earth%20Simulator.doc>
5. <http://en.wikipedia.org>
6. http://www.thocp.net/hardware/nec_ess.html
7. Job Scheduling: <http://www.jamstec.go.jp/es/en/system/scheduling.html>