How Can the Earth Simulator Impact on Human Activities

Tetsuya Sato, Hitoshi Murai, and Shigemune Kitawaki
Earth Simulator Center, Japan Marine Science and Technology Center, 3173-25 Showa-machi, Kanazawa-ku, Yokohama, 236-0001, Japan, {satot, murai, kitawaki}@es.jamstec.go.jp

Abstract. The Earth Simulator (ES) is a vector-parallel supercomputer, consisting of 5120 vector processors. The peak performance of each vector processor is 8Gflops. Eight processors make one node with 16GB shared-memory and 64Gflops peak performance. The total system thus consists of 640 nodes that are connected by a single stage full crossbar network. The development project started in April, 1997 and was completed in February, 2002. In May, 2002 remarkable sustained performance of 35.86Tflops in the Linpack benchmark was achieved, which is a surprising result for a distributed memory parallel system. More surprisingly, the ES has achieved 26.58Tflops for an application program, specifically, an optimized atmospheric global circulation simulation code. This striking performance assures that the ES can bring humans crucial impacts on many fields, such as environmental preservation, human life, manufacturing process, and scientific methodology.

1 Introduction

One hopes eagerly that natural and environmental changes such as typhoons, El Niño, earthquakes, global warming, etc. be predicted accurately well in advance. Current computer capabilities are insufficient to carry out trustable simulations for global changes in climate and generation of earthquakes.

The Earth Simulator (ES) project was planned with aiming at elucidation and prediction of global environmental changes as precisely as possible. The ES is now in operation at the Earth Simulator Center (ESC), Japan Marine Science and Technology Center (JAMSTEC).

2 Overview of the Earth Simulator

2.1 Hardware System

The Earth Simulator is a distributed memory parallel system which consists of 640 processor nodes connected by a 640 × 640 single-stage crossbar switch (Figure 1). Each node is a shared memory system which is composed of eight arithmetic vector processors (AP), a shared memory system of 16GB, a remote access control unit (RCU), and an I/O processor (IOP). The peak performance
of each AP is 8Gflops. Therefore, the total number of processors is 5120 and the total peak performance and the main memory capacity are 40Tflops and 10TB, respectively. A 0.15 micron CMOS technology with Cu interconnection is used for LSIs[1, 2].

The AP contains a vector unit (VU), a 4-way super-scalar unit (SU), and a main memory access control unit which are mounted on a one-chip LSI. The chip size is about 2cm × 2cm and it operates at clock frequency of 500MHz, partially 1GHz. The VU consists of 8 sets of vector pipelines, vector registers, and some mask registers. Vector pipelines have six types of operation pipeline which are add/shift, multiply, divide, logical, mask, and load/store pipelines. Eight operation pipelines of the same kind work together by a single vector instruction and different type of the pipelines can operate concurrently. There are 72 vector registers of 256 vector elements. The SU is a super-scale processor with a 64KB instruction cache, a 64KB data cache, and 128 general-purpose scalar registers. Branch prediction, data prefetching and out-of-order instruction execution are employed. The VU and SU support the IEEE 754 floating point data format.

The memory system (MS) in the node is equally shared by 8 APs and is configured by 32 main memory package units (MMU) with 2048 banks. A 128 mega-bits high speed DRAM operating at 24 nsec bank cycle time is used for the memory chip. The memory capacity of each node is 16GB. Each AP has a 32 GB/s memory bandwidth and 256 GB/s in total. The RCU in the node is directly connected to the crossbar switch by two ways of sending and receiving, and controls inter-node data communications. Several data transfer modes such as three-dimensional sub-array accesses and indirect accesses are supported.

The single-stage crossbar network (IN) consists of two units; One is the inter-node crossbar control unit (XCT) which is in charge of coordination of switching operations. The other is the inter-node crossbar switch (XSW) which is an actual data path. XSW is composed of 128 separated switches, each of which has 1Gbits/s bandwidth operating independently. All the pairs of nodes and switches