Supercomputer SX-9 Development Concept

SX-9 - the seventh generation in the series since the announcement of SX-1/2 in 1983 - is NEC's latest supercomputer that not only extends past SX-series achievements in large-scale shared memory, memory bandwidth and inter-node communication bandwidth, but also delivers more than 100GFLOPS CPU performance for improved ease of use and performance in a wider scope of computation in the field of science and technology. While tracing the advances achieved over the history of the SX series and describing its development concept, this special issue will provide readers with an introduction to the hardware, basic software and storage systems of the new SX-9.

Executive General Manager, 1st Computers Operations Unit, NISHIKAWA Takeshi
Executive General Manager, Computers Software Operations Unit, YAMAMOTO Masato
General Manager, HPC Marketing Promotion Division, HISAMITSU Fumihiko
General Manager, Computers Division, NOGUCHI Takayuki
General Manager, 1st Computers Software Division, AKATSU Motoyasu
General Manager, System Storage Products Division, MARUYAMA Takao

1 Preface

In step with recent advances in IT technology, the usage of computers required for large-scale and high-speed arithmetical calculations and, consequently, the domain of HPC (High Performance Computing) applications have expanded.

In the field of science and technology, which is an early adopter of HPC technology, the importance of HPC in computational simulations as a platform for research and development is increasing. Simulation models and methodologies are growing increasingly more precise and complex, and the scope of supercomputer applications is expanding in diverse fields ranging from the cutting edge of natural science such as physics, nuclear energy, meterology and climatology, environment science and aerospace to industrial fields including automotive collision analysis, chemistry, material development and subsurface resource exploration, electrical/mechanical design, structural analysis in architecture and construction engineering. In addition, there has been growing usage of HPC technology in diverse fields outside of science and technology such as finance, web search engines, and data visualization.

2 NEC SX Series History

In April 1983, NEC made a dramatic entry in the supercomputer market with its simultaneous announcement for the first two models in the series: the SX-1 and the SX-2 - the world's fastest supercomputers at the time. SX-2 achieved peak performance of 1.3GFLOPS (1.3 billion floating-point operations per second) - the first time for a computer to enter the gigaflop territory, and the world-leading speed of the model was verified by benchmark programs in various categories.

The successor SX-3 was announced in April 1990 and rewrote world speed records with peak performance of 22GFLOPS using 4 multiprocessors and maximum 5.5GFLOPS per CPU. Offering a lineup of ISV (Independent Software Vendor) application software, the SX-3 not only featured Japan's first shared-memory-type multiprocessor combined with paral-

lel processing technology, but also pioneered the age of open systems with its 64-bit "SUPER-UX" operating system, which was developed based on UNIX and boasted significant enhancements to respond to the demands of operation on a supercomputer.

Featuring a significant increase in expandability to a maximum of 512 processors, SX-4 was announced in November 1994, achieving its maximum performance in the teraflop territory. The shift from conventional silicon bipolar LSIs to CMOS for the CPU resulted in both a reduction of power consumption and higher density packaging, consequently leading to a cost reduction by allowing the use of air cooling instead of previously used liquid-cooling systems. In 1996, expansion in memory capacity by enabling the use of DRAM instead of SRAM as the memory device and the consequent improvement in a price performance ratio combined with the comprehensive array of application software received high reputation both in Japan and other countries.

In June 1998, SX-5, the fourth generation in this series, was unveiled. Doubling both the clock frequency and the number of vector pipelines, the evolved model achieved CPU vector performance of 8GFLOPS, system peak vector performance of 4TFLOPS with a configuration of parallel processing with 512 CPUs.

The release of the fifth generation SX-6 in October 2001 introduced the integration of the previous model's CPU of 30 LSIs in a single chip. With 1,024 CPUs each boasting a maximum CPU performance of 8GFLOPS, the SX-6 delivered peak performance of 8TFLOPS. Equipped with a maximum of 32 CPUs per node and a maximum of 256Gbytes of shared memory combined with automatic parallelization, the introduction of the SX-7 in October 2002 realized further ease of operation.

Also during this period, delivery of the Earth Simulator, for which NEC was in charge of the hardware and basic software design and manufacture, was completed. The Earth Simulator began operation in March 2002, and its overwhelming performance not only amazed the world, but also continues to make a large contribution to understanding and forecasting climate change on a global scale.

Shipping of the sixth generation SX-8 commenced at the end of 2004, achieving the world's fastest CPU performance of 16GFLOPS and ultrafast peak performance of 65TFLOPS with a maximum of 4,096 CPUs.

Subsequently in October 2007, the seventh generation SX-9 was announced. Featuring the world's first 102.4GFLOPS performance per single core, large-scale shared memory up to 1 Terabyte, and ultrafast interconnects with a blazing 128Gbyte/sec, it delivered near-PetaFLOPS peak performance of 839TFLOPS. In addition, cutting edge LSI technology and high-density packaging technology reduced both power consumption and required installation space to approximately a quarter of that required for a conventional supercomputer.

3 SX–9 Development Concept and Features

3.1 SX–9 Development Concept

Users look to supercomputers as a means to perform their applications at greater speed. The vital point is not the height of their peak theoretical performance, but how simply users can exploit high sustained performance out of them. Tapping all the past experience and know-how accumulated during the development of the SX series, the development of the SX-9 aims at facilitating both more sustained performance and economical operation.

(1) DNA of Ever Faster Processors and Shared Memory

FLOPS (FLoating point number Operations Per Second) is a measure of supercomputer performance and indicates the number of floating point calculations, which are especially used in the field of scientific calculations, can be performed per second. While supercomputers in recent years have appeared claiming PetaFLOPS-class performance based on special benchmark codes, getting more performance out of practical applications is not just a matter of accelerating hardware. With increased performance have come a multitude of factors and considerations that place an increased burden on the people who use (program on) supercomputers, for example, parallelization of programs (using MPI), uniformity of balancing computational load, minimization of communication time and synchronization frequency, etc. Against the background of the expanding fields of application and more general usage of supercomputing, it is increasingly important to lessen the user burden as much as possible and make using and getting the desired results out of a supercomputer simpler. With this in mind, the SX-9 puts a priority on single processor performance, elevating the SX-8's 16GFLOPS to a lofty 102.4GFLOPS.

Also SX-9 inherits the shared memory-type of programming model featured in past SX generations. By providing it with a maximum of 16 CPUs (1,638.4GFLOPS) and 1 Terabyte of shared memory nodes and by enabling the use of an automatic parallelization function by compilers, the development aimed at providing Teraflops-class performance.

(2) Pursuit of Operation and Migration Cost Reductions

CMOS technology, which has been adopted since the advent of the SX-4, has been further enhanced while LSI inte-

gration of hardware functions has achieved further advances. At the same time, high-density packaging technology has contributed to improvements in power consumption, required installation space and other factors related to installation, enabling further reductions in operation costs.

Also from the perspective of reducing software migration costs, while expanding and upgrading SX series architecture, NEC has maintained a superior level of software compatibility enabling customers to continue to use their existing program assets. In addition, SUPER-UX, which is the highly evaluated SX-series operating system, the operating environment, compilers and various tools such as the software development environment can be used as in the past, simplifying migration from both an operational and user perspective.

3.2 SX–9 Highlights

Based on the development concept described in the preceding Section 3.1, SX-9 has been created with the following features.

(1) World's Fastest Single-chip Vector Processor

Using state-of-the-art 65nm CMOS technology, SX-9 realizes 102.4GFLOPS per processor - the world's fastest class of processor performance.

(2) Multiprocessor System with Excellent Scalability

A single SX-9 node has a flat shared memory configuration for a maximum of 16 CPUs with peak single-node performance of 1,638.4GFLOPS, and maximum memory capacity of 1 Terabyte. Also this shared memory node is interconnected with a special high-speed switch at a maximum of 128Gbytes/sec. per node, enabling configurations of up to a maximum 512 nodes (838.8Teraflops) to provide a shared and distributed memory system with excellent scalability.

(3) Superior Energy-saving and Installation Advantages Achieved with Cutting-edge Technologies

In addition to reduction of processor energy consumption and heat generation through adoption of state-of-the-art LSI technology and an energy-saving power consumption design, a system that adopts high-efficiency cooling technology to counter the heat density layer generated by high-density packaging, improving low power consumption characteristics and installation capability.

(4) Proven SUPER-UX Operating System and Rich Software Compatibility

Exploiting experience in the design and production of the operating system, compilers and other fundamental software of the Earth Simulator, NEC has realized an excellent operation environment and software development environment for an ultra-large-scale configuration system. Also by maintaining the superior compatibility of the SX architecture, the SX-9 enables continued usage of the rich array of existing application software finely tuned to take full advantage of the SX series.

4 Hardware Overview

Fig. shows the SX-9 system configuration. The product lineup ranges from a shared-memory type single-node model that tightly integrates a maximum of 16 processors (CPU) and the main memory unit to a multi-node system model connects this single node to special high-speed nodes in a cluster configuration via a high-speed internode crossbar switch (IXS).

The single-node system is available in two models: "A" model with a maximum 16 CPUs (1,638.4GFLOPS), maximum main memory capacity of 1 Terabyte, and a maximum of 32 input/output slots, and "B" model with a maximum 8 CPUs (819.2GFLOPS), maximum main memory capacity of 512Gbytes, and a maximum of 16 input/output slots. Both boast a superior total balance of computational performance, memory throughput and input/output performance and deliver high effective performance.

The multi-node system interconnects 2 to 512 nodes in a cluster configuration by linking multiple single node systems

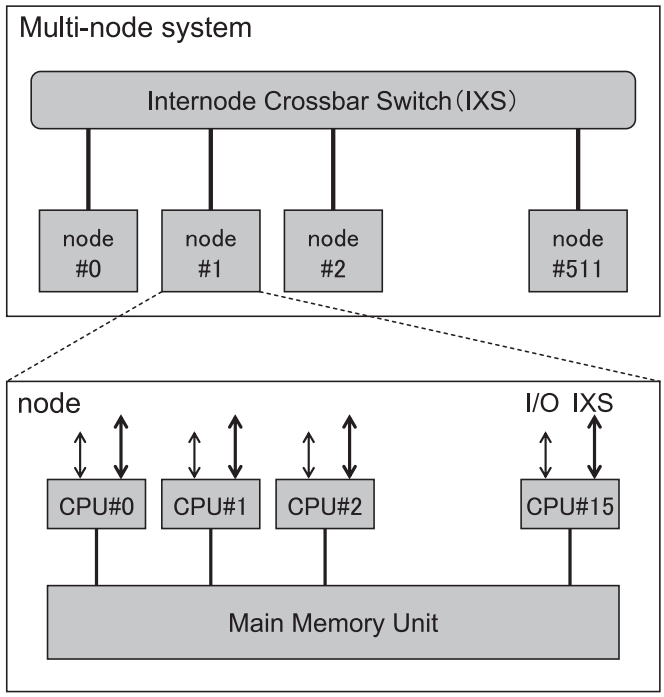

Fig. SX-9 system configuration.

Table Main specifications of the SX-9 System.

Item		SX-9(Single node)						SX-9(Multi- node)		
Model		4B	～	8B	8A	～	16A	32M2	～	8192M512
Processor (CPU)	Number of CPU	4	～	8	8	～	16	32	～	8192
	Peak Vector Performance	409.6GFLOPS	～	819.2GFLOPS	819.2GFLOPS	～	1638.4GFLOPS	3.27TFLOPS	～	838.86TFLOPS
Main Memory Unit (MMU)	Capacity	256GB/512GB			512GB/1024GB			2TB	～	512TB
	Max. data trafer rate	1TB/sec	～	2TB/sec	2TB/sec	～	4TB/sec	8TB/sec	～	2048TB/sec
Input-Output Feature (IOF)	Max. number of slots	8	～	16	16	～	32	64	～	16384
	Max. data transfer rate	32GB/sec	～	64GB/sec	64GB/sec	～	128GB/sec	256GB/sec	～	65536GB/sec
Internode switch	Max. data transfer rate	—						256GB/sec	～	65536GB/sec

via a high-speed internode crossbar switch (IXS) to realize vector computation performance of 839TFLOPS with a maximum of 8,192 CPUs (See **Table**).

5 Basic Software Overview

The SUPER-UX operating system has been developed to enable highly efficient utilization of the Supercomputer SX series hardware. Over the past 20 years, the operating system has been upgraded in step with advances in hardware. While seeing constant improvement in both performance and ease of use, it continues to boast compatibility with earlier versions and deliver stability of quality.

Incorporating powerful batch processing functions, a large-scale file system, job checkpoint/restart functions and automatic operation functions, SUPER-UX enables highly efficient operation management. In addition, it is provided with a comprehensive software development environment, realizing an easy-to-use operating system.

(1) Operating System Basic Functions

SUPER-UX is a UNIX System V port incorporating BSD functions and the addition of other enhanced functionality appropriate for its role as an operating system for a supercomputer. It not only supports large-scale multiprocessors and possesses efficient resource management functions to support multinode operation, but also boasts high-speed input/output functions and provides gang scheduling functions for parallel processing to extract maximum possible performance from the system. With the SX-9, the maximum size of the user's virtual space per process is expanded to 4 Terabytes, and enables the usage of expanded real memory of 1 Terabyte per process. As a result, the system makes the construction of large-scale applications even simpler.

(2) File Management Functions

SUPER-UX makes it possible to create large-scale files and large-scale file systems. While retaining the advantages of a standard UNIX system, SUPER-UX benefits from significantly upgraded functionality, realizing a high-speed file system and high-speed input/output capabilities. Provided as a high-speed shared file system, gStorageFS can realize data transfer without going through the CPU of remote servers with performance comparable to accessing the local file system. Also regarding the user interface, NFS V3 compatibility is maintained.

(3) Batch Processing Functions

For programs that require long hours of processing on a huge scale, SUPER-UX supports the concepts of jobs (sets of processes) at the kernel level. NQSII is the batch processing system that enables appropriate processing by large-scale clusters, and a JobManipulator is provided as an NQSII scheduler extension to maximize system operational efficiency by backfill scheduling. NQSII incorporates enhancements to major functions required for cluster system operation such as NQS job queuing, resource management and load balancing. SSI (single system image) is also enhanced, and system operability has been improved.

(4) Operation Management Functions

SUPER-UX supports a checkpoint/restart function that enables interruption of a program in progress at a point chosen at the user's discretion and its smooth resumption at a later time. Also by using operation management software to provide integrated management of multiple host machines on a network by a single machine and by an automatic operation control device, SUPER-UX contributes to system operation-related labor savings and unmanned operation.

(5) Software Development Environment

The compilers for Fortran, C and C++ provides high level optimization, automatic vectorization and automatic parallelization features to extract maximum performance from the SX-9. The vector data buffering function, which uses the hardware system called ADB (Assignable Data Buffer), is open to the user, enabling highly precise memory access optimization.

Also the MPI Library and the HPF compiler are provided for distributed memory programming to support comprehensive

programming environments for the large-scale multi-node user. It also provides the PSUITE tool for total support of program development, debugging and tuning using GUI, and de facto tools such as TotalView and ITA (Intel Trace Analyzer).

6 Storage System Overview

For the large-scale file system and conducting high-speed input/output of large-volume data, NEC has the iStorage D8, D3 and D1 line of SAN-compatible disk array devices with connectivity with the SX-9 series.

In addition to incorporating state-of-the-art technologies for the high performance and high expandability demanded by scientific computation, these products deliver the high reliability and availability required for stable operation of large-capacity systems. Also in response to the demand for reduced deployment costs, installation has been simplified, eliminating the need for specialized storage know-how.

7 Conclusion

This concludes our review of the history of the SX series, the development concept behind the SX-9 - the latest model and an overview of its hardware, basic software and storage system. In the future, NEC will continue to tackle the cutting edge of technology and drive the evolution of the SX series, introducing new enhancements and innovation to answer needs of customers for supercomputing systems that deliver higher performance and improved ease of use.

---

*UNIX is a registered trademark of The Open Group.

*NFS is a trademark of Sun Microsystems, Inc.

*TotalView is a registered trademark of TotalView Technologies.

*Intel Trace Analyzer is a registered trademark of Intel Corporation.