The SuperComputing System at Osaka University

OSAKA UNIVERSITY

Maintaining the vector-parallel supercomputer:

The Cybermedia Center, Osaka University, has been operating the NEC SX-5/128M8 system since 2001. This system was ranked 8th on the TOP500 list on June 21, 2001 as a historic system breaking the barrier of 1 Tflops even before the Japanese Earth Simulator (JAMSTEC/Earth Simulator Center) was put into operational use in 2002 (http://www.top500.org/2001/06/21/). This was a tremendous achievement, indeed.

Although the NEC SX-5/128M8 system is ranked lower now than its original ranking because of the spread of PC clusters, this SX system still enjoys wide support because of its high sustained performance for real application programs and the ease of programming enabled by the large-scale shared memory architecture. The Cybermedia Center has decided to upgrade this system with the same SX Series supercomputer for January 2007 by focusing on the system's significant leap in sustained performance.

Continuous Performance Enhancement of Vector Computers:

The introduction of the forthcoming system falls into two phases. In the first phase, a 20-node SX-8R system, an enhanced model of the current SX-8 (http://www.hpce.nec.com/47.0.html), will be installed. This system will consist of two sets of computers; one is the SX-8R/64M8 system with a 256GB SDRAM memory unit in each node interconnected by the IXS crossbar switch boasting a 16GB/sec bidirectional communication speed. In the second phase, another system will be made up of the node groups equipped with fast-cycle RAM units of 64GB or 128GB. The Cybermedia Center is also planning to introduce additional, next-generation systems for June 2008, a year and a half later through a phenomenal leap in performance per node. Gain in effective performance for the final configuration is expected to reach a factor of 16 in comparison with the current SX-5/128M8 system.

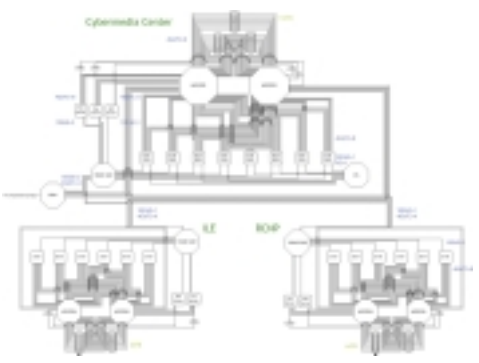
Each node introduced in the first phase is to be interconnected with a storage system through multiple, clustered 1GbE and 2GbFC networks. The 256TB storage system in the first phase is to be deployed at the ILE (Institute of Laser Engineering) and the RCNP (Research Center for Nuclear Physics), both of which are joint operation centers, making up a campus-wide storage area network. In the second phase, an additional 768TB storage system is to be installed, thus resulting in the total storage capacity of 1PB, as well as reinforced clustering interfaces (10Gbps).

Leap of Throughput Enabled by Coupling with PC Clusters and the Grid-compatible Operations:

As for the next system, dramatically improved throughput is expected for coupling simulations with PC clusters by harnessing 1PB storage as a core system. More than 700 nodes of the system will be introduced, including dedicated HPC nodes and shared nodes exploiting idle personal workstations for IT and language education purposes. These systems are to be configured as a clustering system using T310 10GbE TOE-NIC (Chelsio), E600i and S2410 (Force10). The clustering system is expected to realize a broadband and low-latency communication network that surpasses the performance of dedicated interfaces through the enhanced, general-purpose Ethernet technology and RDMA.

As for job management, the NEC NQSII system is capable of handling coupling simulations in a heterogeneous computing environment including the SX Series. NQSII supports the back-filling type job scheduling capability with a JobManipulator option. The fair-share type priority control mechanism is also offered with the ERS (Enhanced Resource Scheduler) option. While very effective system usability has been maintained at the Center with this feature, the highly multiplexed job operations also tend to cause an increase in job turnaround time against expectations by the user. The job manipulation combined with the back-filling type scheduling is expected to mitigate this negative effect without multiplexed job operations. Moreover, the prior allocation of specific job slots enables the concurrent use of Grid-type and conventional environments, and will be realized by the introduction of the NAREGI Grid middleware (beta 2 version) in 2007. This system is expected to become the first vector computer incorporated into the Grid environment, thus enabling seamless collaborations among users at different locations.

One of Initial Plan for Inter-Center Networking



CMC 大阪大学サイバーメディアセンター
Cybermedia Center, OSAKA University
Cybermedia Center, Osaka University
5-1 Mihogaoka Ibaraki Osaka 567-0047 JAPAN
http://www.lcmc.osaka-u.ac.jp/