



T2K - Tsukuba 95 TFLOPS Cluster System at CCS

Overview



T2K-Tsukuba is the latest supercomputer system installed at Center for Computational Sciences, University of Tsukuba, on June 2008. It is a very large scale PC cluster with 10,368 cores to provide 95.4 TFLOPS of peak performance. To execute wide variety of scientific computation, each computation node is configured as a fat-style PC server with 16 cores and 32 GB main memory. A large scale shared file system is provided by Lustre with 800 TB of user space under RAID6 physical file system, which is connected to all computation nodes via Infiniband to provide 16 GB/s of I/O bandwidth.

The operating system is RedHat Enterprise Linux WS5 and the entire cluster system is managed under ACE (Appro Cluster Engine), MVAPICH2 by Ohio State University, and Grid Engine batch job scheduler.

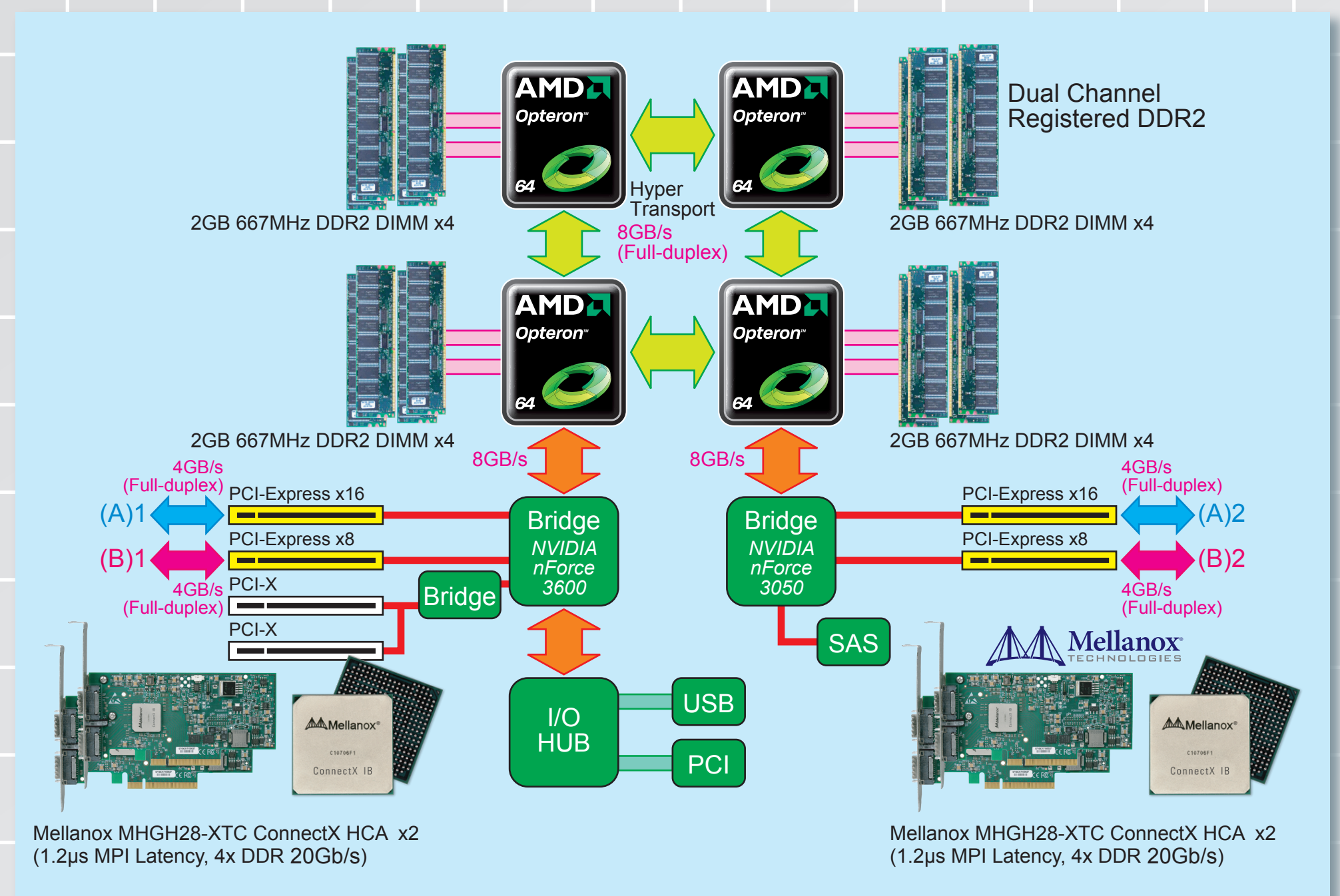
T2K-Tsukuba started its full operation from October 2008, and currently utilized for more than 30 of large scale computing projects. The system is ranked at 20th in TOP500 list on June 2008. As in November 2008, T2K-Tsukuba is the largest single-system cluster in Japan.

Alliance and Vendors

T2K-Tsukuba is designed and procured under T2K Open Supercomputer Alliance contracted with three Japanese universities; University of Tsukuba, the University of Tokyo and Kyoto University.

T2K-Tsukuba is a multi-vendor system; Appro XtremeServer-X3 as computation node, Mellanox MHGH28-XTC ConnectX as Infiniband HCA, Flextronics 24-port IB switch as interconnection switch, DDN S2A9550 RAID6 as shared file server, and the entire system is integrated and maintained by Cray Japan Inc.

Computation Node



Each computation node is equipped with 4-socket of quad-core Opteron (AMD Barcelona, 2.3GHz) in 16-core shared memory configuration with 147.2 GFLOPS of performance. To support such a high-performance node, quad-rail 4xDDR Infiniband HCA is employed with 16 GB/s (bidirection) of communication performance.

Interconnection Network

Each port of quad-rail Infiniband HCA (Mellanox ConnectX 4xDDR) is connected to a full-bisection bandwidth Fat-Tree under automatically configured fault-tolerant interconnection network. It consists of 24-port switches only and configured as 3-stage network. The total counts of switches and cables are 616 and 8554, respectively. This very powerful interconnection network supports very large scale parallel applications with up to 95 TFLOPS performance.

