Cray XT5™ Supercomputer

Cray XT5 System Highlights

Introducing the next revolution in scalable computing — the Cray XT5 supercomputer. The Cray XT5 system combines unprecedented scalability with exceptional manageability, lower cost of ownership, and broader application support. And, as the foundation for the Cray XT5$_h$™ — the industry's most integrated hybrid supercomputer — it creates a new paradigm in high performance computing by combining industry-leading scalar processing capability with high-bandwidth vector processing, reconfigurable FPGA hardware acceleration and alternative parallel programming languages in a single system.

Flexible Processor Configurations

The Cray XT5 cabinet can accommodate both the Cray XT4™ and Cray XT5 compute blades, creating system configurations matched to application requirements.

Designed for Scalability

The Cray XT5 system is designed from the ground-up for extreme scalability, including highly scalable global I/O performance ensuring high efficiency for applications that require rapid I/O access to large datasets.

Linux Environment

A flexible Linux-based operating system makes it easier for a wide variety of applications to benefit from superior scalability. The Linux environment enables streamlined porting of a broad set of ISV codes.

Higher Efficiency, Lower Operating Cost

Superior energy efficiency and lower operating costs through innovative packaging and technologies that reduce power and cooling requirements, reducing energy consumption and operating costs.

Competitive Price/Performance

Competitive price/performance with commodity clusters, while providing superior interconnect, bandwidth, upgradeability, manageability, and scalability.

Cray XT5 Supercomputer

Engineered to meet the demanding needs of capability-class HPC applications, each feature and function is selected in order to enable larger problems, faster solutions, and a greater return on investment. Designed to support the most challenging HPC workloads, the Cray XT5 supercomputer delivers scalable power for the toughest computing challenges.

Scalable Architecture

The Cray XT5 system's 3D torus architecture is designed for superior application performance for large-scale, massively parallel computing. This is accomplished by incorporating two types of dedicated nodes — compute nodes and service nodes. Compute nodes are designed to run MPI tasks efficiently and reliably to completion. Each compute node is composed of one or two AMD Opteron microprocessors (dual or quad core) and direct attached memory, coupled with a dedicated communications resource. Service nodes are designed to provide system and I/O connectivity and also serve as login nodes from which jobs are compiled and launched.

Scalable Compute Nodes

The basic building block of the Cray XT5 system is a compute node. Each compute node is composed of either one AMD Opteron (the Cray XT4 compute blade) or two AMD Opteron processors (the Cray XT5 compute blade), each coupled with its own memory and dedicated communication resource. Cray XT4 blades are optimized for compute and interconnect balance while the new Cray XT5 blades are optimized for memory-intensive and compute-biased workloads. This design eliminates the scheduling complexities and asymmetric performance problems associated with clusters of large SMPs. It also ensures that performance is uniform across distributed memory processes — an absolute requirement for scalable algorithms.

Scalable Interconnect

The Cray XT5 system incorporates a high-bandwidth, low-latency interconnect based on the Cray SeaStar2+™ chip. The interconnect directly connects all the nodes in a Cray XT5 system in a 3D torus topology, eliminating the cost and complexity of external switches and allowing for easy expandability. This allows systems to economically scale to tens of thousands of nodes—well beyond the capacity of fat-tree switches. As the backbone of the Cray XT5 system, the interconnect carries all message passing traffic as well as I/O traffic to the global file system.

Cray SeaStar2+ Architecture
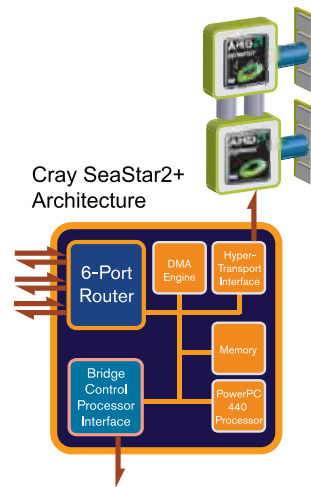
Designed for scalable MPI computing, the Cray SeaStar2+ chip combines communications processing and high-speed routing on a single device. Each communications chip is composed of a HyperTransport™ link, a Direct Memory Access (DMA) engine, a communications and management processor, a high-speed interconnect router, and a service port.

The interconnect router in the Cray SeaStar2+ chip provides six high-speed network links which connect to six neighbors in the 3D torus. The peak bidirectional bandwidth of each link is 9.6 GB/s with sustained bandwidth in excess of 6 GB/s.

Each port is configured with an independent router table, ensuring contention-free access for packets. The router is designed with a reliable link-level protocol with error correction and retransmission, ensuring that message passing traffic reliably reaches its destination without the costly timeout and retry mechanism used in typical clusters. A doubling of the number of virtual channels provides up to a 30% increase in sustained global bandwidth compared to previous generation Cray SeaStar™ routers.

Scalable Software

The Cray XT5 operating environment is designed to run large, complex applications and scale efficiently to more than 240,000 processor cores. The Linux environment features a compute kernel which can be configured to match different workloads. When running highly scalable custom applications, the compute nodes can be run in a lightweight mode, ensuring that operating system services do not interfere with application scalability. This special design ensures that there is virtually nothing that stands between the user's scalable application and the hardware.

Cray XT5 Scalable Architecture

The Cray XT5 architecture directly connects all AMD Opteron processors, removing PCI bottlenecks and shared memory contention to deliver superior sustained application performance at massive scale.
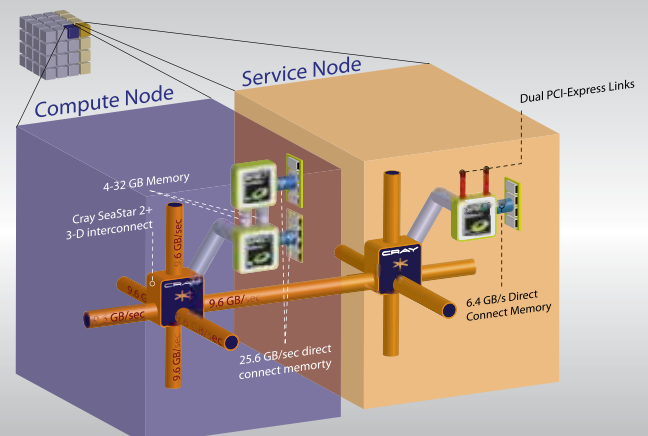
When running ISV applications, the compute nodes can be configured to run a more compatible compute node OS, complete with the necessary services to ensure application compatibility.

Linux Environment
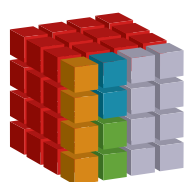
Compute Node
Login Node
Network Node
System Node
I/O Node

Users can submit jobs interactively from login nodes using the Cray XT5 job launch command or through the PBS Pro™ batch program, which is tightly integrated with the system scheduler. The system provides accounting for parallel jobs as single entities with aggregated resource usage.

The Cray XT5 system maintains a single root file system across all nodes, ensuring that modifications are immediately visible throughout the system without transmitting changes to each individual node. Fast boot times ensure that software upgrades can be completed quickly, with minimal downtime.

Programming Environment

The Cray XT5 programming environment includes tools designed to complement and enhance each other, resulting in a rich, easy-to-use programming environment that facilitates the development of scalable applications.

Parallel programming models supported include MPI, SHMEM, UPC, and OpenMP within the node. The MPI implementation is compliant with the MPI 2.0 standard and is optimized to take advantage of the scalable interconnect in the Cray XT5 system.

The Cray XT5 system is compatible with a vast quantity of existing compilers and libraries, including optimized C, C++, and Fortran90 compilers, as well as high-performance optimized math libraries of BLAS, FFTs, LAPACK, ScaLAPACK, SuperLU, and Cray Scientific Libraries.

Cray Apprentice2™ performance analysis tools allow users to analyze resource utilization throughout their code at scale and eliminate bottlenecks and load-imbalance issues.

Scalable RAS & Administration

The Cray RAS and Management System (CRMS) integrates hardware and software components to provide system monitoring, fault identification, and recovery. An independent system with its own control processors and supervisory network, the CRMS monitors and manages all of the major hardware and software components in the Cray XT5 system. In addition to providing recovery services in the event of a hardware or software failure, CRMS controls power-up, power-down, and boot sequences, manages the interconnect, and displays the machine state to the system administrator.

Scalable I/O

The Cray XT5 I/O subsystem scales to meet the bandwidth needs of even the most data-intensive applications. The I/O architecture consists of storage arrays connected directly to I/O nodes which reside on the high-speed interconnect. The Lustre file system manages the striping of file operations across these arrays. This highly scalable I/O architecture allows customers to configure the Cray XT5 with the desired bandwidth by selecting the appropriate number of arrays and service nodes.

Superior Energy Efficiency, Lower Operating Costs

Recognizing the growing need to reduce energy usage and control operating costs, the Cray XT5 family employs innovative packaging technologies and an efficient power conversion train that reduces energy use and total cost of ownership.

System compute blades are packaged with only the necessary components needed for massively parallel processing – processors, memory and interconnect.

In a Cray XT5 cabinet, vertical cooling takes cold air straight from its source — the floor — and efficiently cools the processors on the blades, which are uniquely positioned for optimal airflow. Each processor also has a custom-designed heat sink depending on its position within the cabinet. Each Cray XT5 system cabinet is cooled with a single, high-efficiency ducted turbine fan. It also can take 400/480VAC directly from the power grid without transformer and PDU loss — further contributing to reduced energy usage and reduced cost of ownership.

Existing Cray XT systems can be easily upgraded or expanded to take advantage of the new technologies in the Cray XT5 product.

Finally, because the Cray XT5 offers better scalability and sustained performance than typical clusters, a given level of application performance can be typically achieved with fewer processors – compounding the savings.

CRAY
THE SUPERCOMPUTER COMPANY

Cray Inc.
411 First Avenue S., Suite 600
Seattle, WA 98104-2860 USA
Tel (206) 701 2000
Fax (206) 701 2500
www.cray.com

Sales Inquiries:
North America: 1 (877) CRAY INC
Worldwide: 1 (651) 605 8817
sales@cray.com