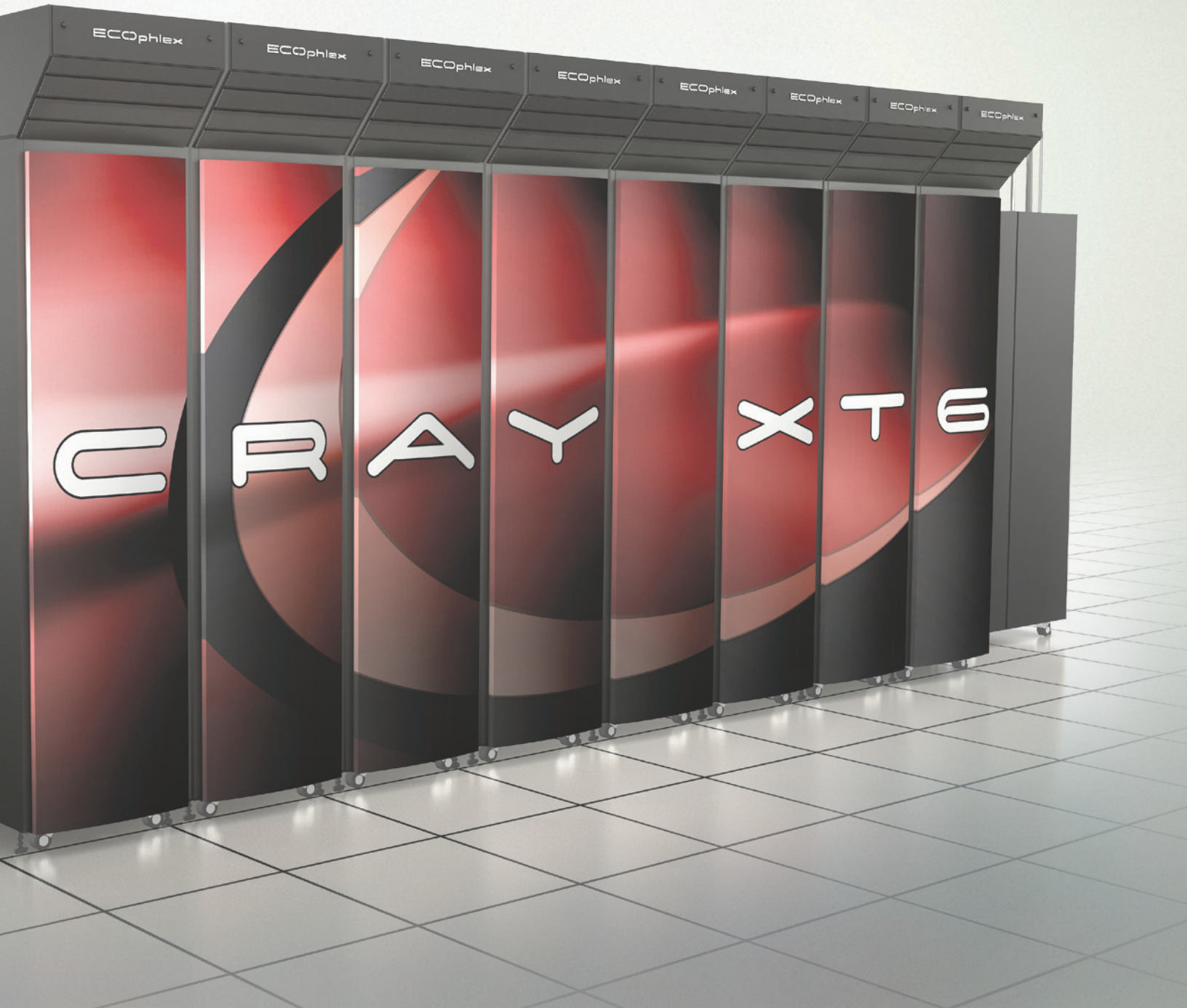
CRAY XT6

Introducing the Cray XT6™ supercomputer, the next evolution of the Cray family that was first to break the petascale performance barrier on real-world applications. Engineered to meet the demanding needs of capability-class High Performance Computing (HPC) applications, each feature and function is selected to enable larger datasets, faster solutions and a greater return on investment. Designed to support the most challenging HPC workloads, the Cray XT6 supercomputer delivers scalable computing power to solve the toughest research and commercial challenges.



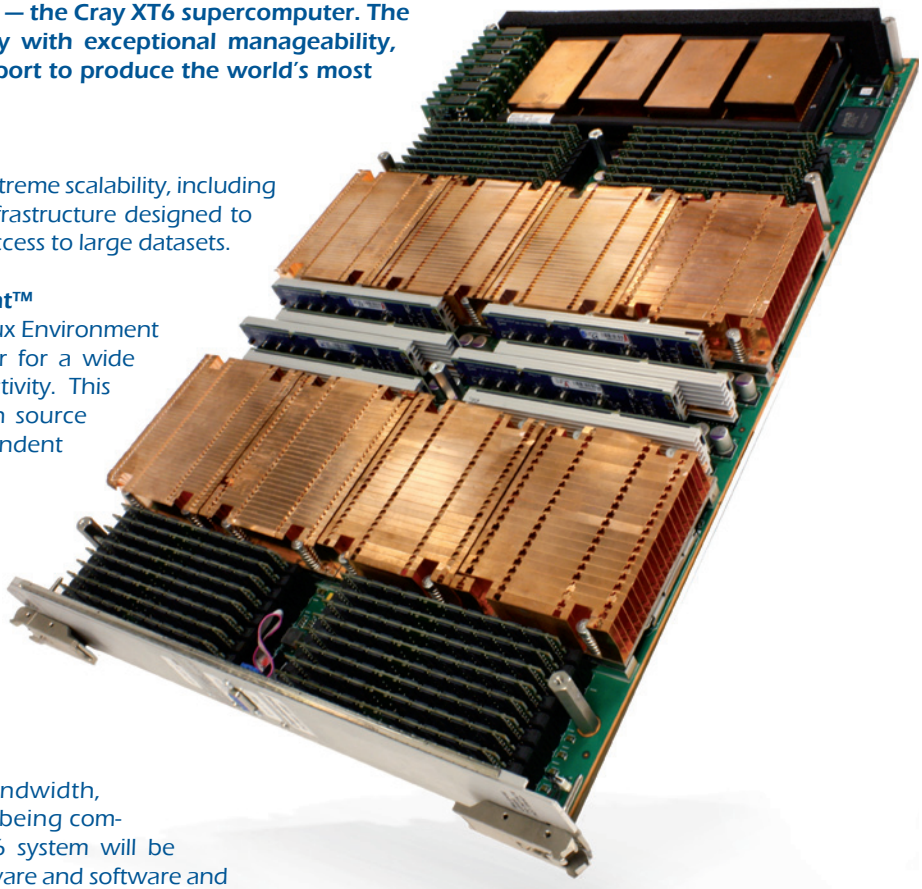
Cray XT6 Specifications

CPU	Eight or 12-core 64-bit AMD Opteron 6100 series processors; up to 192 per cabinet
Cache	64K L1 instruction cache, 64K L1 data cache, 512 KB L2 cache per processor core, 12 MB shared L3 cache
Cores	1,536 or 2,304 processor cores per system cabinet
Peak Performance	12.2 to 20.2 Teraflops per system cabinet
Main Memory	32 GB or 64 GB registered ECC DDR3 SDRAM per compute node
Memory Bandwidth	85.3 GB/sec per compute node
Interconnect	1 Cray Seastar2+ routing and communications ASIC per compute node
	6 switch ports per Cray Seastar2+ chip, 9.6 GB/sec each (57.6 GB/sec switching capacity per Cray Seastar2+ chip)
	3D torus interconnect
External I/O Interface	Gigabit Ethernet
	10 Gigabit Ethernet
	Fibre Channel (FC)
	InfiniBand
Disk Storage	Full line of FC-attached disk arrays with support for FC and SATA disk drives
File System	Lustre, Data Virtualization Service (DVS) allows support for NFS, external Lustre and other file systems
System Administration	Cray System Management Workstation (SMW)
	Graphical and command line system administration
	Single-system view for system administration
	PBS Professional ™ job management system Adaptive Computing Moab job management system
	System software rollback capability
Reliability Features (Hardware)	Cray Hardware Supervisory System (HSS) with independent 100 Mb/s management fabric between all system blades and cabinet-level controllers
	More than 50 measurement points monitored per Cray XT6 system blade
	Full ECC protection in the Cray SeaStar2+ chip
	Redundant power supplies; redundant voltage regulator modules (VRMs)
	Redundant paths to all system RAID
	Variable-speed axial turbofan with integrated pressure and temperature sensors
Reliability Features (Software)	HSS system monitors operation of all operating system kernels
	Lustre file system object storage target failover; Lustre metadata server failover
	Software failover for critical system services including system database, system logger and batch subsystems
	NodeKARE™ - Node Knowledge and REconfiguration
Operating System	Cray Linux Environment: components include SUSE Linux™, HSS and SMW software
	Extreme Scalabiliity Mode (ESM) and Cluster Compatibility Mode (CCM)
Message Passing Libraries	MPI 2.0, Cray SHMEM, other standard MPI libraries using CCM
Compilers	PGI™ Compilers, Cray Compiler Environment, Pathscale: support for Fortran 77, 90, 95; C/C++, UPC, Co-Array Fortran
Power	45 - 54.1 kW (45.9 – 55.2 kVA) per cabinet, depending on configuration Circuit requirements: 3 phase wye, 100 AMP at 480/277 and 125 AMP at 400/230 (3 phase, neutral and ground)
Cooling Requirement	Air-cooled, air flow: 3,000 cfm (1.41 m3/s); intake: bottom; exhaust: top
	Optional ECOphlex liquid cooling
Dimensions (Cabinet)	H 93 in. (2,362 mm) x W 22.50 in. (572 mm) x D 56.75 in. (1,441 mm)
Weight (Maximum)	1,600 lbs. per cabinet (725 kg) air cooled; 2,000 lbs. per cabinet (907 kg) liquid cooled
Acoustical Noise Level	75 dba at 3.3 ft (1.0 m)
Regulatory Compliance	UL 60950-1, CAN/CSA – C 22.2 No. 60950-1, CE-mark, RoHS, WEEE
Safety	FCC Class A, VCCI Class A, ICES-003, EN 50022:2006 Class A, AS/NZS CISPR 22:2006, EN 55024: 1998 +A1:2002 +A2:2003

CRAY

Cray XT6 System Highlights

Introducing the next revolution in scalable computing — the Cray XT6 supercomputer. The Cray XT6 system combines unprecedented scalability with exceptional manageability, lower cost of ownership and broader application support to produce the world's most productive general purpose supercomputer.



Design for Scalability

The Cray XT6 system is designed from the ground up for extreme scalability, including a low contention interconnect and parallel global I/O infrastructure designed to ensure high performance for applications and rapid I/O access to large datasets.

Reliable and Easy to Manage – Cray Linux Environment™

The flexible, scalable and performance-optimized Cray Linux Environment (CLE) is included with every system and makes it easier for a wide variety of applications to benefit from superior productivity. This environment also enables use of a wide range of open source tools and streamlined porting of a broad set of Independent Software Vendor (ISV) applications.

Higher Efficiency, Lower Operating Cost

Cray's High Efficiency cabinets with optional ECOphlex™ (PHase-change Liquid EXchange) cooling provide innovative power, cooling and packaging with exceptional energy efficiency, lowering datacenter cooling requirements and the total cost of ownership.

Price, Performance and Upgradeability

The Cray XT6 system provides superior interconnect, bandwidth, upgradeability, manageability and scalability while still being competitively priced with commodity clusters. The Cray XT6 system will be upgradeable with future planned enhancements to hardware and software and current Cray XT5™ supercomputer customers can upgrade to the Cray XT6 system.

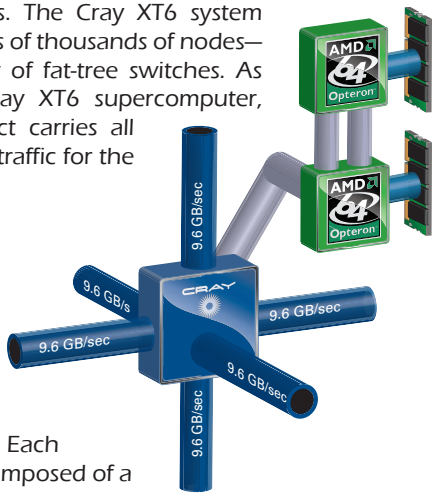
A Scalable System by Design

The Cray XT6 system is scalable by design. Extreme scalability is achieved by optimizing each component of the system (compute, network, software, infrastructure and IO). Cray provides two types of dedicated nodes — compute nodes and service nodes. Compute nodes are optimized to run parallel MPI and/or OpenMP tasks with maximum efficiency. Service nodes provide scalable system and I/O connectivity and can serve as login nodes from which applications are compiled and launched. Cray provides fully integrated networking, using an efficient, low-contention three-dimensional (3D) torus architecture, designed for superior application performance for large-scale, massively parallel applications.

This design eliminates the scheduling complexities and asymmetric performance problems associated with common cluster designs. It also ensures that performance is uniform across distributed memory processes — an absolute requirement for scalable algorithms.

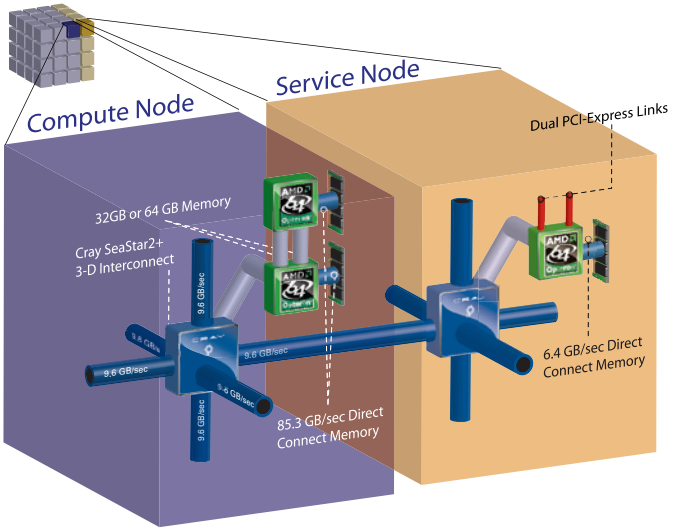
Scalable Interconnect

The Cray XT6 supercomputer incorporates a high-bandwidth, low-latency interconnect based on the Cray SeaStar2+™ chip and has been proven to scale systems beyond two petaflops. All the nodes in a Cray XT6 system are directly connnected in a 3D torus topology, eliminating the inherent inefficiencies of systems that use external switches. The Cray XT6 design allows for easy expandability and is upgradable to future high performance Cray interconnect technologies. The Cray XT6 system economically scales to tens of thousands of nodes—well beyond the capacity of fat-tree switches. As the backbone of the Cray XT6 supercomputer, this powerful interconnect carries all message passing and I/O traffic for the system.



Designed to optimize MPI message passing, 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/sec with sustained bandwidth in excess of 6 GB/sec.

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. The Cray SeaStar communication protocol is connectionless, thus eliminating the need to create and cache queue structures between communicating node pairs. This enables full application scalability to large node counts, even on challenging applications which stress irregular communication patterns.

Cray Software Ecosystem

Each Cray XT6 system supports a complete software ecosystem that includes the latest Cray Linux Environment v3 (CLE3), job schedulers including Altair PBS Professional™, Adaptive Computing Moab™ or Platform LSF™, compilers from PGI®, Pathscale™, and Cray, debuggers from Totalview Technologies™ and Alinea™, many open source programming tools and the integrated Cray Programming Environment (CPE).

Cray has recently announced availability of CLE3, a suite of high performance software which includes a Linux-based operating system based on SUSE Linux™, designed to run large, complex applications and scale efficiently to more than a million processor cores. The Linux environment features a compute kernel which can be configured to match different workloads. When running highly scalable applications, the CLE3 runs by default in Extreme Scalability Mode (ESM), which ensures 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 and has been proven in real world applications to scale to more than 200,000 cores.

CLE3 also includes Cluster Compatibility Mode (CCM). CCM allows the running of most ISV applications out-of-the-box, without recompilation or re-linking and allows for the use of various versions of MPI (MPICH™, Platform MPI™, etc.). At job submission, the compute nodes are configured with CCM to run a more cluster compatible compute node Linux OS, complete with the necessary services to ensure application compatibility. When the application is finished, all nodes are returned to their native ESM state.

Jobs are submitted to the Cray XT6 supercomputer through batch programs such as Altair PBS Professional or Adaptive Computing Moab, which are tightly integrated with the system scheduler, interactively using the Cray XT6 job launch command. The system provides accounting for parallel jobs as single entities with aggregated resource usage.

Each Cray XT6 system includes a fully integrated Cray Programming Environment with tools designed to maximize programmer productivity, and application scalability and performance. This feature-rich, easy-to-use programming environment facilitates the development of scalable applications. Parallel programming models supported include MPI, Cray SHMEM™, UPC Co-Array Fortran and OpenMP. The MPI implementation is compliant with the MPI 2.0 standard and is optimized to take advantage of the scalable interconnect in the Cray XT6 system.

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

The Cray XT6 supercomputer can utilize a wide variety of high performance compilers and libraries, including PGI®, Pathscale™ and the Cray Compiler Environment with support for optimized C, C++, and Fortran90, UPC and Co-Array Fortran, as well as high performance optimized math libraries of BLAS, FFTs, LAPACK, ScaLAPACK, SuperLU, and Cray Scientific Libraries.

Scalable RAS & Administration

Cray XT6 systems are built for reliability and installed systems typically have greater than 99% availability. Key to this reliability are Cray's integrated Hardware Supervisory System (HSS) and innovations in CLE3, including our exclusive NodeKARE™ (Node Knowledge and Reconfiguration) functionality. Cray HSS 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 HSS monitors and manages all of the major hardware and software components in the Cray XT6 supercomputer. In addition to providing recovery services in the event of a hardware or software failure, HSS controls power-up, power-down and boot sequences, manages the interconnect, and displays the machine state to the system administrator.

CLE3 features NodeKARE. Should a user's program terminate abnormally, NodeKARE automatically runs diagnostics on all compute nodes involved in the application removing any unhealthy nodes from the compute pool. This ensures that subsequent jobs are allocated only healthy nodes and run reliably to completion.

Scalable I/O

The Cray XT6 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 XT6 supercomputer with the desired bandwidth by selecting the appropriate number of arrays and service nodes. Additionally, the Cray Data Virtualization Service (DVS) allows for the projection of various other file systems (including NFS, GPFS™, Panasas® and StorNext®) to the compute and login nodes on the Cray XT6 supercomputer.

Superior Energy Efficiency, Lower Operating Costs

Recognizing the growing need to reduce energy usage and control operating costs, the Cray XT6 family employs innovative packaging technologies and an efficient power conversion train that reduces energy use and total cost of ownership.

Cray XT6 supercomputers provide state-of-the-art datacenter flexibility. Each system can be air- or liquid-cooled. In an air-cooled configuration, a single, high-efficiency ducted turbine fan draws cold air straight from its source - the floor - and efficiently cools the processors on the blades, which are precisely positioned for optimal airflow. This design offers unparalleled processor density, using less air per watt than other air-cooled configurations.

The optional Cray ECOphlex technology can dramatically reduce operating costs associated with cooling and provide flexibility in datacenter design and implementation. Each High Efficiency cabinet can be configured with in-line phase-change evaporator coils which effectively extracts virtually all the heat imparted to the airstream as it passes through the cabinet. Coolant is re-condensed in a Heat Exchange Unit (HEU) which is connected to the building chilled water supply. Because a flexible range of building water temperatures is permitted, a modern data center using ECOphlex technology can operate chillers and air handlers much less often, reducing electrical costs. In many climates, cooling towers alone are all that is needed to keep a system fitted with ECOphlex operating at full capacity during much of the year.

Cray XT6 compute blades are designed for maximum power efficiency, with only the necessary components needed for massively parallel processing – processors, memory and interconnect. The 400/480VAC power supplies in each cabinet connect directly from the power grid without transformer and power distribution unit loss, further contributing to reduced energy usage and lower cost of ownership.

Cray XT6 supercomputers allow for easy, flexible upgrade options which let customers prolong the productive lifetime of their investment. The Cray XT6 system can be upgraded or expanded to take advantage of new technologies. Cray XT6 supercomputer owners are well positioned to take advantage of next-generation compute processors, I/O technologies and interconnect as they become available, without replacing the entire Cray XT6 system.

Finally, because the Cray XT6 supercomputer provides better scalability, it can achieve a given level of sustained application performance with fewer processors than a simple commodity cluster. This compounds the value and savings that the Cray XT6 system offers.

Scalable Compute Nodes

Each Cray XT6 blade includes four compute nodes for high scalability in a small footprint – up to 96 processor cores per blade or 2,304 processor cores per cabinet. Each compute node is composed of two AMD Opteron™ processors (eight or 12-core), each coupled with its own memory and dedicated Seastar2+ communication ASIC. Each compute node is designed to efficiently run up to 24 MPI tasks, or alternately can be programmed to run OpenMP within a compute node and MPI between nodes.

The AMD processor's on-chip and highly associative data cache supports aggressive out-of-order execution. The integrated memory controller eliminates the need for separate Northbridge memory chip, and provides a high-bandwidth path to local memory – 85 GB/sec per dual-socket compute node or over 8 TB/sec per system cabinet. This design brings a significant performance advantage to algorithms that stress local memory bandwidth and plenty of headroom for future processor upgrades. HyperTransport™ technology enables a 6.4 GB/sec direct connection between the compute node and the Cray Seastar2+ interconnect, removing the PCI bottleneck inherent in commodity networks.

Each Cray XT6 node can be configured with 32 GB or 64 GB DDR3 memory. Memory on compute nodes is registered and memory controllers provide for the additional protection of x4 device correction, ensuring highly reliable memory performance while retaining the upgradeability, serviceability and flexibility of a socketed component.

The Cray XT6m™ supercomputer is based on the same Cray XT6 supercomputer architecture and software stack used in some of the world's largest supercomputer systems, but with an optimized network for system sizes ranging from 700 to 13,000 processing cores (or one to six cabinets).

CRAY
THE SUPERCOMPUTER COMPANY

Cray Inc. • 901 Fifth Avenue, Suite 1000 • Seattle, WA 98164 • Tel: 206.701.2000 • Fax: 206.701.2500