Introducing the next generation of affordable and productive massively parallel processing (MPP) computing – the Cray XE6m™ supercomputer. Building on the reliability and scalability of the Cray XE6™ supercomputer and using the same proven petascale technologies, these new systems are optimized to support scalable application workloads in the 6.5 teraflop to 200 teraflop performance range, where applications require between 400 and 18,000 cores of processing power.
Cray XE6m Supercomputer
Engineered to meet the demanding needs of capability-class HPC applications, each feature and function is selected in order to enable larger datasets, faster solutions and a greater return on investment. Designed to support the most challenging HPC workloads in a cost effective midrange system, the Cray XE6m supercomputer delivers scalable power for the toughest computing challenges.

Scalable Performance
The Cray XE6m system incorporates the same performance optimized Gemini interconnect found in Cray’s petascale systems which provides superior interconnect, bandwidth, upgradeability and manageability while remaining competitively priced with commodity clusters.

Price, Performance and Upgradeability
The Cray XE6m system provides affordable petascale technologies for truly scalable performance in a fully upgradeable infrastructure that will conserve this HPC investment for years to come.

Reliable and Easy to Manage
Cray Linux Environment™
The flexible, scalable and performance-optimized Cray Linux Environment (CLE) makes it easier for a wide variety of applications to benefit from superior scalability. This environment enables use of a wide range of open source tools and streamlined installation of a broad set of independent software vendor (ISV) applications.

Optimized Architecture
The Cray XE6m system utilizes a two-dimensional (2D) torus architecture, optimized for superior application performance between 400 and 18,000 processing cores. It incorporates two types of dedicated nodes — compute nodes and service nodes. Compute nodes are designed to run parallel MPI and/or OpenMP tasks with maximum efficiency. Each compute node is composed of two AMD Opteron™ microprocessors (16 core) and direct attached memory, coupled with a dedicated communications resource. The Cray XE6m is also built to support future compute nodes which include GPU accelerators. Service nodes are designed to provide scalable system and I/O connectivity and can serve as login nodes from which applications are compiled and launched. 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.

Gemini Scalable Interconnect
The Gemini interconnect is the heart of the Cray XE6m system. Capable of tens of millions of MPI messages per second, the Gemini ASIC is designed to complement current and future massively multi-core processors. Each dual-socket node is interfaced to the Gemini interconnect through HyperTransport™ 3.0 technology. This direct connect architecture bypasses the PCI bottlenecks inherent in commodity networks and provides a peak of over 20 GB/s (10 GB/sec per direction) of injection bandwidth per node. The Gemini router’s connectionless protocol scales from hundreds to hundreds of thousands of cores without the increase in buffer memory required in the point-to-point connection method of commodity interconnects.
The Cray XE6m network provides industry-leading sub-microsecond latency for remote puts and one- to two-microsecond latency for most other point-to-point messages. An internal block transfer engine is available to provide high bandwidth and good overlap of computation and communication for long messages. Advanced features include support for one-sided communication primitives and support for atomic memory operations. The proven 2D torus topology provides powerful bisection and global bandwidth characteristics as well as support for dynamic routing of messages.

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. The Gemini 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 XE6m system supports a complete software ecosystem that includes the latest Cray Linux Environment (CLE), job schedulers including Altair PBS Professional™ and Adaptive Computing Moab™, compilers from PGI®, PathScale™ and Cray, debuggers from Rogue Wave TotalView™ and Allinea, many open source programming tools and the integrated Cray Programming Environment (CPE).

Cray CLE, a suite of high performance software which includes an operating system based on SUSE Linux™ is designed to run large, complex applications and scale efficiently to more than one million processor cores. The Linux environment features a compute kernel which can be configured to match different workloads. When running highly scalable applications, CLE 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 virtually nothing stands between the user’s scalable application and the hardware, and it has also been proven in real world application to scale to more than 200,000 cores.

CLE also includes Cluster Compatibility Mode (CCM). CCM allows most parallel ISV applications to run out-of-the-box without recompilation or relinking 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 XE6m 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 XE6m job launch command. The system provides accounting for parallel jobs as single entities with aggregated resource usage.

Each Cray XE6m supercomputer includes a fully integrated CPE with tools designed to enhance 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 advantageous to algorithms that stress local memory bandwidth and provides significant head room to accommodate future processor upgrades. HyperTransport technology enables a 20 GB/s direct connection between the compute node and the Gemini interconnect, removing the PCI bottleneck inherent in commodity networks.

Each Cray XE6m 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.

**Scalable Compute Nodes**

Each Cray XE6m blade includes four compute nodes for high scalability in a small footprint – up to 96 processor cores per blade or 3,072 processor cores per cabinet. Each compute node is composed of two AMD Opteron processors (16-core), each coupled with its own memory and dedicated Gemini communication ASIC. Each compute node is designed to efficiently run up to twenty-four 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 a separate Northbridge memory chip, and provides a high-bandwidth path to local memory – 85.3 GB/sec per dual-socket compute node. This design brings a significant performance advantage to algorithms that stress local memory bandwidth and provides significant head room to accommodate future processor upgrades. HyperTransport technology enables a 20 GB/s direct connection between the compute node and the Gemini interconnect, removing the PCI bottleneck inherent in commodity networks.

Each Cray XE6m 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.
Higher Efficiency, Lower Operating Cost
Cray’s high efficiency cabinets with optional ECOphlex cooling provide innovative power, cooling and packaging with superior energy efficiency, lowering data center cooling requirements and the total cost of ownership.

optimized to take advantage of the scalable interconnect in the Cray XE6m 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 XE6m 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.

Efficient System Administration
Cray XE6m systems are built for reliability and installed systems typically have greater than 99 percent availability. Key to this reliability are Cray’s integrated Hardware Supervisory System (HSS) and innovations in CLE, including our exclusive NodeKARE® (Node Knowledge and Reconfiguration) functionality.

Cray XE6m 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 XE6m system. 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.

CLE 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.

High Performance, Parallel I/O
The Cray XE6m 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 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 XE6m supercomputer with the desired bandwidth by selecting the appropriate number of arrays and service nodes. Additionally the Cray Data Virtualization Service (DVS), an integrated feature in CLE, allows for the projection of various other file systems (including NFS, GPFS®, Panasas® and StorNext®) to the compute and login nodes on the Cray XE6m system.

Superior Energy Efficiency, Lower Operating Costs
Recognizing the growing need to reduce energy usage and control operating costs, the Cray XE6m supercomputer employs innovative packaging technologies and an efficient power conversion train that reduces energy use and total cost of ownership.

Cray XE6m systems 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 uniquely positioned for optimal airflow. This design offers unparalleled processor density, using less air per watt than other air-cooled configurations.

The optional Cray ECOphlex® (PHase-change Liquid EXchange) technology can dramatically reduce the 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 extract virtually all the heat imparted to the airstream as it passes through the cabinet. Coolant is recondensed in a heat exchange unit (HEU) which is connected to the building’s chilled water supply. Because a flexible range of building water temperatures is permitted, a modern datacenter 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 XE6m compute blades are designed for maximum power efficiency, with only the necessary components needed for massively parallel processing – processors, memory and interconnect. The 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 XE6m supercomputers allow for easy, flexible upgrade options which lower the overall cost of ownership and increase the competitive lifespan of systems. Like the Cray XE6 system, the Cray XE6m midrange supercomputer can be upgraded or expanded to take advantage of new technologies. Cray XE6m system owners are well positioned to take advantage of next-generation compute processors, accelerator blades and I/O technologies as they become available, without replacing the entire system.
### Cray XE6m Specifications

| Processor          | 16-core 64-bit AMD Opteron 6300 Series processors, up to 192 per cabinet; NVIDIA Tesla X2090 GPU, up to 96 per cabinet, depending on configuration  
|                   | 8x64 KB L1 instruction cache, 16x16 KB L1 data cache, 8x2 MB L2 cache per processor core, 2x8 MB shared L3 cache  |
| Memory             | 32 GB or 64 GB registered ECC DDR3 SDRAM per compute node  
|                    | Memory Bandwidth: 102.4 GB/s per compute node  |
| Compute Cabinet    | 3072 processor cores per system cabinet  
|                    | Peak Performance: up to 30.7 teraflops per system cabinet  |
| Interconnect       | 1 Gemini routing and communications ASIC per two compute nodes  
|                    | 48 switch ports per Gemini chip, [160 GB/s internal switching capacity per chip]  
|                    | 2-D torus interconnect  |
| System Administration | Cray System Management workstation  
|                    | Graphical and command line system administration  
|                    | Single-system view for system administration  
|                    | 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  
|                    | Full ECC protection of all packet traffic in the Gemini network  
|                    | Redundant power supplies; redundant voltage regulator modules  
|                    | Redundant paths to all system RAID  
|                    | Warm-swap compute blades  |
| 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 SLES 11, HSS and SMW software)  
|                    | Extreme Scalability Mode (ESM) and Cluster Compatibility Mode (CCM)  |
| Compilers, Libraries & Tools | PGI compilers, Cray Compiler Environment, PathScale  
|                    | Support for Fortran 77, 90, 95; C/C++, UPC, Co-Array Fortran  
|                    | MPI 2.0, Cray SHMEM, other standard MPI libraries using CCM  
|                    | Cray Apprentice, Cray PAT and Cray Compiler included with systems  |
| Job Management     | PBS Professional job management system  
|                    | Moab Adaptive Computing Suite job management system  |
| External I/O Interface | InfiniBand, 10 Gigabit Ethernet, Fibre Channel (FC) and Ethernet  |
| Disk Storage       | Full line of FC-attached disk arrays with support for FC and SATA disk drives  |
| Parallel File System | Lustre, Data Virtualization Service allows support for NFS, external Lustre and other file systems  |
| Power (maximum)    | 45-54.1 kW (45.9 – 55.2 kVA) per cabinet, depending on configuration  
|                    | Circuit requirements: three-phase wye, 100 AMP at 480/277 and 125 AMP at 400/230 (three-phase, neutral and ground)  |
| Cooling            | Air-cooled, air flow: 3,000 cfm (1.41 m3/s); intake: bottom; exhaust: top  |
| 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  |
| Regulatory Compliance | UL 60950-1, CAN/CSA – C 22.2 No. 60950-1, CE-mark, RoHS, WEEE  |
# Cray XE6m-200 Specifications

## Processor
- 16-core 64-bit AMD Opteron 6300 Series processors, up to 192 per cabinet; NVIDIA Tesla X2090 GPU, up to 96 per cabinet, depending on configuration
- 8x64 KB L1 instruction cache, 16x16 KB L1 data cache, 8x2 MB L2 cache per processor core, 2x8 MB shared L3 cache

## Memory
- 32 GB or 64 GB registered ECC DDR3 SDRAM per compute node
- Memory Bandwidth: Up to 102.4 GB/s per node

## Compute Cabinet
- 1,536 processor cores per system cabinet
- Peak Performance: up to 15.4 teraflops per system cabinet

## Interconnect
- 1 Gemini routing and communications ASIC per two compute nodes
- 48 switch ports per Gemini chip, (160 GB/s internal switching capacity per chip)
- 2-D torus interconnect

## System Administration
- Cray System Management workstation
- Graphical and command line system administration
- Single-system view for system administration
- 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
- Full ECC protection of all packet traffic in the Gemini network
- Redundant power supplies; redundant voltage regulator modules
- Redundant paths to all system RAID
- Warm-swap compute blades

## 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 SLES11, HSS and SMW software]
- Extreme Scalability Mode (ESM) and Cluster Compatibility Mode (CCM)

## Compilers, Libraries & Tools
- PGI compilers, Cray Compiler Environment, PathScale
- Support for Fortran 77, 90, 95; C/C++, UPC, Co-Array Fortran
- MPI 2.0, Cray SHMEM, other standard MPI libraries using CCM
- Cray Apprentice, Cray PAT and Cray Compiler included with systems

## Job Management
- PBS Professional job management system
- Moab Adaptive Computing Suite job management system

## External I/O Interface
- InfiniBand, 10 Gigabit Ethernet, Fibre Channel (FC) and Ethernet

## Disk Storage
- Full line of FC-attached disk arrays with support for FC and SATA disk drives

## Parallel File System
- Lustre, Data Virtualization Service allows support for NFS, external Lustre and other file systems

## Power (maximum)
- 27.00 kW (27.27 kVA) per cabinet depending on configuration
- Circuit Requirements: 100 AMP at 200/208 VAC (three-phase and ground)
  - Direct wired standard
  - Optional Power Cords: NA/Japan HUBBELL 4100C9W, EMEA/APAC HUBBELL 563C6W
  - Or, 63 AMP at 400 VAC (three-phase, neutral and ground)

## Cooling
- Air-cooled, air flow: 2,300 cfm (1.09 m³/s); intake: bottom; exhaust: top

## Dimensions (Cabinet)
- H 80.5 in. (2,045 mm) x W 22.50 in. (572 mm) x D 56.75 in. (1,441 mm)

## Weight (Maximum)
- 1250 lbs (567 kg) air cooled

## Regulatory Compliance
- UL 60950-1, CAN/CSA – C 22.2 No. 60950-1, CE-mark, RoHS, WEEE

## Safety
- FCC, VCCI, ICES-003, EN55022, EN55024