Energy-Efficient, Air-Cooled Architecture

- Offers a wide range of flexible configurations based on industry-standard technologies
- Uses innovative technologies at a great price/performance
- Delivers an integrated HPC cluster software stack and reliable cluster management software
- Provides a modular and scalable turnkey solution with worldwide support and services

CS300-AC
Air-Cooled Architecture Delivering Outstanding Price/Performance

**System Overview**

The Cray CS300-AC™ cluster supercomputer is a highly scalable architecture that groups optimized, industry-standard building block server platforms into a unified, fully integrated system. Designed for medium- to large-scale capacity and data-intensive computing workloads, the Cray CS300-AC addresses both massively parallel and Big Data analysis challenges.

From a performance perspective, the Cray CS300-AC system offers over 11,000 compute nodes (nearly 300 racks) and scales to over 25 petaflops. It is integrated with Cray’s HPC cluster software stack including Cray’s Advanced Cluster Engine™ management software, providing a robust management suite delivering network, server, cluster and storage management.

The Cray CS300-AC system turbocharges industry-standard technologies with innovative and flexible HPC designs. A turnkey solution offering compute density, superior energy efficiency and price/performance with high performance network connectivity and the latest processing technologies, the Cray CS300-AC system delivers a wide range of configuration options. Additionally, the air-cooled design can provide savings in power consumption through the use of chilled cooling rack rear door heat exchangers.

**Ideal Environments**

The Cray CS300-AC system is designed for medium-to-large scale capacity and data-intensive computing workloads, addressing both massively parallel applications and Big Data analytics challenges.

**Capacity Computing**

Configurations focus on the latest processors including Intel® Xeon® and AMD Opteron™ with support for Intel® Xeon Phi™ coprocessors or NVIDIA® Tesla® GPU computing accelerators.
- Energy-efficient platform and superior performance per dollar
- Maximizes throughput for massive parallel computing applications

**Data-Intensive Computing**

Configurations focus on memory and bandwidth with excellent memory capacity per FLOP.
- Tight integration with Lustre-based global parallel storage systems and configurations using hierarchical storage – SSD, HD RAS features
- Processes, manages and analyzes large volumes of data – “Big Data”

**Air-Cooled, Energy-Efficient Architecture**

- Efficient design with a wide range of flexible configurations
- Optional 480V power distribution with a choice of 208V or 277V 3-phase power supplies
- Mixed-flow fans provide energy savings and noise reduction

**Flexible and Scalable Configurations**

- Two- and four-socket industry-leading x86 processors and accelerator-based platforms
- Multiple interconnect options including 3D Torus/fat tree, single/dual rail, QDR/FDR, IB/GigE
- Wide range of parallel file storage systems

**Manageable**

- Integrated HPC cluster software stack featuring Advanced Cluster Engine (ACE) management software
- Multi Linux OS support
- Manages heterogeneous nodes with different OS stacks
- Message passing libraries, compilers, debuggers and performance tools
- Network, server, cluster, and storage management
- Fine-grain system power and temperature monitoring
- Export and import of system configurations and images
- Detects hardware, fabric topology configuration errors
- Version control and ability to roll back changes
- Integrates job schedulers such as Grid Engine, SLURM and PBS Pro

**Reliable and Serviceable**

- No single point of failure with fault tolerance capabilities
- All critical components easily accessible and hot swappable
- Built-in multi-generation configuration software
- Worldwide system support and services including software updates, system maintenance, spare parts and repair with factory integration services
Flexible and Energy-Efficient Architecture
The Cray CS300-AC system can be configured in a fat tree or 3D Torus architecture making it optimized for superior application performance. The system incorporates two types of dedicated nodes: compute and service nodes. Compute nodes are designed to run parallel MPI and/or Open MP tasks with maximum efficiency. Service nodes are designed to provide scalability and I/O connectivity and can function as login nodes where applications are compiled and launched.

Highly Configurable Compute Nodes
Each Cray CS300-AC system is composed of two or four processors per node supporting 80 nodes per rack cabinet. Each node can be configured with 32GB, 64GB or 128GB DDR3 memory. Memory controllers ensure highly reliable memory performance while retaining platform upgradeability and flexibility. In addition, the platform supports a variety of hybrid computing configurations.

Scalable Interconnect Performance
The interconnect fabric for the system can be configured as a single- or dual-rail InfiniBand network. If single-rail, the compute nodes have a single QDR or FDR InfiniBand connection to the InfiniBand interconnect fabric. If dual-rail, the compute nodes have two InfiniBand connectors. The InfiniBand I/O channels on the compute nodes are based on low-latency host channel adapters. Each Scalable Unit (SU) modular rack system configuration can be replicated over and over to build reliable and very powerful large-scale systems using industry-standard components with a choice of fat tree or 3D Torus network topologies.

Intel® Xeon® Processor
The Intel Xeon processor features six or eight cores, delivering improved operational efficiency with up to 80% performance boost at the same power level as previous generations. It features cutting-edge technology including Intel® Advanced Vector Extensions, or Intel® AVX, providing up to two times the floating point operations per clock cycle coupled with the new Intel® Integrated I/O and Intel® Data Direct I/O that brings the I/O subsystem onto the processor die for the first time. The result is a significant reduction in I/O latency with the support of PCI Express 3.0 offering up to 80 lanes per two-socket server. In addition, the new processors offer Intel® Turbo Boost Technology 2.0 for maximized, efficient performance for single and multi-threaded applications.

Intel® Xeon Phi™ Coprocessor for Parallel Workloads
The Intel Xeon Phi coprocessor is based on Intel® Many Integrated Core (Intel® MIC) Architecture and it works synergistically with the Intel Xeon processor to increase developer productivity via common programming models and tools. The Intel Xeon Phi coprocessor delivers over 1 teraflops of double-precision peak performance in a single card. It offers many-core compared to multicore with wider vector processing units for greater floating point performance/watt. The Intel® Xeon Phi™ coprocessor is highly parallel and programmable based on open standards with support for data thread and process parallelism with full support from Intel® Cluster Studio XE while delivering outstanding aggregate performance and higher memory bandwidth.

AMD Opteron™ Processor
The AMD Opteron 6300 Series processor features up to sixteen cores per processor within the same package, 1MB of L2 cache per core and a shared 16MB of L3 cache per socket delivering improved performance per watt for multithreaded environments, all in the same power bands with the same power saving features as previous generations. It supports higher DDR3 frequencies and a maximum memory bandwidth of 51.2GB/s for memory-intensive workloads. It features AMD-P 2.0 technology to control power and cooling. Direct Connect Architecture 2.0 with HyperTransport™ Technology to help improve system efficiency and scalability, and AMD Virtualization Technology making it easier to virtually manage your supercomputer. Overall, the AMD Opteron 6300 Series processor delivers performance for the real world and value for real budgets by offering low acquisition costs that help reduce TCO.

NVIDIA® Tesla® K20 GPU Computing Accelerator
The NVIDIA® CUDA architecture enables developers to utilize simplified many-core NVIDIA® GPUs to solve the most complex, intensive computing tasks from the CPU to the GPU. It maximizes bandwidth while working with virtually any PCIe-complaint host system. Each compute server includes two NVIDIA® Tesla™ GPUs, each offering 512 cores and 665 gigaflops of double-precision peak performance, delivering supercomputing performance at one-tenth the cost and one-twentieth the power consumption. The NVIDIA Tesla 20-series GPU makes petascale computing with teraflop processors possible. In addition, it features 6GB of FDDR5 memory per GPU offering protection of data in memory, data integrity and application reliability with register files, L1/L2 caches, shared memory and DRAM; all are ECC protected.
Cray’s Advanced Cluster Engine (ACE) management software is part of Cray’s HPC cluster software stack. ACE is a complete management software suite designed to eliminate the complexity of managing an HPC cluster while providing all the tools necessary to run large, complex applications. ACE software includes command line (CLI) and graphical user interfaces (GUI) providing flexibility for the cluster administrator. An intuitive and easy-to-use ACE GUI connects directly to the ACE Daemon on the management server and can be executed on a remote system running Linux, Windows or Mac OS. The management modules include network, server, cluster and storage management.

ACE Software Supports:
• Multiple network topologies and configurations with or without local disks
• Network failover with high scalability and maximum reliability
• Customizable HPC development environment for industry-standard platforms and software configurations
• Heterogeneous nodes with different software stacks
• System power and temperature monitoring

Applications

Cray HPC Cluster Software Stack
OS, Debuggers, Compilers, MPI, Programming Tools

Advanced Cluster Engine™ - Remote Management Software
Turns Cray’s Clusters into a Functional, Usable, Reliable, and Available Computing System

Server
• Automatic Discovery
• Remote Power Control
• Scalable, Fast, Diskless Booting
• Redundancy and Failover

Cluster
• Partition into Multiple Logical Clusters, Each with Unique OS Configuration
• Integrated Job Scheduler
• Revision System with Rollback

Storage
• Scalable Root File System
• High Bandwidth to Secondary Storage

GUI and CLI
• View/Change/Control
• Monitor Health
• Plugin Interface

Network
• Automatic Discovery
• Redundant Paths
• Load Balancing
• Failover
Cray CS300-AC Cluster Supercomputer

| **Architecture** | Air cooled, up to 80 nodes per rack cabinet |
| **Processor** | Support for 6- or 8-core, 64-bit, Intel® Xeon® processor E5 family |
| | Support for 12- or 16-core AMD Opteron™ 6300 Series Processor |
| | Options for Intel® Xeon Phi™ coprocessors or NVIDIA® Tesla® GPU computing accelerators |
| **Memory** | 32GB, 64GB or 128GB registered ECC DDR3 SDRAM per compute node |
| | Up to 8GB 8-channel ECC GDDR5 memory per coprocessor node |
| **Interconnect** | 1, 10 or 40 GE Gigabit Ethernet |
| | QDR or FDR InfiniBand with Connect X3 or True Scale Host Channel Adapters |
| | Options for single or dual-rail fat tree or 3D Torus |
| **System Administration** | Advanced Cluster Engine (ACE): Complete remote management capability |
| | Graphical and command line system administration |
| | System software version rollback capability |
| | Redundant management servers with automatic load balancing and failover |
| | Automatic discovery and status reporting of interconnect, server and storage hardware |
| | Partition a cluster into multiple logical clusters, each capable of hosting a unique software stack |
| | Integrated job scheduling and management |
| | Remote server control (power on/off, cycle) and remote server initialization (reset, reboot, shut down) |
| | Scalable fast diskless booting for large node systems and root file systems for diskless nodes |
| | Multiple global storage configurations |
| **Reliable, Available, Serviceable (RAS)** | Redundant power, cooling, and management servers with failover capabilities |
| | Redundant root file system |
| | Built-in multi-generation configuration software management |
| | All critical components easily accessible and hot swappable |
| **File System** | NFS, Lustre® and Panasas® PanFS available as global file system |
| **Disk Storage** | Full line of FC-attached disk arrays with support for FC, SATA disk drives and SSDs |
| **Operating System** | RedHat, SUSE or Cent OS |
| **Compilers, Libraries & Tools** | Open MPI, MVAPICH2 or Intel® MPI Libraries |
| | PGI, PathScale, Intel® Cluster Toolkit compilers, CUDA, CUDA C/C++/Fortran, OpenCL, DirectCompute Toolkits, GNU, TotalView, OFED programming tools |
| **Power** | Up to 28KW per cabinet depending on configuration |
| | Optional 480V power distribution with a choice of 208V or 277V 3-phase power supplies |
| **Cooling Features** | Air cooled |
| | Airflow: 3,000 cfm; Intake: front; Exhaust: back |
| | Mixed-flow fans provide energy savings and noise reduction |
| | Optional passive or active chilled cooling rear door heat exchangers |
| **Cabinet Dimensions (HxWxD)** | 78.39”(1991mm) x 23.62”(600mm) x 47.24”(1200mm) standard 42U/19” rack cabinet |
| **Cabinet Weight** | 1,856.3 lbs.; 232 lbs./sq. ft. per cabinet |
| **Support and Services** | Worldwide system support and services including software updates, maintenance, spare parts and repair with factory integration services, system installation and operations training |