Cray® XC™ Series Interconnect and Network Technology

Aries™ Interconnect and Dragonfly Network Drive
Maximum Scalability and Sustained Performance

Cray® XC™ Series Supercomputers
Maximum scalability and sustained performance for today’s demanding production computing environments: Cray’s flagship XC series supercomputers deliver the industry’s most optimized HPC solution.

Cray XC series systems run the world’s most challenging and diverse workloads — and they can be configured across the price/performance spectrum to meet your own unique and exacting requirements. Whether you are a large government datacenter, cutting-edge research facility or competitive commercial organization, Cray can deliver the most optimized computing solution. With their state-of-the-art programming environment and upgradability by design, Cray XC systems deliver a competitive advantage in productivity and TCO.

Aries™ Interconnect and Dragonfly Topology
To provide this breakthrough advantage in performance and scalability, Cray XC series supercomputers integrate the HPC-optimized Aries interconnect. This innovative intercommunications technology, implemented with a high-bandwidth, low-diameter network topology called Dragonfly, provides substantial enhancements on all of the network performance metrics for HPC: bandwidth, latency, message rate and more. Delivering global bandwidth scalability at reasonable cost across a distributed memory system, this network gives programmers global access to all the memory of parallel applications and supports the most demanding global communication patterns.

Cray XC series liquid-cooled (LC) supercomputers utilize the Dragonfly network topology, flexibly constructed from a configurable three-rank mix of backplane, copper and optical links. This provides scalable global bandwidth, avoiding expensive external switches and enabling easy in-place upgrades for growing bandwidth requirements in the future. Cray XC air-cooled (AC) systems utilize the two-rank combination of backplane and copper cabling to reduce costs for technical enterprise applications and commercial environments.

Processing/Interconnect Integration Ensures Flexibility and Performance
Cray’s adaptive computing methodology enables the configuring of both multicore and many-core processors in the same hybrid system, ensuring users have the best compute node architectures available for any kind of workload or application. XC series compute nodes are configured as two sockets for Intel® Xeon® (Haswell/Broadwell) multicore processors and one socket for Intel® Xeon Phi™ (KNL) many-core processors. The XC series LC and AC supercomputers leverage the same best-in-class processing/accelerator technologies and upgradability on interchangeable compute blades, amortizing R&D expenses and reducing the cost of ownership.

The Aries ASIC provides the network interconnect for the compute nodes on XC series compute blades and implements a standard PCI Express 3.0 host interface, supporting a wide range of current and future HPC processing technologies and vendors. The universal nature of the Cray XC series open architecture means the system can be configured with today’s best available devices, then augmented or upgraded in the future with the user’s choice of processors/accelerators using flexible processor daughter cards on the base compute blades.

Learn more about the Cray XC series performance advantage in our white papers on scalability, networks and interconnects at www.cray.com/xc.
The Cray Interconnect/Network Advantage

Developing both high-end, performance-optimized supercomputers and commodity-focused cluster computing systems, Cray has depth and breadth of experience when it comes to understanding the benefits of different interconnects and networks for achieving different goals.

Rather than focus solely on individual internal specifications like signal rates, Cray optimizes everything possible about system hardware, software, OS and IP, taking into account the PCIe performance gate (injection bandwidth) to the processors in the compute nodes. With Cray's Aries/Dragonfly approach we can offer advantages that translate into real-world sustained application performance, and these advantages become even more apparent with the growth of a scaled system. Instead of comparing disparate specs on paper, we prefer to demonstrate our achieved results by benchmarking diverse and demanding industry codes with our customers and prospects.

The Aries interconnect is a Cray-developed ASIC designed in a holistic system approach, integrating four NICs with two router tiles for interconnect. The PCIe-3 compute node connections are flexible for supporting different processor/accelerator types and configurations.

This XC interconnect technology demonstrates a greater advantage on most applications where variable message sizes, connectionless protocol, link-level reliability, fast MPI hardware collectives, MPI vs. I/O traffic and fine-grain adaptive routing become more important to overall system performance.

Dragonfly topology is bandwidth rich, scalable and upgradable, addressing even the toughest network patterns, from nearest-neighbor type communication, irregular and all-to-all communications. Dragonfly groups of 384 compute nodes are constructed from 96 blades distributed over six chassis in a pair of cabinets. An electrical all-to-all network, constructed from backplane links, connects the blades in each chassis. A second all-to-all network, constructed using electrical cables, connects the six chassis in each group. The group-level interconnect is common to all large Cray XC systems. The full system is constructed by connecting Dragonfly groups with an optical all-to-all network.

The global bandwidth of an XC series supercomputer can be tuned by varying the number of optical cables in cabinet group connections. The Dragonfly network helps run workloads at very high system utilization efficiency, as job placement is insensitive, with jobs executing from perfectly grouped placement to worst-case placement with very little runtime variance. Cray optimizes the hardware, network, OS and software stack so users can focus on their algorithms and results.

Cray has developed local shared-memory optimizations in the XC series software stack libraries for MPI collectives that significantly improve performance on high-core-count compute nodes. Additionally, through implementing dynamic allocations (connectionless transfers) rather than static allocations, Cray has reduced the MPI memory footprint and overhead, improving scalable performance.

The Dragonfly network also offers the performance/cost advantages of fewer global traffic hops, no external director switches, and half the optical cables required for an equivalent fat-tree network implementation. Compared to more commodity-based interconnect/network technologies, the Aries interconnect delivers a sound cost/performance advantage, often delivering the option of configurations with more compute nodes for the difference in price of latest-generation cluster interconnects.

Ready for the Most Demanding Applications

The Aries interconnect and Dragonfly network integrated into Cray XC series supercomputers provide cost-effective, scalable global bandwidth. This architecture, together with the tuned Cray software stack and advanced reliability features, enables XC systems to scale multicore/many-core processors to millions of parallel compute cores. The Cray XC programming environment enables efficient use of these resources on the most demanding HPC applications, resulting in consistent, scalable and sustained computing performance.