CRAY®
THE SUPERCOMPUTER COMPANY

Cray® XC40™ Supercomputer

NVIDIA® Tesla® K40 GPU Accelerator Compute Blade

Adaptive Hybrid Computing and Scalable Many-Core Performance

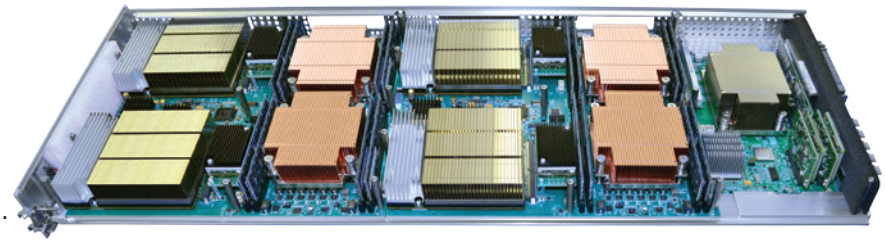
Supercomputer users procure their machines to satisfy specific and demanding requirements. But they need their systems to grow and evolve to maximize machine lifetime and return on investment. To solve these challenges today and into the future, the Cray® XC™ supercomputer series network and compute technology has been designed to easily accommodate upgrades and enhancements. Users can augment their systems in place to upgrade to higher performance processors, or add coprocessor and accelerator components to build even higher performance Cray XC40 supercomputer configurations.

The Cray XC supercomputer series architecture has been specifically designed from the ground up to be adaptive. A holistic approach optimizes the entire system to deliver sustained real-world performance and extreme scalability across the collective integration of all hardware, networking and software.

One key differentiator with this adaptive supercomputing platform is the flexible method of implementing best-in-class processing elements via processor daughter cards (PDCs) supported across the entire family of Cray XC40 and XC40-AC™ supercomputers.

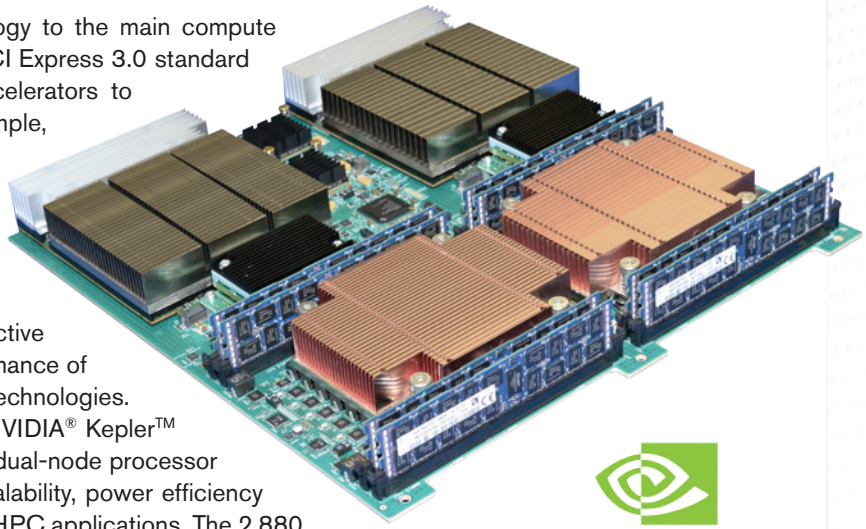
Cray XC Series Compute Blade

The Cray XC series architecture implements two processor engines per compute node and has four compute nodes per blade. Compute blades stack 16 to a chassis and each cabinet can be populated with up to three chassis, culminating in 384 sockets per cabinet. Cray XC40 supercomputers can be configured up to hundreds of cabinets and upgraded to exceed 100 petaflops per system.



Processor Daughter Cards (PDCs)

The Cray XC40 system mates processor engine technology to the main compute blades via two configurable daughter cards. The flexible PCI Express 3.0 standard accommodates scalar processors, coprocessors and accelerators to create hybrid systems that can evolve over time. For example, PDCs can be swapped out or reconfigured while keeping the original compute base blades in place, quickly leveraging the best possible performance technologies.



NVIDIA® Tesla® K40 GPU Accelerators

The Cray XC series expands on the founding legacy of productive hybrid supercomputing, leveraging the best-in-class performance of integrated multi-core scalar and many-core accelerator technologies. Utilizing NVIDIA® Tesla® K40 GPU accelerators with the NVIDIA® Kepler™ compute architecture and 12 GB of GDDR5 memory, this dual-node processor daughter card creates a hybrid unit with the intra-node scalability, power efficiency and flexibility to address the most performance-demanding HPC applications. The 2,880 embedded CUDA cores deliver over 1.43 teraflops of peak double precision floating point execution per GPU.

NVIDIA
TESLA®

Cray Inc. ▪ 901 Fifth Avenue, Suite 1000 ▪ Seattle, WA 98164 ▪ Tel: 206.701.2000 ▪ Fax: 206.701.2500 ▪ www.cray.com