

CRAY XC40

DataWarp™

Applications I/O Accelerator

The Cray® DataWarp™ accelerator delivers application-ready storage I/O to high-velocity data-driven workflows. Through a new balanced and cohesive I/O acceleration tier engineered for the Cray® XC40™ supercomputer, DataWarp software optimizes performance, maximizes system efficiency and reduces total cost of ownership.



The Applications I/O Challenge

As big data and supercomputing applications drive greater infrastructure demands, they also create an ever-widening gap between computational resources and disk-based storage performance. Compute, memory and flash storage technologies have advanced performance on track with Moore's Law. But disk-based storage technologies, while improving in density and capacity, have not kept pace. If you rely on conventional disk and file system technologies, delivering sustainable performance on a spectrum of applications with varying I/O-intensive workloads becomes costly and impractical.

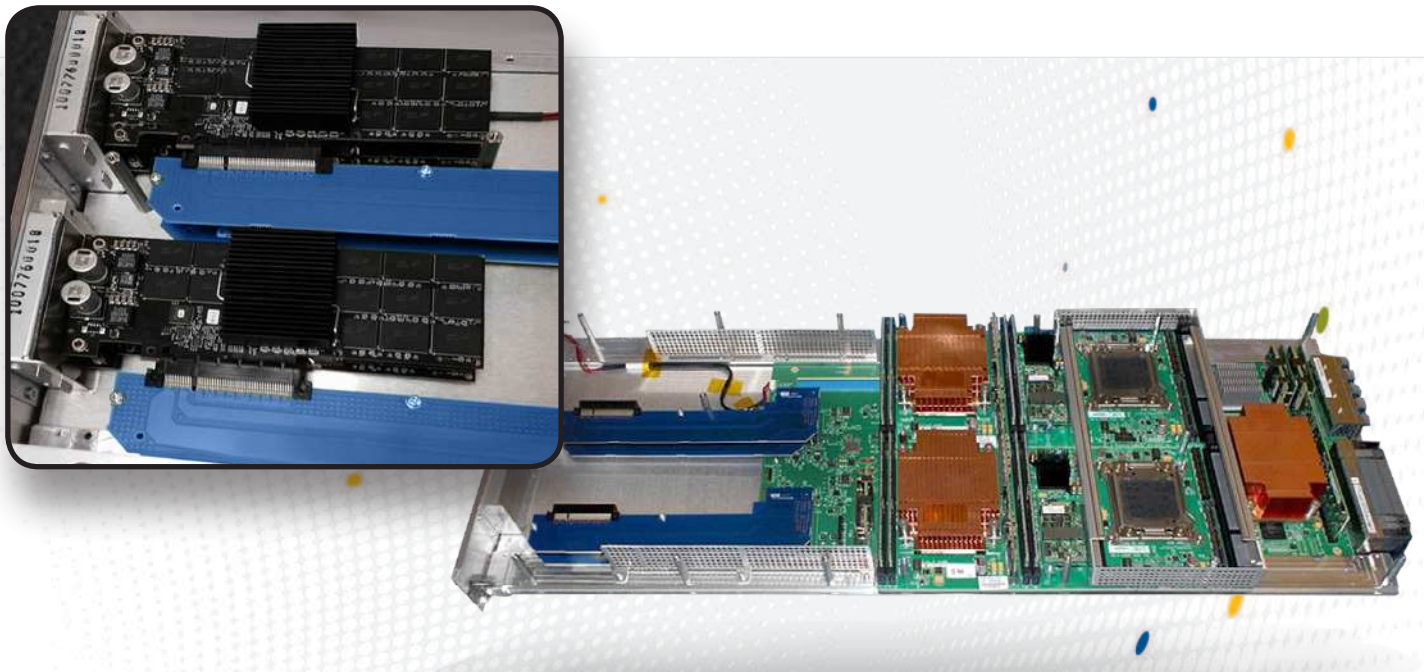
The Cray DataWarp storage system architecture addresses this growing gap by:

- Bringing the fastest applications I/O as close as possible to the compute resources
- Providing a flexible, balanced infrastructure that can map to different applications use cases
- Ensuring a more cost-effective balance of I/O across all storage tiers and mechanisms to move data efficiently and simply across tiers
- Creating a central pool of ultrafast I/O resources that can be allocated in diverse ways to different applications needs
- Meeting the data I/O surge needs of applications and use cases

Performance, Efficiency and Balance for I/O-Intensive Supercomputing

Big data, HPC and supercomputing storage infrastructure managers need a balanced compute-memory-storage system architecture with an innovative cache tier featuring SSD and in-memory flash. The tier should accommodate "bursty" application I/O patterns through a combination of high-bandwidth and low-latency I/O forwarding to deliver thousands to millions of aggregate IOPS to applications. The Cray® XC™ series DataWarp applications I/O accelerator technology addresses these modern I/O needs by delivering a balanced and cohesive system architecture from compute to storage. The DataWarp accelerator allocates storage dynamically in either private (dedicated) or shared modes. Storage performance quality of service can be provided to individual applications, based on the user's policies. The benefits are simple but critical: Workflows are streamlined because applications "just work" — they process faster and access storage on demand. And highly efficient workflows enable productive users.

While leveraging Cray's proven domain expertise in storage, the DataWarp accelerator can be used as a global storage cache for parallel file systems (PFS) such as Lustre®, General Parallel File System (GPFS™) and PanFS™. In these scenarios, the applications I/O accelerator capabilities drive up the overall utilization of the parallel file system by buffering performance across this new tier. This enables new levels of pure performance boosts by delivering the right storage performance to applications. The accelerator improves overall application performance by decoupling application I/O from the corresponding PFS I/O while also improving the performance and resiliency of conventional disk-based solutions.

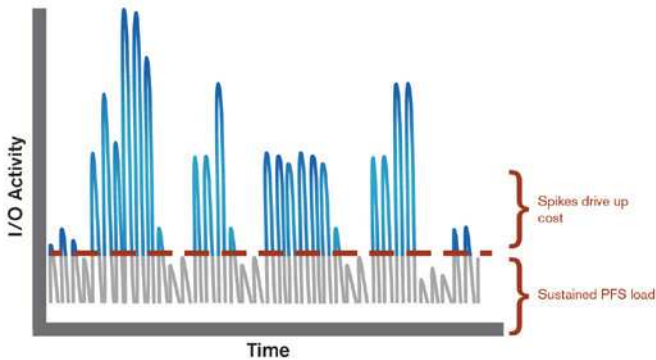


DataWarp I/O blades with SSDs are inserted into XC40 banks of compute blades and all are connected via the high speed Cray Aries HPC interconnect. Every XC series blade type leverages the Aries interconnect.

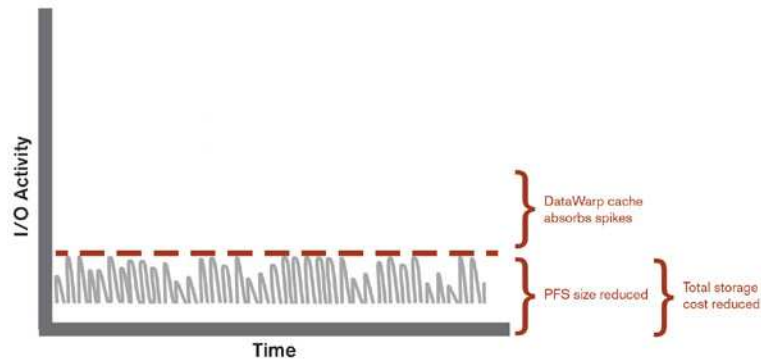


Taking advantage of the quicker price evolution of nonvolatile flash, Cray has designed I/O blades with fast SSD local to the powerful XC series compute blade nodes. This introduces a buffer platform tier that complements the balance of existing on-board memory and off-board disk storage. This adaptive architecture addresses the growing throughput/capacity gap to enable a caching platform that can support burst buffer capabilities (providing temporary nonvolatile data staging) and ensure peak performance of a PFS. Enabled by PCIe interfaces, the new XC series I/O blades can be configured with a variety of flash SSD options, giving each user the optimal balance of speed, performance and cost.

Before I/O Accelerator



After DataWarp I/O Accelerator



The DataWarp applications I/O accelerator advantage absorbs surge peaks, driving down total storage requirements and increasing total storage utilization.

Five Times the Performance of Disk-Only Systems — at the Same Cost

The DataWarp accelerator provides adaptive application-ready I/O at five times the performance of a traditional disk-based PFS at the same cost. The cost per IOPS is reduced up to four times — scaling from 70,000 to 40 million aggregate IOPS in a single supercomputing system. High bandwidth can be delivered with virtually no impact on other I/O executing in the system, ensuring quality of service and sustained bandwidth to specific applications, with a performance improvement of up to 10 times over PFS-based scratch systems. Rebalancing compute-storage systems with optimized levels of memory and flash SSD reduces some of the disk arm seek flailing that can cause invisible performance losses.

The accelerator leverages all the proven performance advantages of the XC series Dragonfly network topology as the system's high-speed network and places the highest-performance storage devices as close to Cray's Aries interconnect as possible. The accelerator also implements sophisticated forwarding capabilities based on Cray's exclusive and proven Data Virtualization Service software. Improving compute node access to data with a local storage tier provides cost-effective bandwidth and exceptional sustained aggregate bandwidth per dollar.

Reduced Cost and Breakthrough Efficiencies in Adaptive Flexibility and Ease of Use

While increasing compute and storage utilization for both peak and sustained workloads, the DataWarp applications I/O accelerator reduces:

- Cost per IOPS. New SSD-based tiers deliver two to five times the IOPS of disk-based systems.
- Cost of bandwidth. The new SSD-based tiers deliver 10 times the bandwidth of disk-based systems.
- Total cost of ownership (TCO) of the supercomputer.

The accelerator provides a new level of flexibility, enabling you to allocate the appropriate type and amount of data storage and I/O movement per job, process, rank or node. Storage is dynamically allocated to maximize compute and storage utilization across the entire system — adaptive flexibility means you can put the best resources where you need them the most. For example, you can put a scratch file system on every node, or a burst capability local to compute nodes for faster checkpoint restart.

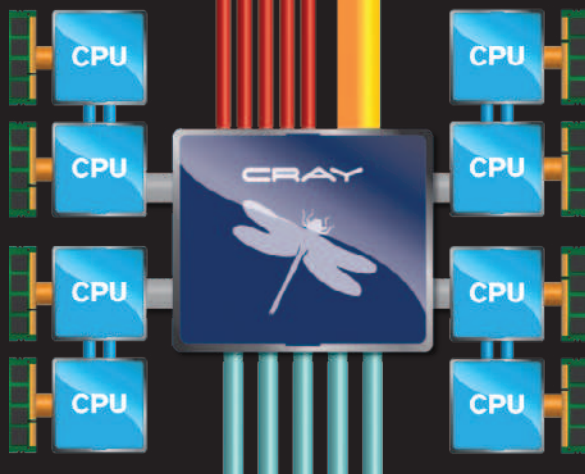
Leveraging a worldwide installed base of leadership-caliber systems and Cray's 40-plus years of domain knowledge in HPC-optimized, high-density design and integrated hardware/software packaging, the I/O accelerator delivers higher performance with fewer cabinets, cables and disks, lowering your TCO. The total machine efficiency for a large capability-class machine, for example, could be increased by 70 to 90 percent, saving millions in compute costs. Moreover, storage is automatically and dynamically provisioned, driving up storage utilization and driving down the cost of storage operations.

The DataWarp accelerator's cohesive software stack enables a broad set of HPC tools — from workload managers to job schedulers to user scripts — to streamline application processing in a range of environments, from single apps to parallel. The system provides low-level (kernel mode) POSIX-compliant I/O forwarding techniques. By cutting down on overhead and improving workload throughput by 10 to 15 percent, users reclaim compute time and improve overall performance. Wall clock execution time is optimized, meaning more runs per day and iterations per compute session, faster time to results and higher productivity. In fact, I/O on demand and features like swap space support help prevent system out-of-memory failures that can kill job runs and waste valuable compute time.

Cray DataWarp Technology Delivers:

- Flash SSD I/O blades and Aries high speed interconnect
- Data virtualization software, management software and I/O forwarding mechanisms
- Ability to scale up a data pool to customer requirements
- Performance gains: I/O and total system compute, bandwidth and IOPS
- I/O system flexibility, data shared across multiple levels
- Integration and support: hardware, software, networking, partners, open source
- Reduced storage system costs and improved TCO
- System balance and efficiency across memory-flash-disk
- Workflow efficiencies and ease of use advances

Customers who purchase the DataWarp accelerator options with their Cray® XC40™ systems will gain access to future releases of the technology and the complete DataWarp accelerator roadmap, which will allow for an expanding repertoire of use cases.



The DataWarp storage tier leverages the HPC performance advantages of the Aries interconnect

The Aries ASIC provides the network interconnect for the compute and I/O nodes on the Cray XC40 system base blades and implements a standard PCI Express Gen3 host interface, empowering connectivity to a wide range of HPC processing compute engines. This universal nature of the Cray XC series open architecture allows the system to be configured with the best available devices today, and then augmented or upgraded in the future with the user's choice of processors/coprocessors utilizing processor daughter cards (PDCs), each with their own independent capabilities and development schedule.

