



## Cray Data Management Platform for Big Data and HPC

Cray's Data Management Platform (DMP) provides a suite of storage connectivity and data management tools for Big Data and high performance computing (HPC) storage scenarios supporting Cray supercomputers and popular x86 Linux® compute environments. Cray's expertise in Lustre™ and storage management infrastructure complements an extensible, open, data management platform for sharing, moving, and archiving data for maximum performance at scale.

# Cray Data Management Platform for Big Data and HPC

Cray's data management platform provides the building blocks for developing complete and scalable open storage solutions, such as Cray Cluster Connect (C3), where storage and data management across compute and storage platforms is required. DMP's components are purpose built for Lustre integration and data sharing scenarios.

The DMP suite is comprised of the following core products and services:

**Lustre File System by Cray:** Cray provides a standard, fully-qualified and thoroughly tested version of Lustre, deployed and supported by Cray on select storage platforms.

**Cray Data Movers:** Data movement systems accommodating a broad range of data management scenarios from data ingest to archive to distribution. Cray Data Movers provide a range of connectivity options to 3rd-party storage and hierarchical storage management (HSM) platforms.

**Cray Development and Login:** Powerful, multi-user development and login environment for Cray XE™/XK™ and XC30™ systems. The external version of the programming environment offers the same development capabilities as delivered natively on Cray supercomputers, but on external x86 servers.

**Cray Routing Services:** Offers scalable, flexible, and non-disruptive InfiniBand routing configurations and architectures for Lustre and associated networking infrastructure.

**Cray Management Services:** A centralized management platform to aggregate data and manage external services, failover, power management, system provisioning, reporting, high availability, and node health monitoring.



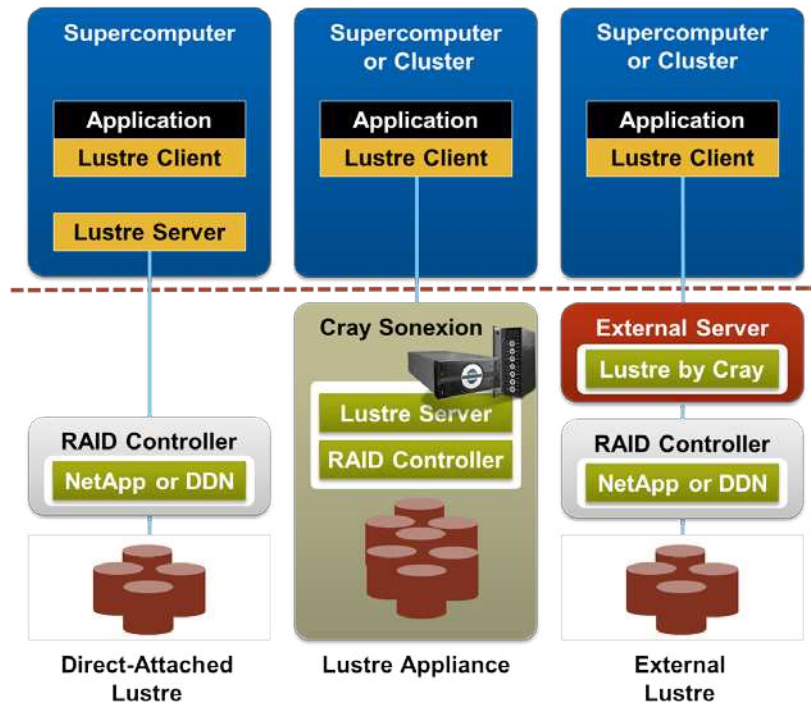
## Cray Lustre Solutions

Cray offers complete Lustre storage solutions for supercomputers, cluster supercomputers, and x86 Linux compute clusters of any type.

### Lustre File System by Cray

Cray provides a standard, fully tested version of Lustre called the Lustre File System by Cray (LFSC). This model provides the building blocks for a complete, compute-agnostic Lustre solution which can be shared within and across multiple compute environments. LFSC offers configuration flexibility and choice of storage platforms. Lustre is offered on a well-defined set of validated servers and storage that Cray sells and supports. Validated storage configurations are offered for DDN SFA and NetApp E-Series block storage arrays. Cray also offers a version of Lustre installed on Cray Sonexion, an integrated Lustre system consolidating all Lustre modules in a compact form factor.

Cray's version of Lustre is based on the standard OpenSFS (Open Scalable File Systems) distribution. Once a version of Lustre meets quality levels determined by Cray, Cray ensures that the version undergoes extensive stress and scalability testing in a variety of conditions. All changes and modifications are contributed back into the OpenSFS source code distribution.



### Compute Considerations

Two compute-dependent models are used for deploying Cray Lustre solutions: internal and external. Internal Lustre solutions embed Lustre capabilities within a Cray supercomputer such as the Cray XC30 system. For example, Lustre servers and LNET routers are installed and executed within the machine.

External Lustre operates the object storage server (OSS) and metadata server (MDS) software on external servers. The Lustre clients are deployed on compute nodes and support all compute platforms, from supercomputers to commodity x86 Linux clusters. External Lustre consolidates all server, LNET and storage modules on network-connected high performance block storage.

### Lustre on Cray Supercomputers

Lustre and its associated components can be optionally installed on dedicated nodes within a Cray supercomputer. Specified nodes perform specific Lustre roles such as OSS or MDS to effectively offload I/O from compute nodes within the supercomputer. This architecture effectively brings data closer to compute and can benefit from a number of Cray Lustre routing architectures and configuration tools.

### Storage System Considerations

A variety of storage systems, storage configurations and data management scenarios are supported by Cray. Cray offers a fully tested and validated version of the Lustre file system which evolves over time, as Cray deems Lustre production ready or feature capable. The two primary storage system architectures for deploying Lustre are component-based and integrated.

### Component Solutions built on Lustre File System by Cray

Component-based solutions offer choice and flexibility where Lustre is installed by Cray on the various brands of components, servers and storage. Cray designs primary parallel storage architectures for NetApp or DDN block storage. These solutions provide interoperability to compute, networking and other storage tiers. In this scenario, Cray deploys its tested and fully supported Lustre file system on servers and storage components, and applies Cray's best practices to supported configurations. Currently Cray supports a range of configurations on a specific set of qualified servers, networking components and storage.

### Pre-Integrated Systems

For customers requiring fully integrated Lustre where all components come pre-configured and assembled in a compact form factor, Cray offers Sonexion®. Cray Sonexion consolidates and integrates all Lustre components and modules — servers, networking and storage. Cray Sonexion reduces the number of managed physical and logical components — hardware and software. Reducing the number of components simplifies deployment and the number of management points (fewer components to manage). Another benefit of Cray Sonexion is scalability. Performance and capacity are scaled in modular, balanced increments. Large I/O bandwidth scales from 5GB/s to 1TB/s aggregate throughput in a single file system.

## Lustre Solution Matrix

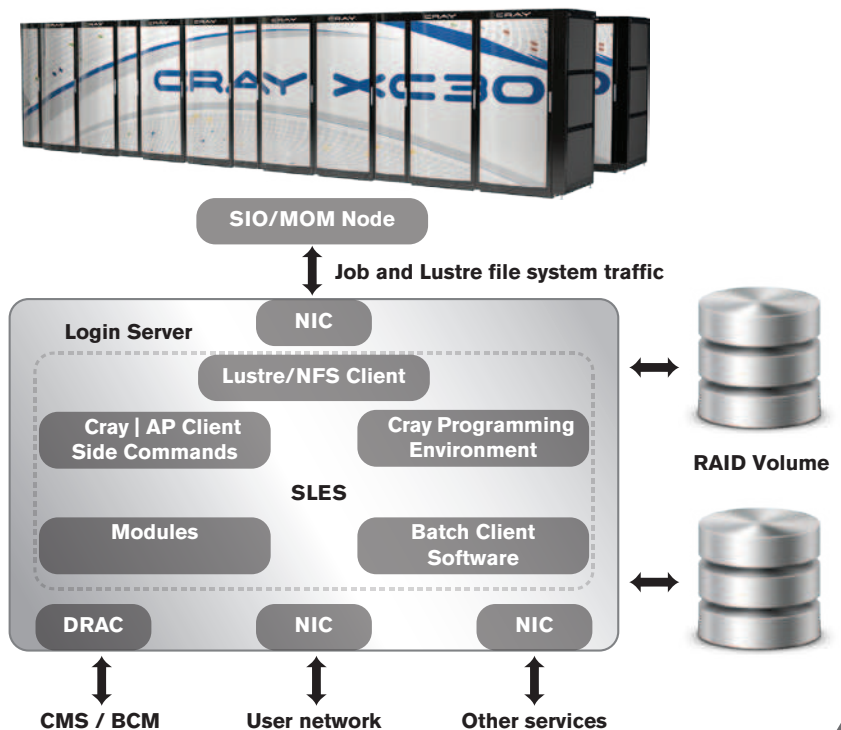
Compute Platform	Lustre Model	External Server Options	Lustre Storage Options
Commodity x86 Linux Cluster	External through Cray Cluster Connect	Cray-supplied external rack-mount servers	Cray Sonexion DDN SFA arrays NetApp E-Series
Cray CS300 x86 Linux Cluster Supercomputer	External through Cray Cluster Connect	Cray-supplied external rack-mount servers	Cray Sonexion DDN SFA arrays NetApp E-Series
Cray XC30 Supercomputer	Direct-attached or externally through Cray Cluster Connect	Cray-supplied external rack-mount servers	Cray Sonexion DDN SFA arrays NetApp E-Series

## Cray Development and Login

Cray development and login nodes are deployed externally from a Cray supercomputer providing a powerful, multi-user development and login environment. The Cray XC30 supercomputer and other Cray systems provide extreme scalability by optimizing each component of the system (compute, network, software, infrastructure and I/O). Optimized compute nodes deliver outstanding computational performance, while optimized service nodes provide system and I/O connectivity. Some of the service nodes also serve as development and "login nodes" from which users compile and launch jobs.

### Capabilities include:

- Distributed login capability physically decoupled from the systems to allow users login access to Cray systems and/or partitions.
- Ability to compile, link and debug applications, and submit and manage jobs on the Cray system – and use a familiar environment as provided on Cray supercomputers; external login nodes offload processing resources to ensure continued productivity on the Cray supercomputer with access to compilers, libraries, debuggers, data, tools, or when the Cray system is unavailable.



A range of pre-configured login nodes are sized to fit the customer's login services load. Networking options include Ethernet and InfiniBand. The nodes offer a combination of hardware with multiple processors, large memories and large capacity local disk for local swap and temporary space as well as high bandwidth network connectivity to facilitate an extremely efficient environment.

At the heart of the Cray development and login software stack is the Novell SUSE Linux Enterprise Server (SLES) operating system providing commonality and compatibility with the Cray Linux Environment™ (CLE) on Cray XC30 systems. In addition, users have access to the Lustre file system on the system as well as access to other file systems in their environment.

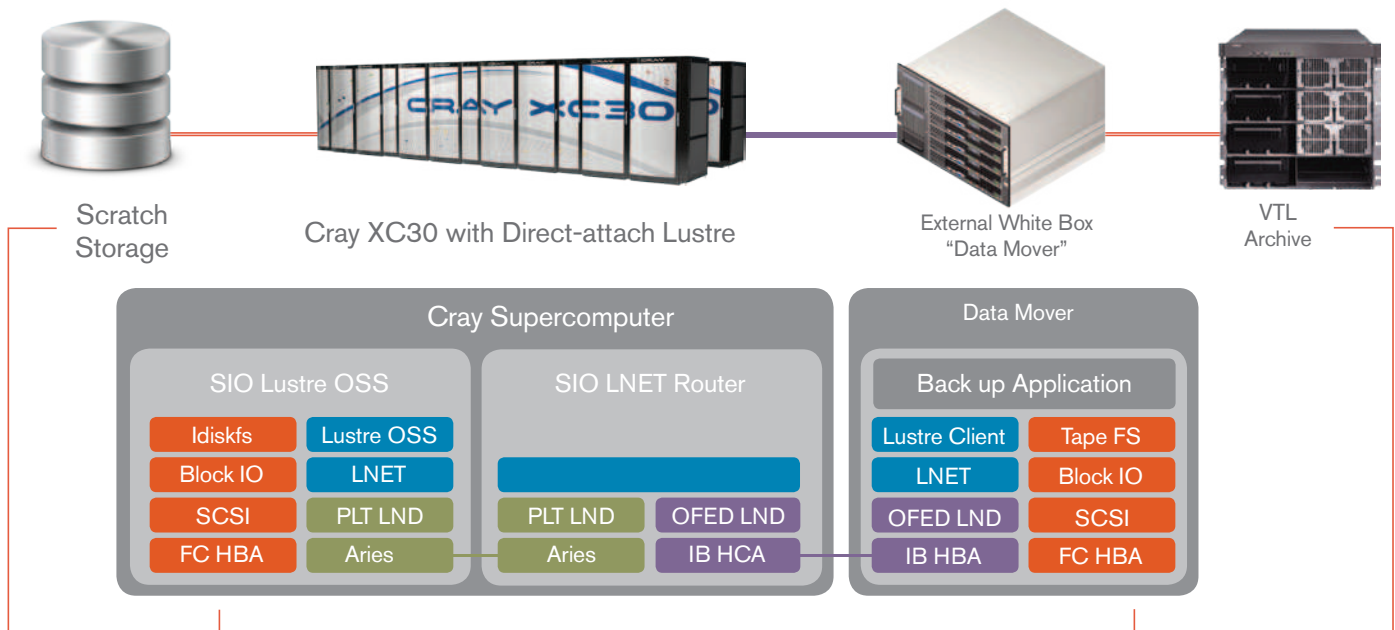
The Cray development and login software stack provides a similar user interface and programming environment as provided by Cray XC30 login nodes. Users familiar with compiling codes, running debuggers and submitting jobs from internal Cray login nodes will be immediately productive using external Cray development and login servers which in some cases offer more powerful computational, storage and I/O capability than internal Cray login nodes.

## Cray Data Movers

Cray Data Movers (CDMs) provide a broad range of connectivity and data movement options to third-party storage systems and Cray storage systems. Connectivity services include data collection for ingest, data movement for processing, backup/restore, interfaces to HSM and data distribution or file transfer to remote sites using a variety of technologies.

To support data movement within a Cray supercomputer environment, such as to share data and copy data to other media including disk, tape and virtual tape libraries, Cray Data Movers can be directly connected to one a Cray I/O node with or without an InfiniBand switch.

Performance for an individual Data Mover depends on several factors — the read bandwidth of the Lustre file system, the write bandwidth of the backup/archive target, the underlying CDM server hardware, and the backup application or Linux utility moving data. Depending on the hardware and software configuration, a single Data Mover can support transfer rates of up to 3GB/s.



## Cray Routing Services

Cray's expertise in Lustre includes routing architectures for InfiniBand and Ethernet. Cray offers scalable and non-disruptive system architectures for InfiniBand and bridges to Ethernet. Benefits include maintaining flexibility, non-disruptively extending InfiniBand fabrics, and routing techniques for Lustre networking (LNET).

### Key capabilities include:

- Bridging of InfiniBand traffic to 10/40GbE.
- Isolation for non-disruptive fabric operations and scalability.
- Reliability and resiliency of routing for continuous operations.
- Adaptive routing: If a router goes down, clients can be configured to use alternative routes.
- Highly available configurations for Lustre.
- Management of router nodes.

LNET routers can be controlled by multiple Cray Management Services (CMS) nodes configured in a highly available configuration using Bright Cluster Manager (BCM) for control and monitoring. Depending on the configuration, routing nodes can deliver over 5GB/s of data to the clients using FDR InfiniBand.

## Cray Management Services

Cray Management Services (CMS) customized around Bright Computing's Bright Cluster Manager provides a centralized management platform to aggregate data and manage external services, power management, system provisioning, reporting, system availability, and node health monitoring.

### CMS scales to manage hundreds of servers responsible for a number of management operations. Basic operations include:

- Power management and node health monitoring.
- Imaging, provisioning and change management, including update installation and management.
- High availability for continuous operations.
- Customizable configuration operations to support customer-specific management requirements.

