



CRAY[®] SONEXION[®]

The Cray Sonexion[®] scale-out Lustre storage system rapidly delivers precision performance at scale to your HPC cluster of choice, with the fewest number of components. Brought to you by Cray, the world's leading experts in large-scale parallel storage solutions for HPC, Sonexion's modular, compact design keeps costs low while delivering the right performance for compute clusters and applications of all types. Performance scales in modular building blocks, reducing the number of hard drives needed to achieve sustained performance at scale. Sonexion scales in predictable, performance-optimized configurations, ensuring precise levels of performance as capacity expands — and ensuring reliability, availability and stability at scale. Finally, Sonexion's compact form factor reduces the total storage hardware infrastructure (cables, servers, components and racks) required for sustaining production-grade, petascale deployments — by fifty percent on average over custom, component-based configurations.



CRAY[®]

SONEXION[®]

Precision Performance — Modular Scalability

The Cray Sonexion scale-out system maximizes the performance and capacity scaling capabilities of the Lustre file system. This integrated and modular storage solution is composed of high performance Scalable Storage Units (SSU), a Metadata Management Unit (MMU) and a network-ready rack that includes all storage and processing needs for a complete, production-grade parallel storage system. This architecture results in a high performance Lustre solution that is easier to deploy, scale and manage over time than custom-designed, component-based Lustre solutions.

Utilizing Cray's best practices for deploying and scaling Lustre, each storage unit consolidates and integrates networking, storage and Lustre into an embedded, rigorously tested, appliance-like module called the SSU. The SSU combines high-density stor-

age building blocks with 6GB/s-capable controllers. Features include dual hot spare drives and mirrored solid state disk that help accelerate file system journaling. With the range of options available, SSUs can be configured to deliver an unmatched level of performance that scales almost linearly with minimal storage or network performance degradation.

Each SSU is physically capable of delivering 6GB/s of peak bandwidth. Using the IOR benchmark, Cray's performance team has benchmarked real-world, sustained file system performance of each SSU to be greater than 5GB/s per SSU and up to 1TB/s sustained performance to a single file system. Each standard rack can store up to 1.4 petabytes of file system capacity.

Compact Design for Density at Scale

Every Cray Sonexion rack comes pre-assembled, integrated, configured and tested. The modular architecture, combined with the Cray Sonexion System Manager (CSSM), simplify both the initial deployment and ongoing operations — ensuring non-disruptive, performance-optimized scalability. Cray Sonexion storage systems run efficiently utilizing both air- and water-cooled options. This power efficiency and extreme density for Lustre reduces the cost of operating petascale systems.

As of 2012, Cray has shipped more than 50 petabytes storage capacity and delivered Cray Sonexion storage systems to customers in the energy, government and academic market segments. As part of the Blue Waters supercomputer at the University of Illinois' National Center for Supercomputing Applications, Cray installed a 36-cabinet, 25 petabyte Sonexion storage system capable of delivering more than one 1 terabyte of I/O bandwidth performance to a single file system and supporting more than 25,000 compute nodes.

At Blue Waters, the datacenter footprint — which is measured in total number of racks required to store 25 petabytes and achieve 1TB/s performance to a single file system — was delivered in one-third the total amount of hardware compared to another leading storage systems vendor.

Easier to Deploy and Manage at Scale

Because Sonexion comes pre-integrated, pre-tested and pre-configured, deployment time is greatly reduced. All components — networking, storage, RAID, operating system and Lustre — come pre-configured and pre-cabled. There are no external servers, switches or ad hoc systems to manage. As your Lustre file system grows, Sonexion enables easy scale-out of performance and capacity. Simply add storage building blocks — SSUs, expansion modules and racks — to achieve prescribed performance and capacity objectives. Cray's expertise integrating and configuring Lustre is embedded in the design — enabling customers to focus on research instead of Lustre.

The Cray Sonexion System Manager (CSSM) software application simplifies the end-to-end experience of deploying, managing

and operating a large-scale Lustre solution.

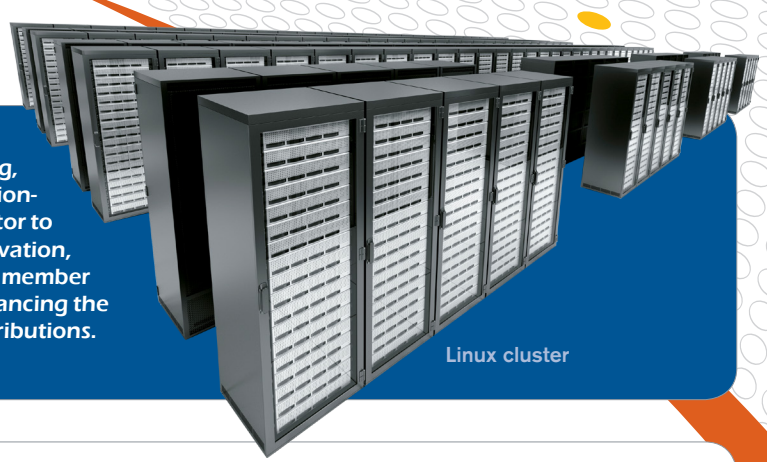
The Cray Sonexion System Manager offers system administrators an intuitive interface and alternative command-line set of tools to monitor and optimize the entire storage system.

CSSM provides status and control of all system components, including storage hardware, RAID, operating system and the Lustre file system in an integrated administrator interface. A web client hosted on one of the controller modules in the MMU interfaces with all distributed system manager component services. CSSM also integrates a comprehensive set of community-developed tools that collect, index and analyze fast-moving data to help administrators keep the system stable and balanced.



Large-scale Lustre Solution Experts

The Cray Sonexion family leverages Cray's expertise in designing, deploying and managing the world's largest and fastest production-grade parallel storage solutions. Cray has been a leading contributor to the Lustre open source community, collaborating on Lustre innovation, testing and configuration for HPC. Additionally, Cray is a founding member of the OpenSFS consortium which is dedicated to continually enhancing the Lustre roadmap and ensuring consistent, high-quality Lustre distributions.



Linux cluster

Reliability, Availability, Serviceability and Stability — through Quality

Cray's expertise in designing, deploying, optimizing and supporting large-scale parallel storage systems has enabled Cray to deliver a highly reliable, available and stable clustered storage system. The modular and redundant architecture of Cray Sonexion systems provides the highest reliability and resiliency to Lustre storage solutions for HPC and technical computing. In addition, each component, module and subsystem undergoes exhaustive factory testing under the most demanding test conditions to ensure maximum system robustness. Cray's quality assurance team simulates real-world, thermal and failure stress testing matching worst-case, production operations scenarios.

The MMU module hosts the metadata and management server operations. SSU modules operate as active/active-integrated server modules with redundant and independent system interconnections, providing maximum reliability while delivering maximum performance. The CSSM is tightly integrated into the system stack — from storage and embedded server modules to the Lustre file system and the entire storage cluster — enabling rapid, accurate monitoring and diagnosis down to the component level. System-wide software and firmware upgrades are executed through a simple and single interface in the CSSM system, removing the complexity and risks of traditional large Lustre implementations.

Configuration

The Cray Sonexion comes pre-configured and optimized for scaling Lustre without redesign. Cray's Lustre design ensures optimum performance configurations across the spectrum of Lustre: from initial deployments to multi-petabyte file systems. Storage operators simply add SSUs (capacity and performance) or expansion modules to meet the performance and capacity objectives of the storage system. Each file system will vary based on number of SSUs and expansion modules to meet the individual bandwidth and usable capacity requirements for each storage system. Each 42U rack comes with two InfiniBand switches for linking the MMU, SSUs and Lustre clients. The rack also contains all power supplies, InfiniBand and Ethernet cabling, and a dual Gigabit Ethernet switch for management system connections to individual components.

A base MMU is configured in a 2U 24-bay 2.5" drive enclosure with 22 10K RPM disk drives and two solid state drives. A 5U MMU is also available with up to 80 15,000 RPM SAS drives, enabling higher metadata performance.

A base SSU is housed in a 5U 84-bay 3.5" drive enclosure with 80 drives used to provide data storage in an 8 x (8+2) RAID6 target configuration, two global hot spares and two solid state drives. An SSU expansion enclosure, which has the same drive configuration as the base SSU, can be added via SAS connections to double the usable capacity for a given bandwidth.

Open Systems Open Opportunities

Parallel storage solutions from Cray support industry-standard Linux compute clusters from Dell, HP and IBM through a standard Lustre client. Cray Sonexion is available for a wide range of x86 Linux InfiniBand-attached compute clusters.

Customers also need to ensure investments made today support

solutions of the future. A key tenet of open systems is taking advantage of collaborative innovation through open source solutions, where possible. Cray is leading the open source community through OpenSFS, ITBA (InfiniBand Trade Association), OFA (Open Fabric Alliance) and other partner-driven innovation groups.

Trustworthy and Dependable Service and Support

Cray provides a single point of service and support for multivendor solutions — and takes a system-wide solution approach to service, integration and support.

Cray offers a flexible set of customer support options and will tailor these to meet the needs of a range of customers across industries. Cray can provide expert support and services for performance tuning, tailoring configurations to application needs, as well as offering pre- and post-sales support for facilities and datacenter integration.

Cray Sonexion Specifications

Rack	Sonexion 1300	Sonexion 1600
Height	42U or 1,867 mm (73.5 in)	42U or 1,867 mm (73.5 in)
Width	600 mm (23.6 in)	600 mm (23.6 in)
Depth	1,200 mm (47.2 in)	1,200 mm (47.2 in)
InfiniBand Switches	Dual 18 port QDR switches standard Dual 36 port QDR switches optional	Dual 36 port FDR switches standard
Rack Switches	Dual 24 port Gigabit Ethernet switches standard Dual 48 port Gigabit Ethernet switches optional	Dual 24 port Gigabit Ethernet switches standard Dual 48 port Gigabit Ethernet switches optional
Standard Cooling	Passive	Passive
Water-Cooled Option	Rear Door Heat Exchange Unit	Rear Door Heat Exchange Unit
Full Rack Weight (Standard Air-cooled)	1,141 kg (2,510 lbs)	1,141 kg (2,510 lbs)
Full Rack Weight (Water-cooled Door)	34.5 kg (76 lbs) additional	34.5 kg (76 lbs) additional

Metadata Management Unit (MMU)

Metadata Controller Height	2U or 89 mm (3.5")	2U or 89 mm (3.5")
2U24 Metadata Disk Enclosure Height	2U or 89 mm (3.5"), 24 drive slots	2U or 89 mm (3.5"), 24 drive slots
2U24 Metadata Disk Enclosure Drives	22 x 2.5" 450GB or 600GB 10K drives	22 x 2.5" 450GB or 600GB 10K drives
2U24 Metadata Disk Enclosure Storage	14 drives, RAID10 7+7	14 drives, RAID10 7+7
2U24 Metadata Disk Enclosure MGT	2 drives, RAID1 1+1	2 drives, RAID1 1+1
2U24 Metadata Disk Enclosure CSSM	4 drives, RAID10 2+2	4 drives, RAID10 2+2
2U24 Metadata Journal SSDs	2 100GB SSDs, RAID1 1+1	2 100GB SSDs, RAID1 1+1
5U84 Metadata Disk Enclosure Height	5U or 222 mm (8.75"), 84 drive slots	5U or 222 mm (8.75"), 84 drive slots
5U24 Metadata Disk Enclosure Drives	82 x 2.5 inch 300GB 15K SAS drives	82 x 2.5 inch 300GB 15K SAS drives
5U24 Metadata Disk Enclosure Storage	72 drives, RAID10 36+36	72 drives, RAID10 36+36
5U24 Metadata Disk Enclosure MGT	2 drives, RAID1 1+1	2 drives, RAID1 1+1
5U24 Metadata Disk Enclosure CSSM	6 drives, RAID10 3+3	6 drives, RAID10 3+3
5U24 Metadata Journal SSDs	2 100GB SSDs, RAID1 1+1	2 100GB SSDs, RAID1 1+1
Maximum Number of Files per File	6 billion	6 billion

Scalable Storage Unit (SSU)

Base/Expansion Unit Height	5U or 222 mm (8.75")	5U or 222 mm (8.75")
Base/Expansion Unit Drive Slots	84 x 3.5" drives	84 x 3.5" drives
Base/Expansion Unit Data Drives	82 x 1TB, 2TB or 3TB, 7.2K RPM NL-SAS, RAID6	82 x 1TB, 2TB, 3TB, or 4TB 7.2K RPM NL-SAS, RAID6
Base/Expansion Unit SSDs	2 x 100GB, RAID1, N/A for Expansion	2 x 100GB, RAID1, N/A for Expansion
IOR Read Bandwidth	3.5 GB/s peak; 3.0GB/s sustained	6 GB/s peak; 5.0GB/s sustained

Power Consumption

Rack with Switches	19 kilowatts	20 kilowatts
---------------------------	--------------	--------------

Heat Dissipation

Rack	64,300 BTU	68,300 BTU
-------------	------------	------------

System Availability

Hot Swappable	Disk Drives, Power Supply Units, Fans, Power Cooling Modules, SBB Controller Modules	
----------------------	--	--

Software & Support Information

Software	CSSM, Linux and Lustre included, 1 year renewable	
Hardware	1 year renewable	

General System Environmental Specifications

Altitude and Temperatures	Operational Altitude	-30 to 3048m (-100 to 10,000ft)
	Operating Temperature Range	5-35° C
	Temperature Variance	De-rated by 1° C/300m above 900m below the specified maximum temperature
	Humidity	20% to 80% non-condensing
Power Options Per Rack	US	208V AC 24A, 3 Phase delta, 6 Circuit/Rack, NEMA L21-30P
	EMEA	415V AC 32A, 3+N+E Phases, 4 Circuit/Rack, IEC60309 32A