About Xyratex

Xyratex is a leading provider of enterprise class data storage subsystems and hard disk drive capital equipment. The Networked Storage Solutions division designs and manufactures a range of advanced, scalable data storage solutions for the Original Equipment Manufacturer (OEM) community. As the largest capital equipment supplier to the industry, the Storage Infrastructure division enables disk drive manufacturers and their component suppliers to meet today’s technology and productivity requirements. Xyratex has over 25 years of experience in research and development relating to disk drives, storage systems and manufacturing process technology.

Founded in 1994 in an MBO from IBM, and with headquarters in the UK, Xyratex has an established global base with R&D and operational facilities in Europe, the United States and South East Asia.

Xyratex Headquarters
Langstone Road
Havant
Hampshire PO9 1SA
United Kingdom
T +44 (0)23 9249 6000
F +44 (0)23 9249 2284

Principal US Office
46831 Lakeview Blvd.
Fremont, CA 94538
USA
T +1 510 687 5200
F +1 510 687 5399
www.xyratex.com

©2011 Xyratex (The trading name of Xyratex Technology Limited). Registered in England & Wales. Company no. 03134912. Registered Office: Langstone Road, Havant, Hampshire PO9 1SA, England. The information given in this brochure is for marketing purposes only. It is not a specification. The products and their details are subject to change. For a detailed specification or if you need advice on a specific requirement please contact Xyratex. No part of this document may be reproduced, copied, or used in any form, or by any means, for any purpose (other than for the purpose for which it was designed), without the written permission of Xyratex Technology Limited.
Abstract

The performance requirements of HPC environments with ever-larger compute clusters have placed unprecedented demands on the I/O performance of supporting storage systems. Storage I/O has become the performance bottleneck for many HPC clusters, as traditional storage systems cannot keep pace with the increasing scale of these deployments. Clearly, a state-of-the-art, scalable HPC storage solution is required. What would a high-performance storage architecture that offers a new level of component integration and performance optimization in HPC environments look like?

In this paper we introduce the ClusterStor™ 3000 storage solution, which integrates the highly scalable, parallel Lustre® file system with Xyratex’s innovative Scale-Out Storage Architecture to create the next generation of HPC storage. It outlines how the ClusterStor 3000 delivers unparalleled storage performance in a turnkey, easy-to-manage solution designed to meet the demanding I/O requirements of HPC cluster environments.
Introduction

The HPC market has undergone a paradigm shift. The adoption of low-cost, Linux-based clusters that offer significant computing performance and the ability to run a wide array of applications has extended HPC’s reach from its roots in scientific laboratories to smaller workgroups and departments across a broad range of industrial segments, from biotechnology and cloud computing, to manufacturing sectors such as aeronautics, automotive, and energy.

With dramatic drops in server prices, the introduction of multi-core processors, and the availability of high-performance network interconnects, proprietary monolithic systems have given way to commodity scale-out deployments. Users wanting to leverage the proven benefits of HPC can configure hundreds, even thousands, of low-cost servers into clusters that deliver aggregate compute power traditionally only available in supercomputing environments.

As HPC architecture has evolved, there has been a fundamental change in the type of data managed in clustered systems. Many new deployments require large amounts of unstructured data to be processed. Managing the proliferation of digital data – documents, images, video, and other formats, places a premium on high-throughput, high-availability storage. The explosive growth of large data has created a demand for storage systems that deliver superior I/O performance. However, technical limitations in traditional storage technology have prevented these systems from being optimized for I/O throughput. Performance bottlenecks occur when legacy storage systems cannot balance I/O loads or keep up with high-performance compute clusters that scale linearly as new nodes are added.

Despite the obvious advantage in application performance offered by HPC cluster environments, the difficulty in optimizing traditional storage systems for I/O throughput, combined with architectural complexities, integration challenges, and system cost have been barriers to wider adoption of clustered storage solutions in industrial settings. With the introduction of the ClusterStor 3000, a fully integrated, Lustre-based storage cluster solution, these challenges have been met. The ClusterStor 3000 delivers:

- **Industry-leading IOPS performance and solution bandwidth** - distributed I/O processing and balanced workloads as the system scales linearly

- **Faster, easier system implementation and reliable expansion** - factory-installed and tested components, reduced cable complexity with solution setup in hours, not days or weeks, plus streamlined system expansion with turnkey Scalable Storage Units (additional storage and processing nodes)

- **Fully integrated hardware and software platform** - smaller solution footprint (fewer components, lower cost, and improved investment protection)

- **Simplified cluster management** – ClusterStor Manager automates configuration and node provisioning, and provides performance statistics, system snapshots, and syslog reporting

- **Extreme reliability, availability, and serviceability** - No single point of failure, high-availability, and redundancies with rapid failover
Scale-Out Storage Architecture Overview

At the heart of the ClusterStor 3000 is the Scale-Out Storage Architecture, an advanced storage design that provides higher density and system reliability at lower cost. Fully integrating the Scale-Out Storage Architecture with the Lustre file system creates a next-generation ClusterStor solution that delivers simplified system installation and operation, optimized HPC performance, and non-disruptive cluster expansion.

Limitations of Traditional Server/Storage Systems

Traditional storage clusters are made up of a number of disparate building blocks, including servers to run the file system and associated software, a high-speed storage interconnect, such as InfiniBand, connected to a RAID controller, and a high-density storage system housing the disk. Each subsystem in this hardware/software combination adds complexity and potential bottlenecks in terms of balancing I/O, reliability, and system scalability, as shown in Figure 1.

<table>
<thead>
<tr>
<th>Storage subsystem</th>
<th>Potential Bottlenecks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interconnect fabric IB/10GbE/FC ...</td>
<td>Under-provisioned networks</td>
</tr>
<tr>
<td>File system server(s)</td>
<td>Unbalanced fabrics</td>
</tr>
<tr>
<td>RAID controller(s) HW or SW</td>
<td>SD or DDR InfiniBand</td>
</tr>
<tr>
<td>Disk system SATA/SAS</td>
<td>Gigabit Ethernet</td>
</tr>
<tr>
<td></td>
<td>Old and slow systems</td>
</tr>
<tr>
<td></td>
<td>Lack of memory</td>
</tr>
<tr>
<td></td>
<td>Old software versions</td>
</tr>
<tr>
<td></td>
<td>Too few servers for the underlying storage systems</td>
</tr>
<tr>
<td></td>
<td>Too many disks behind each controller</td>
</tr>
<tr>
<td></td>
<td>Slow disk connectivity (3Gb SAS, 4Gb FC, SATA)</td>
</tr>
<tr>
<td></td>
<td>Too may disks for each expander</td>
</tr>
<tr>
<td></td>
<td>Too little bandwidth available to each drive</td>
</tr>
<tr>
<td></td>
<td>SAS dongle</td>
</tr>
<tr>
<td></td>
<td>SATA drives</td>
</tr>
</tbody>
</table>

Figure 1: Typical storage I/O bottlenecks
ClusterStor Scale-Out Storage Architecture – Distributed I/O Processing at Scale

The ClusterStor Scale-Out Storage Architecture is a consolidated hardware and software environment, designed from the ground up to uniquely address the demand for simple Lustre scalability and performance, which features an integrated Scalable Storage Unit (SSU). Each SSU supports two industry-standard x86 Embedded Server Modules (ESMs), which connect directly through a common midplane to all drives in the SSU, and share a redundant high-speed interconnect across the midplane for failover services. The ESMs are capable of running industry-standard Linux distributions, and each module has its own dedicated CPU, memory, network, and storage connectivity.

When new SSUs are added to the cluster, performance scales linearly as incremental processing network connectivity and storage media are added with each unit. This modular design removes the performance limitation of traditional scale-out models in which servers or RAID heads quickly become the bottleneck as more drives are added to the cluster. The Scale-Out Storage Architecture combines enclosure and server enhancements with software stack optimizations to deliver balanced I/O performance (even on large data workloads), and outperform traditional storage topologies by adding easy-to-install, modular SSUs that scale ESMs as HPC storage scales, distributing I/O processing throughout the solution (see Figure 2).

Figure 2: Traditional storage architecture vs. ClusterStor Scale-Out Storage Architecture
Introducing the ClusterStor 3000

The ClusterStor 3000 is a comprehensively designed storage solution that integrates the Lustre file system, which offers scalable capacity, superior I/O performance, and simplified storage management, with the innovative Scale-Out Storage Architecture, which boosts processing power with storage capacity. Deploying the ClusterStor 3000 results in the highest performance, highest density per square meter storage solution at the lowest cost of ownership, and offers these additional benefits:

- **Comprehensively engineered Lustre storage solution** – manufactured and configured precisely to deliver predictable performance, stability, and reliability
- **Easy installation and configuration** - ClusterStor systems are fully provisioned with a complete software stack and pre-configured for optimal performance and reliability, with installation in hours, not days
- **Fully integrated system management and monitoring** - ClusterStor Manager offers a unified system view to manage every layer of the storage cluster (disks, storage enclosure, RAID data protection, file system, servers, and network), and comprehensively monitor the solution (I/O throughput, system performance, and networking)
- **High-availability and reliability** - ClusterStor systems offer full HA end-to-end in the solution, managed by the ClusterStor Manager, with optimizations to ensure fast service failover.

Highlights of the Lustre File System

The ClusterStor 3000 features the Lustre file system, a client/server based, distributed architecture designed for large-scale compute and I/O-intensive, performance-sensitive applications. The Lustre architecture is used for many different types of clusters, but it is best known for powering 8 of the Top 10 HPC clusters in the world and over 70 of the 100 largest HPC installations, with tens of thousands of client systems, petabytes (PB) of storage, and hundreds of gigabytes per second (GB/sec) of I/O throughput.

Lustre file system scalability has made it a popular choice in the oil and gas, manufacturing, rich media, and finance sectors. Lustre has also been used as a general-purpose data center backend file system at a variety of sites, from Internet Service Providers (ISPs) to large financial institutions; but known complexities in installing, configuring, and administering Lustre clusters has limited broader adoption of this file system technology. With the introduction of the ClusterStor 3000 storage solution, users can now leverage the advantages of the Lustre file system without facing the integration challenges inherent to a multi-vendor environment.

A Lustre cluster is an integrated set of servers that process metadata, and servers that store data objects and manage free space. Together, the metadata and object storage servers present file systems to clients. A Lustre cluster includes the following components:

- **Management Server (MGS)** - The MGS stores configuration information for all Lustre file systems in a cluster. Each Lustre server contacts the MGS to provide information. Each Lustre client contacts the MGS to retrieve information.
• **Metadata Server (MDS)** - The MDS (typically co-located with the MGS) makes metadata available to Lustre clients from the Metadata Target (MDT). The MDT stores file system metadata (filenames, directories, permissions, and file layouts) on disk and manages the namespace. The MDS provides network request handling for the file system.

• **Object Storage Server (OSS)** - The OSS provides file I/O service and network request handling for one or more local Object Storage Targets (OSTs). The OST stores data (files or chunks of files) on a single LUN (disk drive or an array of disk drives).

• **Lustre Clients** – Although not part of the ClusterStor 3000, Lustre clients are computational, visualization, or desktop nodes that mount and use the Lustre file system. Lustre clients see a single, coherent namespace at all times. Multiple clients can simultaneously read and write to different parts of the same file, distributed across multiple OSTs, maximizing the collective bandwidth of network and storage components.

**Overcoming Lustre’s Challenges**

Despite the obvious advantages of using Lustre as a backend file system in HPC deployments, legacy installations have shown limitations and challenges to users. Integrating Lustre into an HPC solution that overcomes these challenges drove the development and design of the ClusterStor 3000.

**Installation and Administration**

Lustre is traditionally installed and administered strictly from a command-line interface, requiring fairly advanced knowledge of Linux administration and experience provisioning clusters and managing distributed nodes. System administrators may lack this knowledge or not understand the complexities of the Lustre software, requiring them to rely heavily on professional services and external support to deploy and administer the file system. Additionally, integrating Lustre into a storage system with heterogeneous hardware adds system and support complexity. It can take weeks to get the system up and running reliably, with additional effort required to tune system performance.

**ClusterStor Advantage:**

ClusterStor systems come installed and configured from the factory, fully optimized to work with Lustre, and ready to deploy right from the loading dock. This innovative HPC storage design leaves behind the complexities of other systems that integrate third-party equipment and require tuning of file system performance in a heterogeneous, multi-vendor infrastructure.

Additionally, the ClusterStor 3000 includes a browser-based GUI that features intuitive wizards to guide administrators through automated Lustre network configuration and user authentication, greatly easing cluster setup. Once the file system is running, the ClusterStor Manager is used to administer the Lustre cluster (including failover and other tasks), monitor operations, analyze performance, gather statistics, and generate reports.
Data Redundancy and Protection

Lustre maintains the integrity of data in the file system, but it does not back up data. Lustre relies exclusively on the redundancy of backing storage devices. To protect data from loss, Lustre users must attach external RAID storage.

ClusterStor Advantage:
The Scale-Out Storage Architecture uses a highly-available RAID 6 array to protect system data and provide optimal fault tolerance and performance. The Lustre file system’s OSSs are configured for active-active failover, enabling all logical drives to be seen by both OSSs at all times. If one OSS node fails, the active OSS takes over OST management and operations of the failed node. In normal non-failure mode, I/O load is balanced between the OSSs.

System Monitoring and Troubleshooting

Lustre generates events to indicate changing states and conditions in the file system, but it does not provide comprehensive file system monitoring. In the event of file system errors, log descriptions may clearly indicate the problem, but many entries are difficult to interpret without Lustre troubleshooting experience. Additionally, Lustre does not isolate file system errors from faults in the client/server hardware, network interconnect, or backing storage system. In a large-scale file system, determining root cause may be difficult.

ClusterStor Advantage:
The ClusterStor Manager provides a unified system view of all Lustre nodes, with combined syslogs available to monitor events in the cluster. To enhance the usability of the Lustre logs, filtering and record sorting is available. Additionally, the ClusterStor Manager rolls up log information into diagnostic payloads, ensuring that support engineers have all the information needed to help customers diagnose and resolve issues quickly.

As part of the ClusterStor 3000, Lustre is deployed on a hardware platform offering sophisticated fault isolation to pinpoint the location of a failure, with easy field serviceability to quickly replace failed parts and system wide high-availability, enabling uninterrupted cluster operations if a component fails.
Software Architecture

The ClusterStor 3000 software architecture consists of an integrated, multi-layer software stack:

- ClusterStor Manager
- Lustre file system
- Data protection layer (RAID)
- System management software (GEM)
- Linux OS

The ClusterStor 3000 runs Lustre 2.1 software in a standard Linux environment (Scientific Linux 6.0 operating system). The file system is fully integrated with the ClusterStor Manager, GEM, and RAID layers in the stack.

ClusterStor Manager

The ClusterStor Manager provides a single-pane-of-glass view of the ClusterStor 3000 infrastructure. It includes a browser-based GUI that simplifies cluster installation and configuration, and provides consolidated management and control of the entire storage cluster. Additionally, the ClusterStor Manager provides distributed component services to manage and monitor system hardware and software.

Simplified Cluster Management and Troubleshooting: The ClusterStor Manager includes intuitive wizards to guide users through configuration tasks and node provisioning. Once the cluster is running, administrators use the GUI to effectively manage the storage environment:

- Start and stop file systems
- Manage node failover
- Monitor node status
- Collect and browse performance data

Additionally, the dashboard reports errors and warnings for the storage cluster and provides extensive diagnostics to aid in troubleshooting, including cluster-wide statistics, system snapshots, and Lustre syslog data.
Health Monitoring and Reporting: To ensure maximum availability, ClusterStor Manager works with GEM, the platform’s integrated management software, to provide comprehensive system health monitoring, error logging, and fault diagnosis. On the GUI, users are alerted to changing system conditions and degraded or failed components.

Data Protection Layer (RAID)

The ClusterStor 3000 uses RAID to provide different data protection layers throughout the solution. The RAID subsystem configures each OST with a single RAID 6 array to protect against double disk failures and drive failure during rebuilds. The 8 + 2 RAID sets support hot spares so that when a disk fails, its data is immediately rebuilt on a spare disk and the system does not need to wait for the disks to be replaced.

This layer also provides cache protection in the event of a power failure. When a power event occurs, Xyratex’s innovative Metis Power Protection technology supplies reserve power to protect in-flight storage data (enabling it to be securely stored on persistent media), a significant advantage over traditional cache memory protection or having to use external UPS devices within the storage rack.

Additionally, the ClusterStor 3000 uses write intent bitmaps (WIBS) to aid the recovery of RAID parity data in the event of a downed server module or a power failure. For certain types of failures, using WIBS substantially reduces parity recovery time from hours to seconds. In the ClusterStor 3000, WIBS are used with SSDs (mirrored for redundancy), enabling fast recovery from power and OSS failures without a significant performance impact.

Unified System Management Software (GEM)

Extensive ClusterStor system diagnostics are managed by GEM, sophisticated management software that runs on each ESM in the SSU. GEM monitors and controls the SSU’s hardware infrastructure and overall system environmental conditions, providing a range of services including SES and HA capabilities for system hardware and software. GEM offers these key features:

- Manages system health, providing RAS services that cover all major components such as disks, fans, PSUs, SAS fabrics, PCIe busses, memories, and CPUs, and provides alerts, logging, diagnostics, and recovery mechanisms
- Power control of hardware subsystems that can be used to individually power-cycle major subsystems and provide additional RAS capabilities – this includes drives, servers, and enclosures
- Fault-tolerant firmware upgrade management
- Monitoring of fans, thermals, power consumption, voltage levels, AC inputs, FRU presence, and health
- Efficient adaptive cooling keeps the SSU in optimal thermal condition, using as little energy as possible
- Extensive event capture and logging mechanisms to support file system failover capabilities and to allow for post-failure analysis of all major hardware components
Hardware Architecture

The ClusterStor 3000 hardware architecture consists of a pre-configured, rack-level storage cluster that can be easily expanded using modular storage node building blocks. The principal hardware components include:

- Cluster Management Unit
- Scalable Storage Unit
- Network Fabric Switches
- Management Switch

The ClusterStor 3000 is differentiated from other file system solutions by its innovative Cluster Management Unit and Scalable Storage Unit architectures.

Figure 4: ClusterStor 3000 in a rack (schematic)
Cluster Management Unit

The Cluster Management Unit (CMU) features the MDS node, which stores file system metadata and configuration information, the MGS node, which manages network request handling, and the ClusterStor Manager software, which is the central point of management for the entire storage cluster, monitoring every storage element within the cluster. The CMU hardware contains a pair of servers, embedded RAID, and one shelf of high-availability shared storage. Cluster interface ports support InfiniBand or 10GbE data networks and 1GbE management network connections.

The CMU is fully redundant and fault-tolerant. Each node is configured for active-passive failover, with an active instance of the node running on one system and a passive instance of the node running on the peer system. If an active node fails, for example, the MDS goes down, then the passive MDS node takes over the MDT operations of the failed MDS. The RAID protects the CMU’s cache and, in the event of a power outage, writes it to persistent storage. The shared storage of the CMU supports a combination of Small Form Factor (SFF) SAS HDD and SSD drives, protected using RAID 1, for management data, file system data, and journal acceleration.

The ClusterStor 3000 supports InfiniBand or 10GbE connections to the MDS and MGS nodes. Each server is configured to operate with either network fabric. Additionally, each server connects, via Ethernet, to dedicated private management networks supporting IPMI.

The CMU provides these key features:

- Industry-leading IOPS platform
- SFF SAS HDD and SSD support
- Redundant network links per node
- Data protection/integrity (RAID 1 and 10)
- Redundant flash MDS journal disk
- Scalable MDS (as SSUs are added and system expands)
- Centralized HA management node for all storage elements in the cluster

![Cluster Management Unit (schematic)](image-url)
Scalable Storage Unit

The ClusterStor 3000’s core building block is the Scalable Storage Unit (SSU). Each SSU is configured with identical hardware and software components, and hosts two OSS nodes. The SSU contains two ESMs, with RAID-protected, high-availability shared storage, and interface ports to support InfiniBand or 10GbE data networks and 1GbE management network connections.

The ESMs are Storage Bridge Bay (SBB) compliant with SAS expanders that enable both modules to directly access all drives in the enclosure (a differentiator among fully fault-tolerant systems). The ESMs connect through a common midplane, eliminating the need for extra external cables, and share a redundant, high-speed interconnect across the midplane for failover services. This efficient and highly reliable design enables the SSU’s SAS infrastructure to deliver robust performance and throughput – 2.5 GB/sec per SSU for reads and writes.

The SSU is fully redundant and fault-tolerant, thus ensuring maximum data availability. Each ESM serves as a Lustre OSS node, accessing the disk as shared OST storage and providing active-active failover. If one ESM fails, the active module manages the OSTs and the disk operations of the failed node. In non-failure mode, I/O load is balanced between the modules. The OSS cache is protected by Xyratex’s unique Metis Power Protection technology; in the event of a power outage, the OSS can write its cache to persistent storage. The shared storage of the SSU consists of high-capacity SAS disk drives, configured in a RAID 6 array to protect against double disk failures and drive failure during rebuilds.

SSU Platform

The SSU platform is the Xyratex-designed, 5U 84 drive, ultra-dense storage enclosure, purpose built for the demands of HPC applications. The 5U 84 enclosure’s innovative design reflects extensive Xyratex knowledge, developed from pioneering the industry’s first high-density storage system design in 2003 (the 4U 48 RA-4835), and substantial expertise derived from designing five generations of SAS-based storage systems.

The SSU platform includes unique dampening technologies that minimize the impact of rotational vibration interference (RVI). Maintaining disk drive performance while reducing RVI is a key design challenge in high-density storage system design. All Xyratex enclosures, including the SSU platform, maximize disk drive performance by mitigating RVI sources, including cooling fans and other disk drives, and other enclosures mounted in the same rack.

Within the SSU enclosure, all drives are individually serviceable and hot swappable. Additionally, each drive is equipped with individual drive power control, enabling superior availability with drive recovery from soft errors. The SSU platform uses “Nearline” SAS-based drives — offering the cost/capacity benefits of traditional, high-capacity SATA drives, but with a native SAS interface to mitigate data integrity risks and performance limitations associated with using SATA as the drive interface protocol. Additionally, the SAS drives are natively dual-ported with multi-initiator support, to facilitate the fast and reliable failover of drives between OSSs and overcome the need for discrete SATA/SAS multiplexor modules, which are required when using SATA drives in high-availability architectures.

Disk drive reliability is a key determinant in storage cluster performance and uptime. The disk drive modules in the SSU are CERT-qualified and offer significantly enhanced disk drive reliability. CERT is Xyratex’s unique disk drive testing
process that leverages over 30 years of drive reliability experience from our Disk Drive Test equipment division. This proprietary process has been proven to offer an Average Failure Rate (AFR) that is 3-4 times lower than the industry norm.

**SSU ESM**

The ClusterStor 3000 supports up to two ESMs in each SSU. Each ESM has its own dedicated x86 CPU complex with an integrated RAID XOR engine, high-speed, low-latency cache (protected by Metis Power Protection technology), and SAS controllers/switches with either 40G QDR InfiniBand or 10GbE data network host connections. Additionally, each ESM connects, via 1GbE ports, to the dedicated management and IPMI networks. The SSU enclosure includes multiple high-speed intercontroller links across a common midplane for communication between ESMs for synchronization and failover services. Each ESM is a Field Replaceable Unit (FRU) and is easily accessible at the rear of the SSU enclosure for field service and upgrade.

The SSU offers these key features:

- Industry-leading density/capacity/performance platform (5U 84 high-density enclosure with 3.5" drives)
- Enterprise class, fault-tolerant design
- High-capacity and high-performance SAS drive support
- Redundant network links per SSU
- Massive internal SAS bandwidth
- High-availability, active-active ESMs
- Data protection/integrity (RAID 6, 8+2) with two dedicated global hot spares
- Redundant flash OSS journal disks
- Extensive diagnostic services
- Metis Power Protection technology

![Figure 6: Scalable Storage Unit (schematic)](image-url)
Network Fabric Switches

The Network Fabric Switches (InfiniBand or 10GbE) manage I/O traffic and provide network redundancy throughout the ClusterStor 3000. To maximize network reliability, the ESMs in the SSU are connected to redundant network switches. If one switch fails, the second module in the SSU (connected to the active switch) manages the OSTs of the module connected to the failed switch.

Additionally, to maintain continuous management connectivity within the solution, the network switches are fully redundant at every point and interconnected to provide local access from the MDS and MGS nodes to all storage nodes.

Management Switch

The Management Switch consists of a dedicated local network on a 1GbE switch (with an available redundant second switch option) that is used for configuration management and health monitoring of all components in the ClusterStor 3000. The management network is private and not used for data I/O in the cluster. This network is also used for IPMI traffic to the SSU’s ESMs, enabling them to be power-cycled by the ClusterStor Manager.

Figure 7: ClusterStor 3000 network connections (schematic)
Conclusion

The ClusterStor 3000 is an integrated Lustre storage appliance that delivers extreme I/O performance in a highly-available, linearly scalable, reliable, easy-to-manage solution. With a comprehensively engineered design, featuring an innovative scale-out storage architecture and on-board Lustre file system optimized for performance and reliability, ClusterStor 3000 users can realize increased application performance and improved productivity in HPC cluster environments.
About Xyratex

Xyratex is a leading provider of enterprise class data storage subsystems and hard disk drive capital equipment. The Networked Storage Solutions division designs and manufactures a range of advanced, scalable data storage solutions for the Original Equipment Manufacturer (OEM) community. As the largest capital equipment supplier to the industry, the Storage Infrastructure division enables disk drive manufacturers and their component suppliers to meet today’s technology and productivity requirements. Xyratex has over 25 years of experience in research and development relating to disk drives, storage systems and manufacturing process technology.

Founded in 1994 in an MBO from IBM, and with headquarters in the UK, Xyratex has an established global base with R&D and operational facilities in Europe, the United States and South East Asia.

Xyratex Headquarters
Langstone Road
Havant
Hampshire PO9 1SA
United Kingdom
T: +44 (0)23 9249 6000
F: +44 (0)23 9249 2284

Principal US Office
46831 Lakeview Blvd.
Fremont, CA 94538
USA
T: +1 510 687 5200
F: +1 510 687 5399

www.xyratex.com

©2011 Xyratex (The trading name of Xyratex Technology Limited). Registered in England & Wales. Company no: 03134912. Registered Office: Langstone Road, Havant, Hampshire PO9 1SA, England. The information given in this brochure is for marketing purposes and is not intended to be a specification nor to provide the basis for a warranty. The products and their details are subject to change. For a detailed specification or if you need access to a specific requirement, please contact Xyratex. No part of this document may be transmitted, copied, or used in any form, or by any means, for any purpose other than for the purpose for which it was delivered, without the written permission of Xyratex Technology Limited.