THE SUN™ I/O STACK AND ITS BENEFITS FOR EMC STORAGE

White Paper
September 2006

Abstract

To fully support Fibre Channel (FC) SANs and new SAN features, Sun Microsystems developed a new, open, standards-based I/O framework and device driver stack called the Sun StorageTek™ SAN Foundation software (commonly known by its code name Leadville). Sun is committed to preserving standards compliance for all layers of the stack to give organizations the flexibility to choose the storage components that best suit their needs, and Sun and EMC are cooperating at the engineering level to help EMC comply with these standards. This paper describes the Sun I/O stack and its benefits for EMC storage.

Table of Contents

Executive Summary

Most operating systems in use today were designed before storage area networks (SANs) were widely implemented. Prior to SANs, storage was directly attached (DAS) and communicated with servers through a static interface such as Small Computer System Interconnect (SCSI). Unfortunately, even though storage architectures are moving to SANs, many operating systems still treat the storage devices as SCSI devices. The problems with continuing to use a SCSI driver framework are:

- The need to reboot to add more storage — no flexibility
- A small number of LUNs (Logical UNits) — no scalability
- No multipathing — limited availability
- Host Bus Adapters (HBAs) are not hot pluggable — impacting availability
- Requires a driver for each device — increasing complexity and manageability
- Requires the administrator to manually edit configuration files — increasing deployment times, errors, and administration costs
- Requires the administrator to add third-party patches — increasing deployment times, errors, and administration costs

To fully support Fibre Channel (FC) SANs and new SAN features, and alleviate the problems mentioned above, Sun Microsystems developed a new, open, standards-based I/O framework and device driver stack called the Sun StorageTek™ SAN Foundation software (commonly known by its code name *Leadville*). It includes the Leadville I/O stack, Sun StorageTek™ Traffic Manager software (multipathing or MPxIO), Storage Networking Industry Association (SNIA) libraries, and utilities. The Leadville stack is now fully integrated into the Solaris™ 10 Operating System (OS) and provides:

- ***The ability to dynamically add new devices without rebooting*** — so there is no need to interrupt applications to add capacity or change a failed device
- ***Automatic persistent binding*** — hosts with access to the same tape drives can use the same logical names, simplifying administration
- ***Virtually limitless scalability*** — systems can see thousands of devices on the FC fabric
- ***Integrated multi-pathing*** — for availability and scalability
- ***No third-party patches*** — resulting in lower cost, easier installation and configuration, and a simpler support model

Established in 1993, EMC's global partnership with Sun has provided thousands of joint customers with advanced information storage management solutions on Solaris Operating System platforms. Sun and EMC have come to an agreement on a number of hardware, software, and support initiatives designed to help enterprises more easily deploy technologies together. Both companies are collaborating on interoperability to help ensure compatibility between the EMC family of networked storage platforms and the Solaris 10 Operating System.

This alliance helps ensure the highest levels of solution performance, availability, and cost-effectiveness, along with a superb total customer experience through:

- A new cooperative support agreement to provide one-stop, 24/7 mission-critical support to EMC and Sun joint customers.
- Industry-leading interoperability testing between EMC's eLab and Sun servers to help ensure consistent reliability and storage connectivity between Sun and EMC products.
- EMC is qualifying its hardware platforms and porting storage software to the Solaris 10 OS on both SPARC® and AMD Opteron™ x64 processor-based systems. Shared qualification of products means innovative new products go to market faster.
- EMC is porting agreed-upon software components to the Solaris 10 OS.
- Engineering cooperation between EMC, Sun, and HBA vendors to implement Leadville for the best performance and reliability.

From Sun's perspective, the Leadville stack exemplifies Sun's commitment to providing open, standards-based technologies. Sun is committed to preserving standards compliance for all layers of the stack to give organizations the flexibility to choose the storage components that best suit their needs, and Sun and EMC are cooperating at the engineering level to help EMC comply with these standards. Standardizing on the Leadville I/O stack delivers the following benefits to enterprises choosing to implement a Sun and EMC environment:

- ***Common driver set for HBAs*** — Helps reduce complexity and management tasks by eliminating the need to buy, install, and track firmware and patches for each HBA and array.
- ***Better I/O stack*** — Since Sun develops the operating system for its servers, Sun is in the ideal position to develop a better I/O connection, or stack, into its servers. In essence, Sun is providing a clean, well standardized, open interface point at the SAN port on Sun servers with single vendor accountability, reducing complexity and simplifying support for customers who choose to use storage from other vendors such as EMC.
- ***Sun and EMC move in lock-step to introduce new HBAs*** — In general, when vendors introduce new products it can take quite some time for other vendors to test, qualify, and support those products. Because Sun is responsible for the interface all the way to the SAN port on the HBA, and has a cooperative agreement with EMC, the two companies can quickly introduce new HBAs and functionality to the market.
- ***Greatly simplifies interoperability matrix between Sun and EMC products*** — With the Leadville I/O stack, the interoperability matrix on the EMC side is extremely simplified, making it easier to purchase and implement a joint solution.
- ***Best value in multipathing*** — Since the Leadville I/O stack is integrated into the Solaris OS, there is no added cost and it is also fully supported by Sun.

The Leadville I/O stack was first introduced by Sun in 2000 and has now been integrated into the Solaris 10 Operating System. Since its introduction many Sun customers, including major businesses in the financial and telecommunications industries, have been successfully using the Leadville I/O stack to lower costs and the cost of managing storage, as well as easily integrating heterogeneous storage based on industry standards. And, with the Sun and EMC agreement, these businesses can implement a solution where the server is completely supported by Sun to the Fibre Channel plug and EMC storage is supported by Sun via the support agreement for one-stop support.

Chapter 1
Introduction

A Little History

The frameworks for most operating systems in use today were designed before IT started migrating storage architectures to SANs. Operating system designers created the storage I/O framework based on the common storage method at the time — DAS connected to servers through static interfaces such as the SCSI protocol. While DAS and SCSI are excellent for single systems with limited storage needs such as desktops and small, isolated workgroup servers, the technologies present the following limitations for networked storage architectures such as SAN:

- Device sizes are fixed and LUNs cannot be added without rebooting
- SCSI-based storage devices support a relatively small number of LUNs
- No multipathing support — SCSI-based storage devices are typically single ported and visible through only one path at a time, limiting scalability and failover capability
- HBAs and storage interconnects are not hot-pluggable — impacting availability and flexibility
- There is no concept of zoning or LUN-masking — affecting security

Prior to the Solaris 8 OS, and like other operating systems of its time, the Solaris OS had a driver framework optimized for SCSI DAS. The operating system scanned storage at boot time and created a directory tree reflecting how storage was physically connected to the server through SCSI busses, HBAs, etc. The OS assumed this tree would not change until the system was powered down for maintenance. The administrator was required to create and maintain device driver configuration files (`sd.conf`, etc.) to tell the framework the properties of each storage device because the SCSI protocol did not provide a mechanism for devices to report themselves or their properties.

When the device tree was created at boot time, the OS used the driver configuration file to create a node for each line in the file and passed this node to the driver to probe the device. If the probe was successful, a file in the `/devices` directory was created. Configuring devices in this manner was not efficient and had an adverse effect on boot time because it required probing of each line in the configuration file, even if a device was no longer attached to the system. It also consumed more memory because in order to allow for devices that were offline during the boot process, the nodes were not destroyed for devices that were probed but not present.

Fibre Channel SANS

As performance improvements in computer systems spawned increasingly data- and network-intensive applications, the limitations of SCSI were exacerbated. The Fibre Channel standard was developed to keep pace with improvements in computer

networking, addressing the need for very fast transfers of large amounts of data to storage devices connected to multiple systems. Fibre Channel enabled storage vendors to create more dynamic and highly available solutions using large pools of consolidated storage, as opposed to DAS. Characteristics of a typical Fibre Channel environment are:

- Storage devices can dynamically connect, disconnect, or move to a new port
- SAN configuration changes are possible by adding or reconfiguring paths to storage devices
- Storage devices support multiple ports per LUN and multiple paths to each LUN
- Storage can grow dynamically and expose new LUNs
- Servers can see thousands of LUNs via the SAN
- Large, central storage devices can allow remote booting by many servers
- Hot-pluggable SAN HBAs and interconnects can be used for higher availability

Although the nature of the physical storage interconnects changed from SCSI to FC, the device drivers for early Fibre Channel HBAs are written to the old SCSI driver frameworks. In other words, the operating system SCSI driver framework still views the new FC devices as SCSI devices, forcing the FC configuration to appear as a DAS SCSI configuration. This use of drivers to bridge SCSI and Fibre Channel works well enough to meet initial needs. However, as enterprise SANs grow in scale and complexity, this method presents several problems, particularly in rack-based environments with large numbers of small servers. For example, storage applications still follow a path based on the direct physical connection, even though multiple, redundant connections might be available, negating any load balancing or failover capabilities. Another limiting factor is the continued requirement to manually modify device driver configuration files and reboot the system in order to change the configuration. Finally, the number of components in the I/O stack, e.g., the Solaris OS, VERITAS Volume Manager, PowerPath, Qlogic or Emulex driver, HBA, and EMC storage, increases the complexity of the solution and are all supported by their respective vendors.

New Leadville Framework

Beginning with the Solaris 8 OS, Sun added capabilities to fully support Fibre Channel SANs, making the Solaris Operating System Fibre Channel aware via the Leadville framework and driver stack. Key new features include transparent multipathing support, dynamic device node creation, support for 16,000 LUNs, and dynamic reconfiguration. The I/O framework and device driver stack provide the following capabilities:

- Present persistent device nodes for storage, regardless of the physical path or configuration changes after the node is created
- Automatically identify devices and create device nodes without requiring manual administration

- Suspend I/O to an HBA and allow it to be removed and replaced without taking the system off-line
- Provide high availability through multipathing software
- Allow seamless operation with both legacy and new storage applications
- Present thousands of LUNs to a single host
- Allow fabric boot from large, central storage devices

Prior to the Solaris 10 OS, the Leadville stack was available separately as Sun StorageTek SAN Foundation software. As of the Solaris 10 OS, it is fully integrated into the operating system, making it even easier for system administrators to use, and making it available on SPARC, x64, and x86 systems. With the Solaris 10 OS, instead of all of the components listed earlier, the storage stack consists of the Solaris OS, a Sun HBA, and EMC storage — and all of these components interoperate properly and are fully supported by Sun. In addition, because the Leadville stack is integrated into the operating system, it is part of all Solaris OS update releases, enabling continuous innovation and predictable quality.

Chapter 2

The Leadville I/O Stack

Figure 1 depicts an overview of the Leadville FC stack in the Solaris Operating System. The stack is well integrated into the Solaris OS for high performance, and provides a simplistic means of extending the stack to support new features of the FC and SNIA standards as they emerge. The layers that make up the stack are described below.

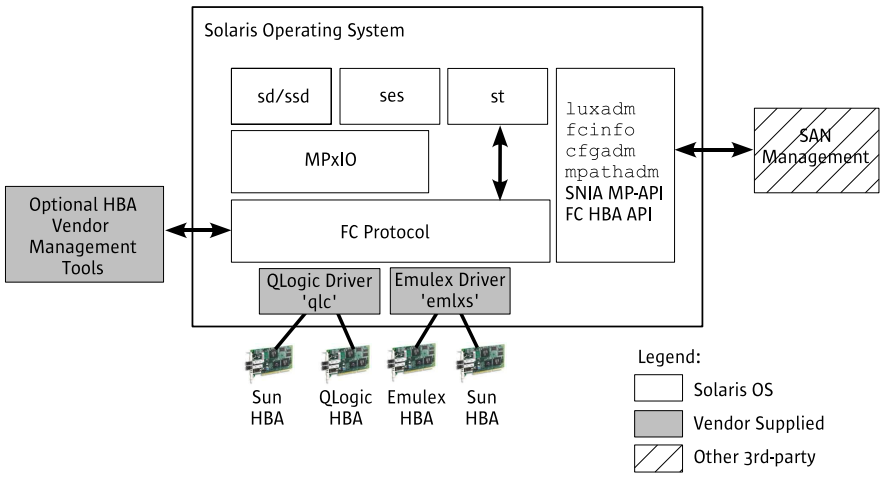

Figure 1. Basic Stack Diagram

Vendor Drivers

The first layer, working from the bottom of Figure 1 up, consists of the HBA vendor drivers. The Leadville framework provides a well-defined interface to enable HBA vendors to quickly support new cards and FC technology without needing to understand the internals of the Solaris OS. This allows the vendors to focus on developing new hardware because they spend minimal time coding a light-weight driver to interface into the Leadville framework for complete functionality. These vendor drivers are included in the Solaris 10 OS and are written to include support for all HBAs sold by Sun or the vendor. This is called the Leadville *Common HBA Driver*. A benefit to EMC is that they only qualify one driver stack regardless of the number of cards they list in their ESM support matrix or test in the eLab.

FC Protocol Layer

At the heart of Leadville I/O stack is the Fibre Channel Protocol layer, which is a common Sun component acting as an interface to provide consistent behavior for how systems running the Solaris OS attach to a SAN. The FC Protocol layer provides a rich interface to the FC HBAs and reduces the complexity of SANs by utilizing well-known interfaces into the core of the Solaris OS.

This layer is designed to provide the following key benefits:

- *Automatic discovery* — automatic device node creation without the need for rebooting. This eliminates the need for `.conf` file entries as required by SCSA drivers.
- *Consistent and persistent binding* — utilizing entries in `devlink.tab` files, logical links can be defined as needed by the user.
- *Supports 16,000 LUNs* — the FC Protocol layer correctly utilizes the FC addressing structures without internal limitations imposed by SCSI drivers (which limits the number of LUNs to 255).
- *Supports fabric boot* — OS images can be stored on remote storage devices.
- *Supports dynamic reconfiguration* — if the user needs to replace or add an HBA, it can be accomplished without shutting the server down.
- *Low footprint* — the layer is well-integrated into the Solaris OS and utilizes all available interfaces to perform its function without inventing new interfaces or utilizing non-public interfaces.
- *Standards-compliant* — the layer is developed and tested against the FC T11 protocol, rather than proprietary HBAs, switches, or storage devices.
- *Part of the native operating system* — no patches are required to integrate with other features of the Solaris 10 OS, such as Solaris ™ ZFS, Solaris™ Containers, Sun™ Cluster software, disaster recovery, and all other features of the Solaris 10 OS. Because the stack is integrated into the operating system, it benefits from intensive testing process that the Solaris OS undergoes before every release, across all products. This helps to ensure the highest quality and ease of use for customers. It also provides consistency and coherency in `syslog` messages among the host driver components.
- *Common interface for management APIs/CLIs* — the behavior of the management tools does not vary by HBA. It is now consistent across HBAs. These standard APIs enable Solaris system utilities, such as `fcinfo` or `mpathadm`, to provide basic FC HBA and multipathing support in a common fashion regardless of which HBA the devices are connected to.

MPxIO

MPxIO is the multipathing driver within the Solaris Operating System. It is protocol independent, but written to support FC as the first protocol. The driver is located just above the FC protocol layer and below the target drivers. This layer provides the following key benefits:

- Availability — the key benefit is having MPxIO provide multipath support by not allowing broken paths to interrupt I/O operations. Since this module is built into the Solaris OS, path failure is detected and recovered from very quickly.

- Native array support — most of the FC arrays in the industry are supported by this module including all EMC arrays, as well as standard support for symmetrical and asymmetrical (T10 Asymmetric LUN Access or T10 ALUA) devices. T10 ALUA does not require any proprietary failover to be encoded in the Solaris stack. It simply works for the array vendors that provide the support for these devices.
- Load-balancing — MPxIO provides several options for load-balancing including round-robin and preferred paths (with T10 ALUA).
- Efficient enumeration — because the driver is located below the target drivers, as the FC Protocol layer auto-discovers devices, MPxIO can efficiently create devices nodes without duplication due to multiple paths. This simplifies administration and utilizing less memory as new paths are added.
- Co-existence with all other multipathing drivers — MPxIO can be disabled by the user allowing third-party software to be employed.
- Aggregation — performance can be increased by adding more paths.

Target Drivers

All target drivers are supported by the FC stack in the Solaris OS. The disk driver is supported by the SSD driver although there is no difference between the SSD and the SD (SCSA) driver (both are built from the same source code base) with the exception of not needing to create or maintain `.conf` file entries with the SSD driver. The ST (tape) driver does not use MPxIO but benefits from all of the other features of the FC Protocol layer. A particularly useful benefit is consistent binding for the ST driver. This enables hosts with access to the same tape drive to use the same logical name for higher-level applications and system administrator convenience.

Management Interfaces

Leadville provides open, standard interfaces to administer Solaris OS-based systems connected to a SAN. Both T11.5 and SNIA standard interfaces are supported — the Multipath Management API (MP API) and FC HBA API are integrated into the Solaris OS along with CLIs encompassing these APIs to enable easier administration. SNIA MP API defines standard interfaces for multipath devices, associated path discovery, and path administration on a host. Because the APIs are open standards, this allows EMC storage and Leadville management APIs to be utilized by higher-level applications such as Sun StorageTek Enterprise Storage Manager software and AppIQ. The management tools that are built into the Solaris OS are:

- MP API — provides a single management view of co-existing multipathing solutions. It provides for multipathing administration, as well as path enable and disable. The CLI built on top of this API is `mpathadm`.
- FC HBA API — provides FC HBA local and remote port discovery, FC HBA identification, and device discovery. The CLI built on top of this API is `fcinfo`.

- `luxadm` — this CLI provides SCSI Enclosure Services (SES) administration on the storage subsystems within the Sun Fire™ V490 and V890 systems, as well as older Sun StorEdge A5x00 arrays. It also provides support for MPxIO-enabled FC devices on systems running the Solaris 8, 9, and 10 OS.
- `cfgadm` — this CLI provides the ability to dynamically reconfigure hardware resources. `cfgadm_fp` is the `cfgadm` plug-in that enables configuration support for FC ports, allowing FC fabric devices to be dynamically reconfigured.

Chapter 3

Multipathing

Sun StorageTek Traffic Manager software, or MPxIO, is the native multipathing solution built into the Solaris Operating System. It operates between the target drivers and the individual protocol stacks, as illustrated in Figure 2. Today, MPxIO provides multipathing for the following protocol stacks in the Solaris OS: Fibre Channel, iSCSI, and Infiniband.

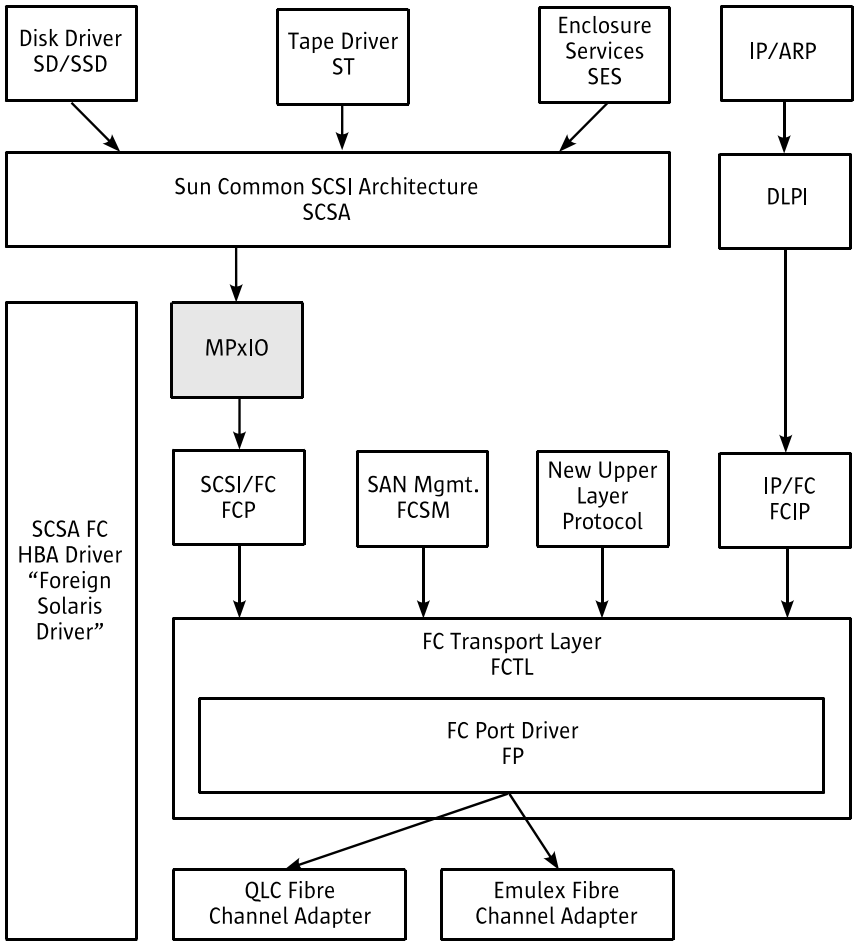

Figure 2. Leadville and MPxIO Block Diagram

MPxIO virtualizes and abstracts multiple paths to a disk or LUN. Instead of representing a device node for each and every path to a single device, a single device node is created that represents the sum of all paths to the disk or LUN. MPxIO provides a path-independent name in both the `/dev` and `/devices` directories and uses only one device node of the corresponding device driver (ssd, st, etc.). This is ***not*** true with third-party *foreign Solaris drivers*, which might utilize more than one node and in addition must provide equivalent multipathing support. The advantage of using MPxIO is that

applications, volume managers, file systems, and databases do not need to be aware of multiple paths. Other advantages over third-party solutions include:

- Command throttling — target device driver queues are maintained per device and not administered per path. This reduces the potential for over-runs on the target device.
- Bandwidth aggregation — routing I/O through multiple connections can increase bandwidth by adding more controllers. In addition, using dynamic reconfiguration on the server side, more HBAs, and therefore more paths, can be added without shutting the server down.
- Dynamic device node creation — as new devices are discovered or removed on the SAN, the resulting node create or destroy is performed automatically within Leadville and MPxIO. This reduces system administration.

Standards for the Sun FC Stack and MPxIO

MPxIO and Leadville are based on open standards for communicating with devices, device management, and multipathing. The following T11 and T10 standards, listed in Table 1 and Table 2 are supported:

Table 1. T11 Standards

Title	Standard
FC-GS-4	ANSI INCITS 387-2004
FC-DA	ANSI/INCITS TR-36-2004
FC-FS	ANSI/INCITS 373:2003
FC-AL	ANSI X3.272:1996
FC-HBA	INCITS 386:2004

Table 2. T10 Standards

Title	Standard
SCSI Primary Commands — 3 (SPC-3)	ANSI INCITS 408-2005
SCSI Block Commands — 2 (SBC-2)	INCITS 405-2005
SCSI Enclosure Services (SES)	ANSI INCITS 305-1998/AM1-2000 [R2003]
SCSI Stream Commands — 2 (SSC-2)	ANSI INCITS 380-2003
SCSI-3 Architecture Model (SAM)	ANSI INCITS 270-1996
SCSI Fibre Channel Protocol — 2 (FCP-2)	ANSI INCITS 350-2003
Multipath Management API	INCITS 412

MPxIO and Storage Support

To discuss how MPxIO supports different types of storage, it is first helpful to understand the difference between them. Most storage arrays that implement redundant controllers fall into one of two categories. At the high-end are symmetric arrays, such as the Sun StoragTek™ 6920 system and EMC Symmetrix, which support

concurrent access to volumes through two or more SAN ports. The second category is asymmetric arrays, which allow access through only one port at a time.

MPxIO storage devices can be classified into one of four categories:

- Symmetric — all paths to the storage device are active and I/O commands can be issued through any path.
- Standard asymmetric — some paths to storage are active while others are standby. Standby paths cannot accept all commands. The mechanism to switch a path state from standby to active or vice-versa is as specified by SCSI standards — SCSI III primary commands (SPC 3) for asymmetric devices.
- Sun proprietary asymmetric — some paths are active while others are standby as in the standard asymmetric category. The mechanism to switch path state from standby to active or vice versa is Sun proprietary. For example, through Sun StorEdge T3 array explicit LUN failover.
- Non Sun proprietary asymmetric — some paths to storage are active while others are standby as in the standard asymmetric category. The mechanism to switch path state from standby to active or vice versa is proprietary to a vendor other than Sun.

MPxIO provides support for all Sun storage devices as well as all third-party storage solutions that offer fully symmetrical controllers, such as EMC Symmetrix and any third-party storage solution that offers standard asymmetric target port group support. In addition, MPxIO also provides proprietary failover support for EMC CLARiiON arrays.

EMC PowerPath

EMC PowerPath is a server-based software solution that combines multiple path I/O capabilities, automatic load balancing, and path failover functions into one integrated package. PowerPath is an independent software layer that sits above the target drivers. PowerPath creates virtual devices that manage multiple paths to a storage system. Applications reference the virtual device and PowerPath allocates paths to the storage system. PowerPath's key features are:

- Multiple paths — PowerPath enables multiple paths from the host to logical devices allowing access to the logical device even if a specific path is unavailable. Multiple paths can also be used to share I/O workloads and improve performance.
- Load balancing — PowerPath intelligently allocates I/O requests across all available paths with a dynamic load-balancing algorithm. A number of load balancing policies are supported, including Symmetrix/CLARiiON optimization, adaptive, round-robin, least I/Os, least blocks, and basic failover.
- Proactive I/O path testing and automatic path recovery — PowerPath periodically tests failed paths and automatically restores them when they are operational again. It also periodically tests idle paths to report and address defective paths.
- Automatic path failover — PowerPath automatically redirects I/O to alternate paths in the event of a failure.

- High availability cluster support — PowerPath's path failover capability avoids node failover in clustered environments.

MPxIO and PowerPath Co-Existence

Multipathing with MPxIO can be enabled or disabled on a system-wide or per-port basis for finer control. If multipathing is enabled on a specific HBA controller port, all supported devices connected to the port have multipathing enabled on them. Configuring multipathing by port enables MPxIO to co-exist with other multipathing solutions such as EMC PowerPath. However, devices and paths should not be shared between MPxIO and other multipathing solutions. Co-existence can occur only when MPxIO is managing non EMC storage and PowerPath is controlling all EMC storage.

Chapter 4
iSCSI Through MPxIO

The Solaris 10 Operating System, Update 1 includes an iSCSI initiator (client of an iSCSI network) device driver designed to enable the Solaris OS to read and write to storage (iSCSI target) through any TCP/IP interface, including a standard Ethernet network interface card (NIC). iSCSI is an IP-based storage networking standard for linking data storage subsystems. The iSCSI protocol is implemented underneath the SCSI protocol and on top of the TCP/IP protocol. Multiple TCP connections are allowed and can be used for load balancing or failover. iSCSI can potentially be used to access native iSCSI storage, as well as Fibre Channel storage through bridges.

Although iSCSI includes native multipathing as an optional feature (Multiple Connections per Session or MC/S), it is not widely implemented by other vendors and is not currently supported in the Solaris OS. To provide some multipathing support, Solaris IP multipathing (IPMP) can be configured for failover by configuring two or more NICs for higher availability, as shown in Figure 3. However, though it has the capability to multipath NICs, it is limited in that it cannot multipath the target ports on the storage devices. This is where MPxIO becomes useful.

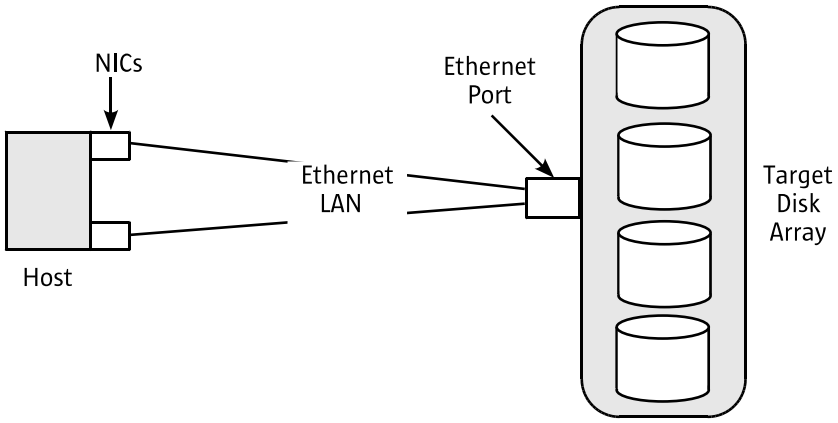

Figure 3. IPMP — Failover and Load Balancing at the NIC

Because MPxIO operates above the protocol layer, it can support Fibre Channel, Infiniband, and iSCSI. Fibre Channel and iSCSI drivers register LUNs with MPxIO, which matches paths to the same logical unit at the SCSI protocol layer. As explained before, MPxIO hides multiple paths to one device, presenting only one view of a device to the target driver and storage applications. The iSCSI initiator driver registers an instance with MPxIO for each LUN for every unique target port identifier (consisting of a concatenation of the target node name and the target portal group tag). For example: iqn.1721-02.com.sun.1234,1, where iqn.1721-02.com.sun.1234 is the target node name and 1 is the target portal group tag.

iSCSI targets can share a common IP address and port, or can have multiple IP addresses. In order for MPxIO to enumerate more than one path for an iSCSI target, the iSCSI initiator must establish multiple iSCSI sessions to that target, as illustrated in Figure 4. One way to do this is to allocate ports into multiple target portal groups to make the target port identifier unique. A portal group is a set of ports that support a coordinated session across the ports. In this case two ports with the same IP address could have unique target portal identifiers. For example, iqn.1721-02.com.sun.1234,1 and iqn.1721-02.com.sun.1234,2.

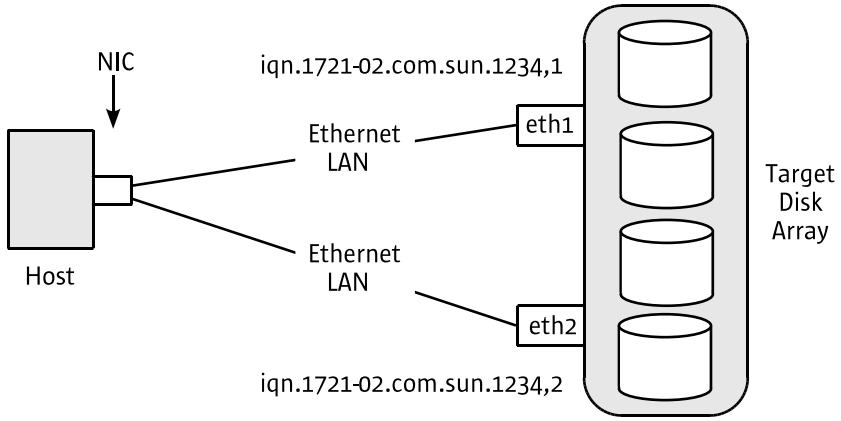

Figure 4. Utilizing Multiple Target Portals

If an iSCSI target supports the native iSCSI multipathing (MC/S), it presents all of its target portals in a single target portal group in which all ports form one logical SCSI target port. The Solaris iSCSI initiator driver registers only one instance of a LUN with MPxIO for each logical SCSI target port.

In addition, because MPxIO is above the protocol layer, it can support different protocols to the same device. For example, one LUN could appear to the host via Fibre Channel and iSCSI paths and MPxIO utilizes both paths, as depicted in Figure 5. Note, however, that because MPxIO utilizes round robin load balancing, that MPxIO is most useful where all of the links between the host and storage have equal bandwidth and latency.

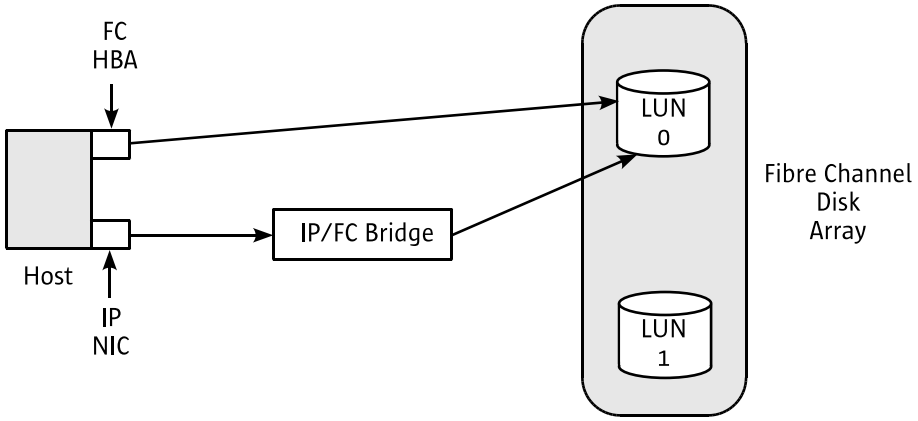

Figure 5. Multiple Protocols to the Same Device

For more information on iSCSI in the Solaris Operating System see: *Using iSCSI Multipathing in the Solaris 10 Operating System, Sun BluePrints™ OnLine, September 2005*.

Chapter 5
Conclusion

As one of the first proponents and premier innovators in network computing, it should be no surprise that Sun is once again leading the pack — creating ingenious, standards-based, cost-effective solutions to today's networked storage problems. Sun StorageTek SAN Foundation software, featuring the Leadville I/O stack and MPxIO, provides an open I/O stack and driver set, as well as multipathing support for heterogeneous storage arrays and systems, bringing the following benefits to Sun and EMC joint customers:

- No need to reboot when adding or reconfiguring — for industry-leading availability
- Support for thousands of devices — for industry-leading scalability
- Integrated multipath I/O — decreases cost and complexity, increases availability
- Integrated native SAN stack — decreases cost and complexity
- 64-bit SPARC and AMD Opteron support — platform choice and performance
- Fibre Channel and iSCSI connectivity — for more choice in hardware
- Validated, standards compliance — better interoperability, lower cost and complexity
- Single stack for all SAN devices — decreases cost and complexity, and increases utilization

The bottom line for IT and storage managers is things just got easier, less expensive, and more highly available on a heterogeneous platform. Isn't that just what every CIO is asking for?

For More Information

For more information on the topics contained in this paper, see the following Web sites and papers.

- Sun and EMC Alliance: *http://www.sun.com/emc*
- Using iSCSI Multipathing in the Solaris 10 Operating System, Sun BluePrints OnLine, September 2005
- Increasing Storage Area Network Productivity, Sun BluePrints OnLine, July 2004
- Solaris Fibre Channel and Storage Multipathing Administration Guide, 819-0139-10, available on *http://docs.sun.com/app/docs/doc/819-0139*
- T10 Technical Committee: *http://www.t10.org*
- T11 Technical Committee: *http://www.t11.org*

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA **Phone** 1-650-960-1300 or 1-800-555-9SUN (9786) **Web** sun.com

Printed in USA□XX/XX□ XX0000-0/