



Redundancy in enterprise storage networks using dual-domain SAS configurations

technology brief

Abstract.....	2
Introduction.....	2
Why dual-domain SAS is important	2
Single SAS domain	3
Dual-domain SAS.....	4
Examples of redundant architectures using serial storage.....	5
Dual-domain SAS.....	5
Dual-domain SAS: servers in two-node cluster.....	6
HP dual-path with cascaded JBODs.....	7
Summary	8
Appendix A: Nomenclature and definitions	9
Dual-path	9
Dual-port	9
JBOD	9
Multiple SCSI domains.....	9
Port Identifier	9
phy.....	10
SAS domain	10
SCSI domain.....	10
SCSI initiator.....	10
SCSI initiator port.....	10
SCSI target port	10
Appendix B: Enabling technology.....	11
SAS and SATA technology.....	11
Connectors.....	12
Backplane	13
Devices.....	13
Drives and controllers	13
Expander	13
Cluster Interconnect	13
Firmware and Software	14
Serial Tunneling Protocol (STP).....	14
SPC-3 Target port group commands	14
For more information.....	15
Call to action	15

Abstract

This technology brief explores the issue of single points of failure within storage networks and how pathway redundancy addresses this problem. Appendix B of this brief explains the technologies that enable redundant storage networks and why the implementation is dependent upon serial drive technology. The differences between Serial Attached SCSI (SAS) and Serial ATA (SATA) drives are explored as well. This brief also provides insight into how HP will implement redundant storage networks in various enterprise configurations.

Introduction

Using dual-domain and dual-path architectures creates redundant pathways from servers to storage devices. The redundant paths created by these configurations reduce or eliminate single points of failure within the storage network. The technologies discussed in this brief employ single and dual-ported serial hard drives and dual controllers, or dual-ported controllers, to provide dual-paths from the server. This provides increased levels of high availability with redundant paths from the controller to the drives.

Dual-domain SAS implementations make it possible to tolerate host bus adapter (HBA) failure, external cable failure, expander failure, failure in a spanned disk (JBOD) environment, and failure in Redundant Array of Independent Disks (RAID) environments.

Dual-path SATA implementations provide some lower cost solutions that cover cable failure, but not the full redundancy of a SAS dual-domain solution. The dual-path solution uses a single domain method of providing tolerance to cable failure.

One of the key differences between dual-domain and dual-path architectures is the use of SAS or SATA drives. Dual-port SAS drives can provide the advantages of complete pathway redundancy in a dual-domain configuration. Customers seeking lower cost solutions can employ single port SATA drives in a dual-path configuration.

Why dual-domain SAS is important

The INCITS T10 Technical Committee defines a SCSI domain as “the interconnection of two or more SCSI devices and a service delivery subsystem.”¹ When any part of the pathway in this single domain fails, there is no secondary, or redundant, pathway to allow continued data transfer. Historically, the rates of failure are higher in connectors, cabling, and highest of all in hard drives. The higher failure rates are the result of the mechanical connection points and moving parts inherent in these components.

Dual-domain SAS creates an additional domain to address this single domain pathway failure. The additional domain uses an open port on an HP Smart Array controller that is capable of supporting dual-domain SAS. The second port on the dual-domain capable Smart Array controller generates a unique identifier and can support its own domain. To take advantage of multiple domains, SCSI devices must be dual-ported and connected to pathways in both domains. SAS drives can meet that requirement.

Dual-domain SAS solutions continue to offer higher reliability, performance, and data availability than SATA, while SATA provides higher capacities and a lower cost point. In addition the Enterprise class SAS drives, HP has broadened the hard drive offerings with a new Midline class of SATA and SAS drives. These Midline drives are high capacity, ranked in cost as lowest dollar per GB for their performance class, and designed with higher reliability and performance than entry level drives. These drives are appropriate for high-capacity environments such as external storage, backups,

¹ INCITS T10 Technical Committee- <http://t10.t10.org/ftp/t10/drafts/smc2/smc2r07.pdf>

archival, reference, and redundant highly-available applications. They are intended for use in an environment where the drive workload is 40 percent or less.

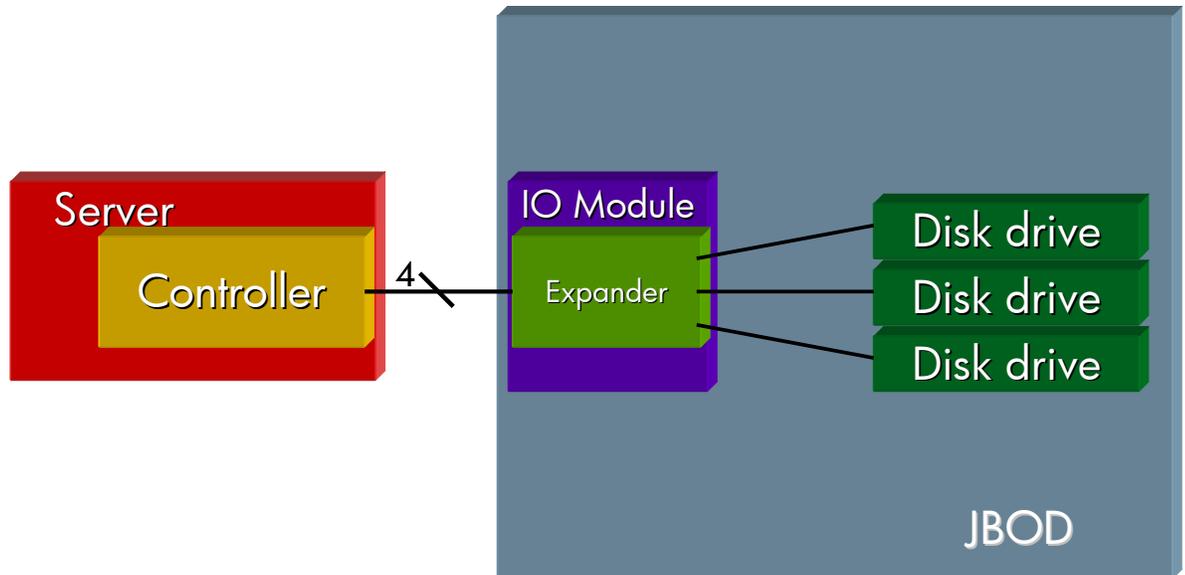
Single SAS domain

There are many single points of failure in a single domain. Failure can originate from any controller, expander device, JBOD or cable, as shown in Figure 1.

NOTE:

JBOD is the acronym for “Just a Bunch of Disks.” It refers to a group of hard disks, usually in an enclosure, that may or may not be RAID configured.

Figure 1. Single domain SAS storage network



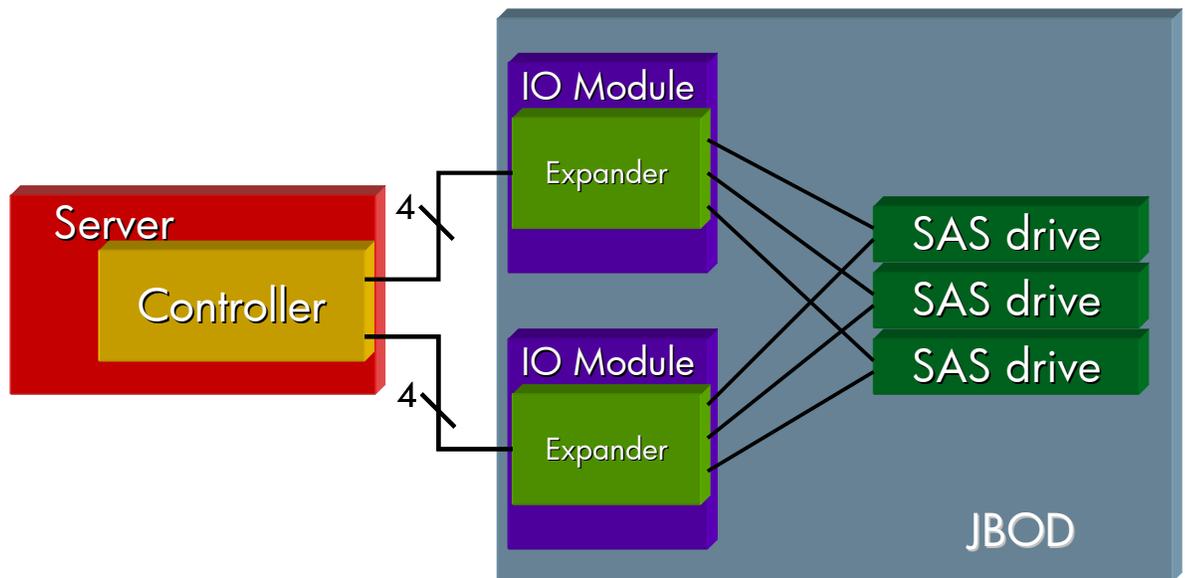
NOTE:

All figures in this technical brief using the “4↗” notation indicate a 4-lane bundled path. There are 4 wires per lane, or 16 total wires in a 4x cable.

Dual-domain SAS

Using dual-ported drives and dual controllers, or a controller capable of supporting dual domains, provides dual-path redundancy from the server. Figure 2 shows an example of a single controller capable of dual-domain support. The dual-domain SAS configuration in Figure 2 can tolerate simultaneous single port failure in a dual-port, dual-domain capable controller, external cable failure, and expander failure.

Figure 2. SAS dual-domain SAS to JBOD storage network



NOTE:

Currently, the P800 controller provides dual-domain support

For additional information on Serial technology and how it makes dual-domain SAS possible, see "Enabling technology" in Appendix B.

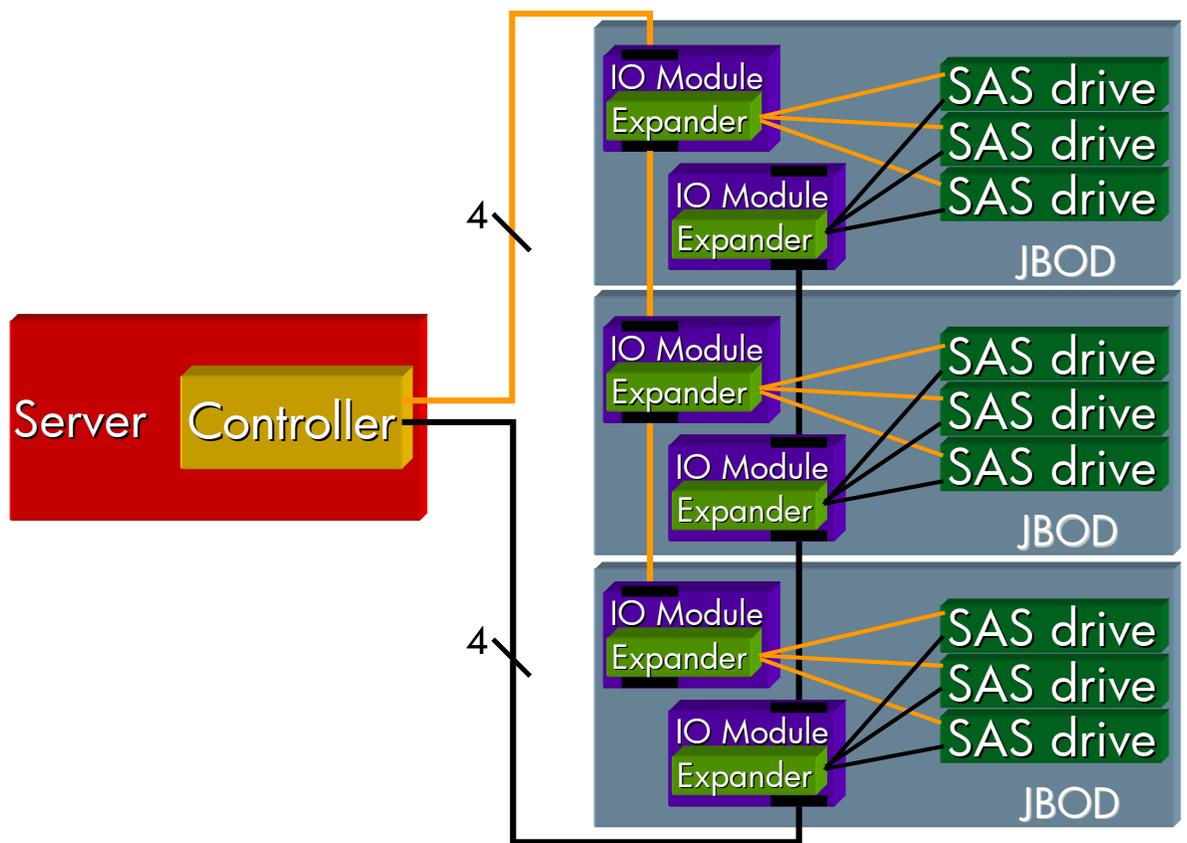
Examples of redundant architectures using serial storage

This section illustrates some of the redundant storage scenarios made possible by serial storage, Smart Array, and cluster computing technologies.

Dual-domain SAS

This dual-domain SAS architecture provides redundant pathways for cascaded JBODs. Using two IO modules for each enclosure and SAS dual-port drives, this configuration provides redundancy throughout the storage network and eliminates any single point of failure.

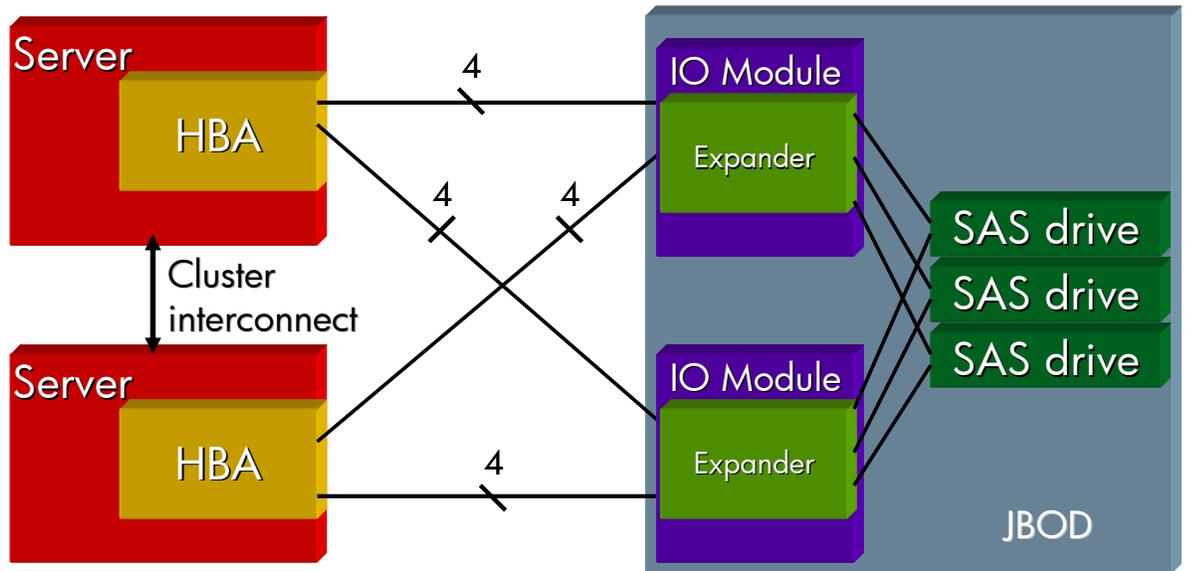
Figure 3. Dual-domain SAS for cascaded JBODS



Dual-domain SAS: servers in two-node cluster

This configuration, which is also known as a high availability (HA) cluster, at least one server must have access to the storage network. The cluster interconnect provides redundancy in the event of host bus adapter (HBA) or cable failure, as shown in Figure 4. Dual-domain SAS requires "active/active" configurations. The "active/active" configurations permit both controllers to process I/Os, and either controller can act as a standby. To use the dual-ports for redundant pathways, all drives must be SAS.

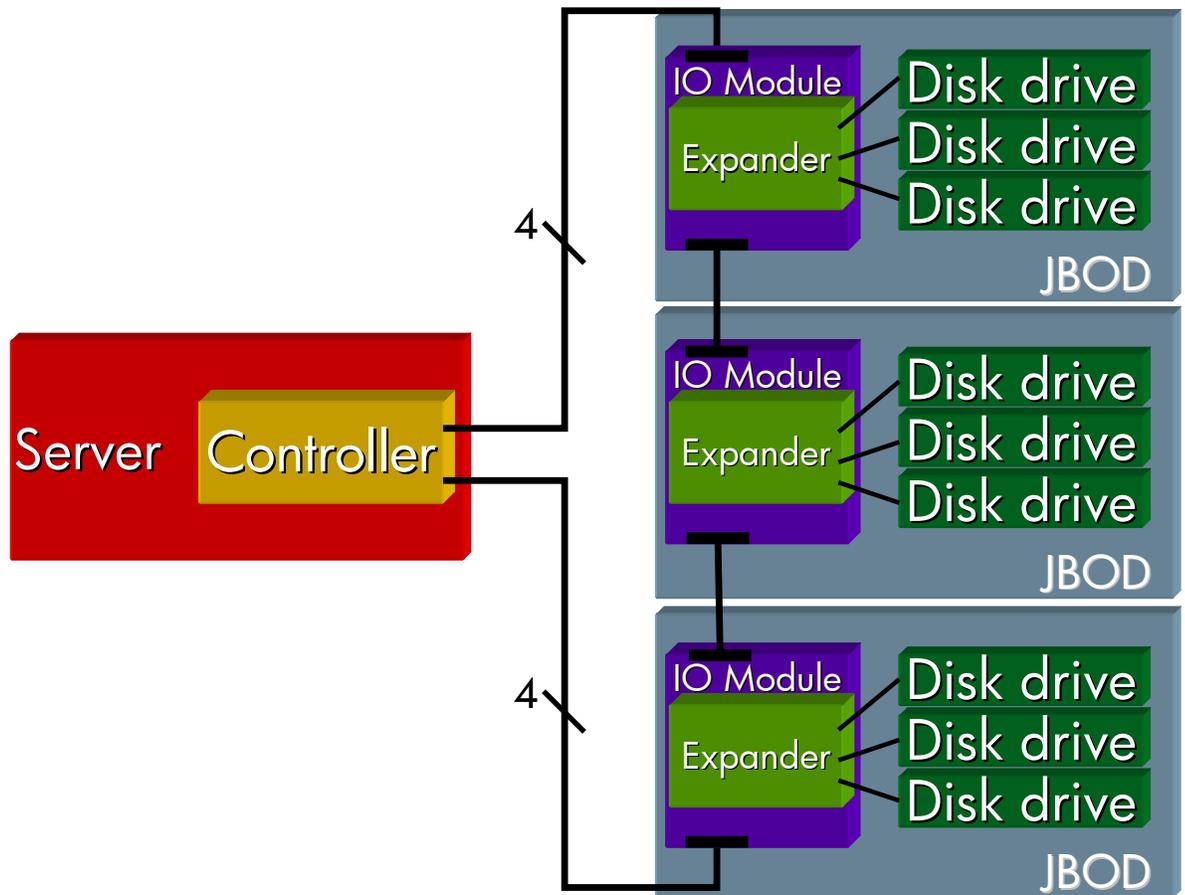
Figure 4. Multi-Path, 2-Node, HA cluster configuration



HP dual-path with cascaded JBODs

The dual-path approach can prevent a single point of failure occurring in complex enterprise configurations such as cascaded JBODs. The configuration is supported for both SAS and SATA drives. In this configuration a controller is connected to the IO module at each end of a set of cascaded JBODs. Each JBOD contains a single expander, as shown in Figure 5. This configuration, which uses half of the IO modules and expanders required for other redundant storage, also reduces costs.

Figure 5. Dual-path scenario for cascaded JBODs



Summary

HP introduced dual port SAS hard drives during the last half of 2007. HP is now introducing products that support SAS dual port operation. As of this writing, the Smart Array P800 controller is capable of supporting two domains. The HP StorageWorks Dual Domain I/O Module Option Kit enables the 60 Modular Smart Array and 70 Modular Smart Array to support redundant pathways from servers to storage devices. See the "For More Information" section for more complete information on all of these products.

The dual-domain and dual-path configurations reviewed in this paper can provide effective solutions for systems administrators and other customers looking for higher levels of redundancy, reliability, and increased data availability within the storage network. Storage network redundancy achieved through the implementation of dual-domain can eliminate single points of failure and make it possible to tolerate HBA failure, external cable failure, expander failure and failure in a spanned disk (JBOD) environment. This can be a crucial enhancement for businesses such as healthcare, manufacturing, and retail, where data availability is essential to uninterrupted operation.

Appendix A: Nomenclature and definitions

New technologies generate new descriptions and nomenclature surrounding those technologies. This section provides definitions for related concepts, technologies, and infrastructure terms related to storage network redundancy and dual-domain SAS configurations.

Dual-path

Dual, redundant paths between initiator and target ports can be in the same or different SCSI domains.

Dual-port

Target devices, such as SAS hard drives, usually incorporate two Phys, or ports. The drives are referred to as “dual-port” because each Phy reports a unique SAS address.

JBOD

The acronym stands for “Just a Bunch of Disks”, JBOD typically refers to a group of hard disks, usually in an enclosure, that are not RAID configured.

Multiple SCSI domains

If two ports cannot communicate, then they are in different SCSI domains

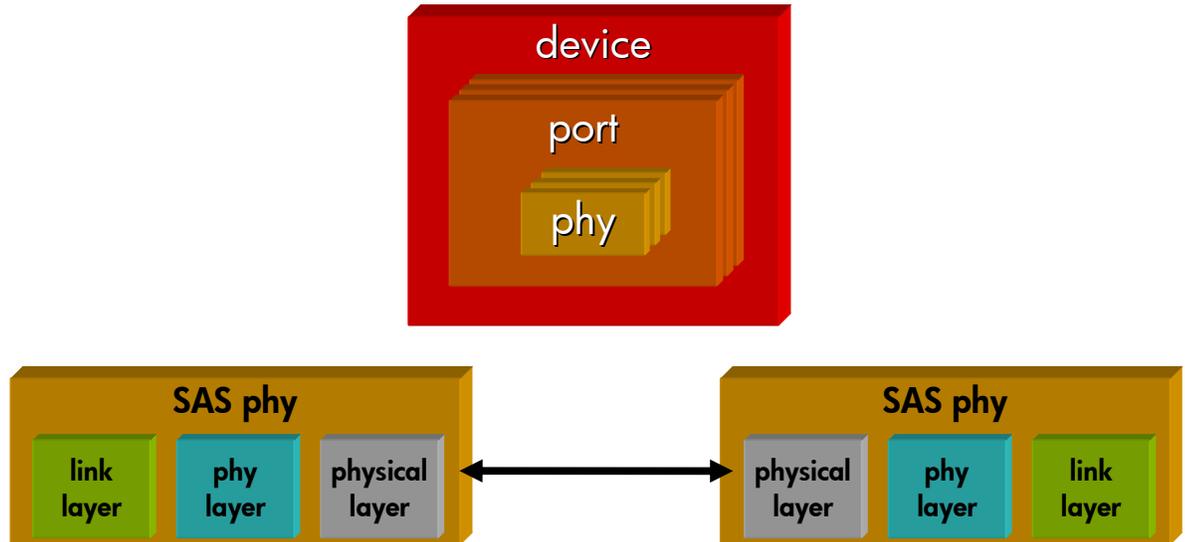
Port Identifier

This is a value by which a SCSI port is referenced within a domain. The SCSI port identifier is either an initiator port identifier or a target port identifier.

phy

phy is a generic electronics term referring to an electronic integrated circuit or a functional block of a circuit that encodes and decodes transmissions between a pure digital domain (on-off) and a modulation in the analog domain. A 'SAS phy' is a combination of the physical layer, phy layer and link layer functions. A minimum of two phys (one at the initiator, the other at the target) is required to complete a SAS physical connection pathway, as shown in Figure A-1.

Figure A-1. phys are required to complete a SAS physical connection



SAS domain

A SAS domain consists of one or more SAS devices and expander devices and can be contained within a SCSI domain

SCSI domain

A SCSI domain contains a service delivery subsystem and one or more SCSI devices. Additionally, this domain contains a set of SCSI initiator ports and target ports that can communicate with each other when interconnected by a service delivery subsystem.

SCSI initiator

The SCSI initiator is the host-side endpoint of a SCSI session. SCSI initiators request data transfers from SCSI targets. An HBA is an example of a SCSI initiator.

SCSI initiator port

A SCSI initiator port acts as the connection between application clients and the service delivery subsystem through which requests and responses are routed. In all cases when this term is used it refers to an initiator port or a SCSI target/initiator port operating as a SCSI initiator port.

SCSI target port

A SCSI target port contains a task router. It acts as the connection between device servers and task managers and the service delivery subsystem through which requests and responses are routed. When this term is used it refers to a SCSI target port or a SCSI target/initiator port operating as a SCSI target port.

Appendix B: Enabling technology

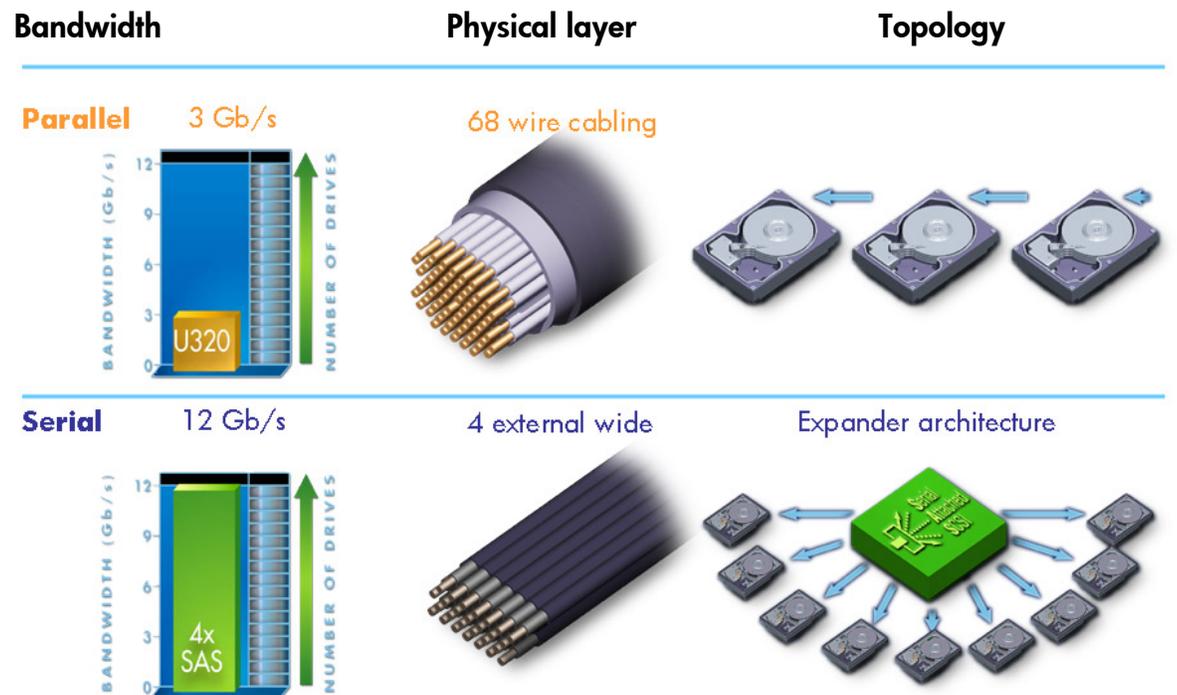
The advent of SAS and SATA drives makes the implementation of dual-domain SAS technology possible. Hardware and protocol commonality enables the redundancy and reliability of dual-domain SAS technology and reduces the number of possible single points of failure within a system.

SAS and SATA technology

Parallel storage technology has reached hard limitations with respect to volume and speed. The physical architectural differences found in SCSI and ATA drives also mean that there are limitations for connectivity. Serial technology has eliminated these connectivity barriers to connectivity (see Figure B-1). Serial attached SCSI and ATA are complementary technologies based on a universal interconnect, so SAS users can choose to deploy cost-effective SATA drives in a SAS storage environment.

Requirements for physical connectivity have dropped from the 68 wires per cable required by parallel technology to 4 wires for serial. In addition to reducing the size of wiring and connectors, serial technology reduces the incidence of crosstalk, signal skew, and other forms of signal error found in parallel storage networks. Serial "star" topology, enabled by expanders, is more flexible, easier to manage, and common to both SAS and SATA drives.

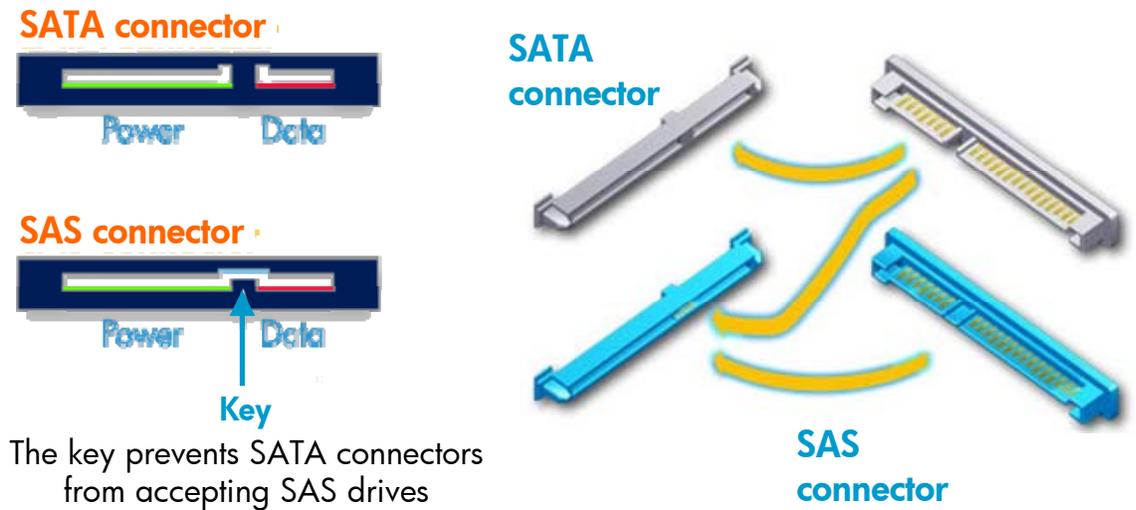
Figure B-1. Comparison of parallel and serial technologies



Connectors

SATA drives are compatible with SAS connectors, but SATA connectors will not accept SAS drives (Figure B-2). Preventing SAS drives from connecting to a SATA network is important because SATA drives are, without the intervention of port selectors, single-ported and cannot accommodate the dual-ported pathways used by SAS drives. These blind-pluggable connectors enable a common topology including controllers, HBAs, backplanes and both SAS and SATA drives. Such commonality permits the implementation of dual-domain SAS technology by providing a single standard for all storage network pathways.

Figure B-2. SAS and SATA connectors



Backplane

The same keyed connection layout is applied to system backplanes allowing them to accept both SAS and SATA drives (Figure B-3).

Figure B-3. Common backplane for Serial drives



Devices

The following devices make up dual-domain SAS networks. They also fulfill requirements for switching and data flow over the available paths.

Drives and controllers

Controllers, HBAs and hard drives (SAS drives) all require dual-ports to have a dual-domain configuration. SATA drives can act as dual-port devices with the intervention of appropriate port selectors or switches.

Expander

An expander is a device used to facilitate communication among large numbers of SAS devices. Expanders contain two or more external expander ports. Each expander device contains at least one SAS Management Protocol target port for management and may contain SAS devices. For example, an expander may include a Serial SCSI Protocol target port for access to a peripheral device. An expander is not necessary to interface a SAS initiator and target, but a connected expander helps a single initiator communicate with more SAS and SATA targets. By analogy, expanders are like an Ethernet hub in a network that allows connection of multiple systems using a single port in the network. Expanders can be implemented as ASIC (application-specific integrated circuit) components on initiator devices such as an HBA.

There are two different types of expanders: Edge Expanders and Fanout Expanders.

- An Edge Expander allows the SAS initiator to communicate with up to 128 SAS addresses.
- A Fanout Expander can connect up to 128 sets of Edge Expanders, known as an Edge Expander Device Set, to allow addressing even more SAS devices.

Cluster Interconnect

A computing cluster must have a dedicated cluster interconnect to which all cluster members are connected. This interconnect serves as a private communication channel between cluster members and HP Smart Array controllers directing data traffic on existing domains

The cluster interconnect is normally used for five high-level functions:

- Cluster Executive components
- Cluster I/O
- Cluster Networking
- Reliable Datagrams
- Application-specific traffic.

Firmware and Software

A storage network with redundant paths must employ protocols and command sets that provide a basis for communication and retrieval of information about the state of each pathway and all of the devices on those pathways. Further, these protocols and command sets must have the ability to direct traffic by controlling switches and links to dictate “active and inactive” pathways.

Serial Tunneling Protocol (STP)

STP allows a SAS host to communicate directly with SATA Targets (SATA devices addressable in a SAS Expander topology). When data is directed to a SATA drive connected to a SAS backplane through an edge expander, an STP connection immediately opens to enable SATA frames to pass through the connection to the drive. STP operates transparently in the storage network with virtually no impact on system throughput.

SPC-3 Target port group commands

Within the dual-domain SAS network of available paths between initiator and target devices, a command set must exist to control the switches and links causing these paths to be set in active, inactive or standby states. The SPC-3 Target port group commands serve those functions.

A target port group is a set of target ports that are always in the same state (for instance, a multi-ported controller card).

A “logical unit group” is a set of logical units (LUs) whose target port groups are always in the same state. For example, in an active-optimized/active-non-optimized RAID controller, one controller may be optimized for one set of LUs and the other for the rest of the LUs.

For more information

For additional information, refer to the resources listed below.

Resource description	Web address
Serial Attached SCSI Architecture Overview	http://www.scsita.org/aboutscsi/sas/tutorials/SAS_General_overview_public.pdf
HP StorageWorks 60 Modular Smart Array - Overview & Features	http://h18002.www1.hp.com/storage/disk_storage/msa_diskarrays/drive_enclosures/msa60/index.html
HP StorageWorks 70 Modular Smart Array - Overview & Features	http://h18002.www1.hp.com/storage/disk_storage/msa_diskarrays/drive_enclosures/msa70/index.html
Modular Smart Array dual domain Q&A	http://h18002.www1.hp.com/storage/disk_storage/msa_diskarrays/drive_enclosures/msa60/qa.html#13
HP Smart Array P800 controller Overview & Features	http://h18004.www1.hp.com/products/servers/proliantstorage/arraycontrollers/smartarrayp800/index.html
HP SAS Drives Questions & Answers	http://h18004.www1.hp.com/products/servers/proliantstorage/serial/sas/questionsanswers.html

Call to action

Send comments about this paper to TechCom@HP.com.

© 2008 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

TC080501TB, May 2008

