

Managing Access Control Through SAS Zoning

White Paper

by Heng Liao, Tim Symons, Rachelle Trent

September 2005

© PMC-Sierra, Inc.

1 Managing Access Control

Serial Attached SCSI (SAS) is gaining popularity in small storage area network (SAN) server environments. With its rise in popularity comes the need to segregate and manage device traffic in a similar fashion to what is already done in larger Fibre Channel networks by using zones or in Ethernet using virtual LANs. By doing this, IT administrators can create much more flexible, scalable, and efficient server networks that meet their business needs. SAS zoning, a proposal to the T10 Technical Committee for inclusion in the SAS-2 specification, provides this capability.

1.1 Ease of SAS

SAS offers many benefits for enterprise IT infrastructures that need to operate leanly and with high productivity and flexibility. SAS, a point-to-point protocol, is a high performance, scalable, flexible and economical solution for deploying storage systems.

SAS is a replacement for parallel SCSI, maximizing the ease with which data storage capacity and throughput is created. Its target application is direct attached storage (DAS) systems where one server is connected to multiple disks. However, SAS provides a powerful switching capability using *expanders*, which act as switches to end devices, enabling quick aggregation of many drives in a single SAS domain (up to 16,384 devices). These expanders are fully capable of connecting multiple hosts to multiple targets. SAS has gained significant interest as a mechanism for connecting large groups of targets in small storage area networks (SANs), in cluster or in blade server environments, allowing these servers to share resources across their targets.

In these larger server environments, it is necessary to provide more management and control services. Frequently, the traffic needs to be segregated for efficiency. Furthermore, since storage resources are shared, server resources need to be controlled by putting limitations on which resources each server can control preventing unauthorized access. As well, not all servers need to access all data. Some data may need to be dedicated on one server; other data may need to be shared across multiple servers. Access controls, such as those provided for the SAS zoning specification, prevent unauthorized access, malicious attacks or corruption of data by operator error on the server. Access controls ensure that if a server is compromised, only data that is accessible by the compromised server is at risk of being lost and not the entire network.

The SAS transport protocol supports the following protocols over its serial interface:

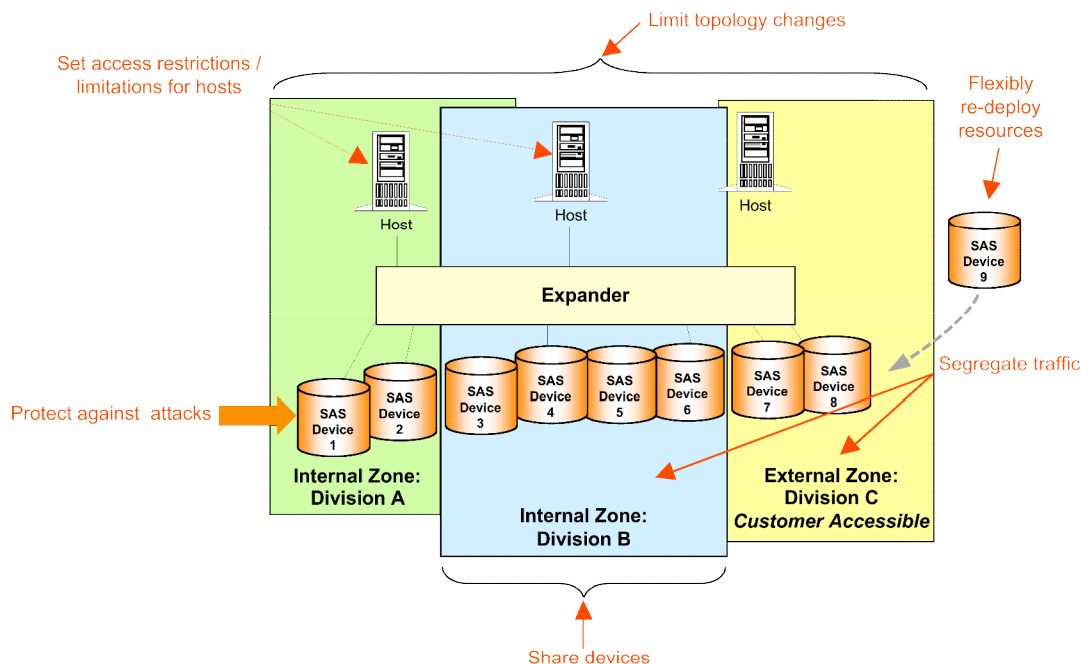
- Serial SCSI Protocol (SSP) to communicate with SAS devices and existing SCSI software
- SCSI Tunneling Protocol (STP) to identify and communicate with SATA devices
- SCSI Management Protocol (SMP) to manage SAS' point-to-point topology

The zoning features discussed in this paper are managed by an extension to the SMP commands.

1.2 Benefits of SAS Zoning

What are the benefits for access control and what makes SAS zoning attractive? To answer these questions, envision a SAS expander topology as shown in Figure 1.

Figure 1 SAS Access Control Features



Access control provides *traffic segregation* for data, functions and broadcast traffic. It logically separates traffic between hosts and resources and provides common access privileges to all end devices. There are many ways to segregate traffic. Figure 1 shows how traffic can be segregated between internal and external zones.

Access control provides *flexible re-deployment of resources*. For example, if one hard disk drive (HDD) is assigned to a server in a group, but more capacity is required for that server, it is easy to add another disk to the group.

Access control provides *controlled sharing of resources* and imposes *access restrictions/limitations*. Users can limit the resources that each host “sees” by configuring different policies for each SAS zone. By assigning SAS PHYs into groups and applying access control policies to restrict these groups, the system designer can ensure that only authorized users can access certain parts of the system. By grouping end devices, system designers can also save on the amount of resources required for an expander implementation.

Zoning allows users to limit access to the SMP control plane such that only the authenticated management devices can issue SMP control commands.

Access control *limits the impact of topology change*. A SAS network uses a mechanism called *topology discovery* to determine which devices are part of the topology. Each time a new device is added, removed, or lost, a broadcast event is generated to notify all the expanders and host devices that re-discovery must be performed to determine which device has changed. This is a time consuming process and requires both the hosts' and expanders' resources as well as increases SMP traffic. It makes sense, therefore, to limit the impact of topology changes so that when a device is added or removed, only its host device has privileges to see the topology change

Finally, access control also provides *protection against attacks* by limiting the propagation of Broadcast traffic. Broadcast events are very disruptive and consume a great amount of link bandwidth when being sent frequently, also known as a *broadcast storm*. If broadcast storms are not limited, any "misbehaving" device in the topology can disrupt the operation of the entire network. It becomes difficult to support larger networks if broadcast storms are not controlled.

1.3 How Zoning Works

Access control functionality is fully implemented in the expanders. Expanders are used for control as it is difficult to know if a host is authorized or not. Therefore, access control does not require hosts to intervene or change their behavior. By allowing expanders to control zoning, legacy SAS and SATA devices, which do not understand zoning, can operate within the SAS domain.

From the perspective of a SAS system administrator, the zoning model requires no change to the end devices in the network. Initiators continue to perform normal SAS discovery, and initiators and targets send and receive OPEN address frames as usual. However, unlike a typical SAS system, initiators and targets do not see the entire SAS domain, also known as a service delivery subsystem. Instead, they only see the portions of the domain, otherwise known as groups, that they have been given permission to see based on a *permission table* that is configured for each zoning expander. Zoning operation is determined by the configuration made to the zoning attributes for each PHY port of the expander (called *PHY zone configuration*).

2 The Zoning Implementation

To enable zoning in a network, the system designer/administrator must:

- Implement the service delivery subsystem using *zoning expanders*. (Note, although up to 128 zones are supported by the SAS-2 zoning proposal, the number of supported zones differs for each zoning expander vendor.)
- Define the portions or groups required in the zoning domain by analyzing the devices that need to share common access privileges. This information must then be translated into the zoning permission table contained within the expander. The permission table defines whether communication is allowed between groups. Group assignment is strictly based on the expander ports of the *zoning service delivery subsystem*. The service delivery subsystem can be a portion of a SAS domain or a complete SAS domain that provides normal services of a SAS system including zone management and access control services. Each PHY can only belong to one group.
- Define the attributes for each expander PHY:
 - Define the *supervisor(s)* for the service delivery subsystem. The supervisor is a management entity that is either an expander device inside the zoning subsystem or a device attached to the subsystem. The supervisor is an SMP initiator that is capable of generating SMP commands for SAS zoning configuration and management. More than one supervisor can exist in a fabric. Note that a *supervising expander* is used to coordinate the supervisors is automatically elected based on the largest SAS address in the topology and is responsible for propagating zone permission table changes to all zoning expanders in the subsystem.
 - Designate the PHYs attached to the zoning expanders as trusted or not trusted. Devices inside the boundary of the zoning subsystem are designated as trusted whereas those outside the subsystem are listed as untrusted. This exercise defines the boundary of the zoning service delivery subsystem.

The SAS-2 zoning proposal that has been submitted to the T10 Technical Committee provides complete details about the updated SMP commands that are used for zone management configuration and topology discovery.

New and Modified SMP Commands for SAS Zoning

- REPORT GENERAL function
- DISCOVER function
- CONFIGURE PHY ZONE function
- CONFIGURE ZONE PERMISSION function
- REPORT ZONE PERMISSION function
- REPORT ZONE ROUTE TABLE function

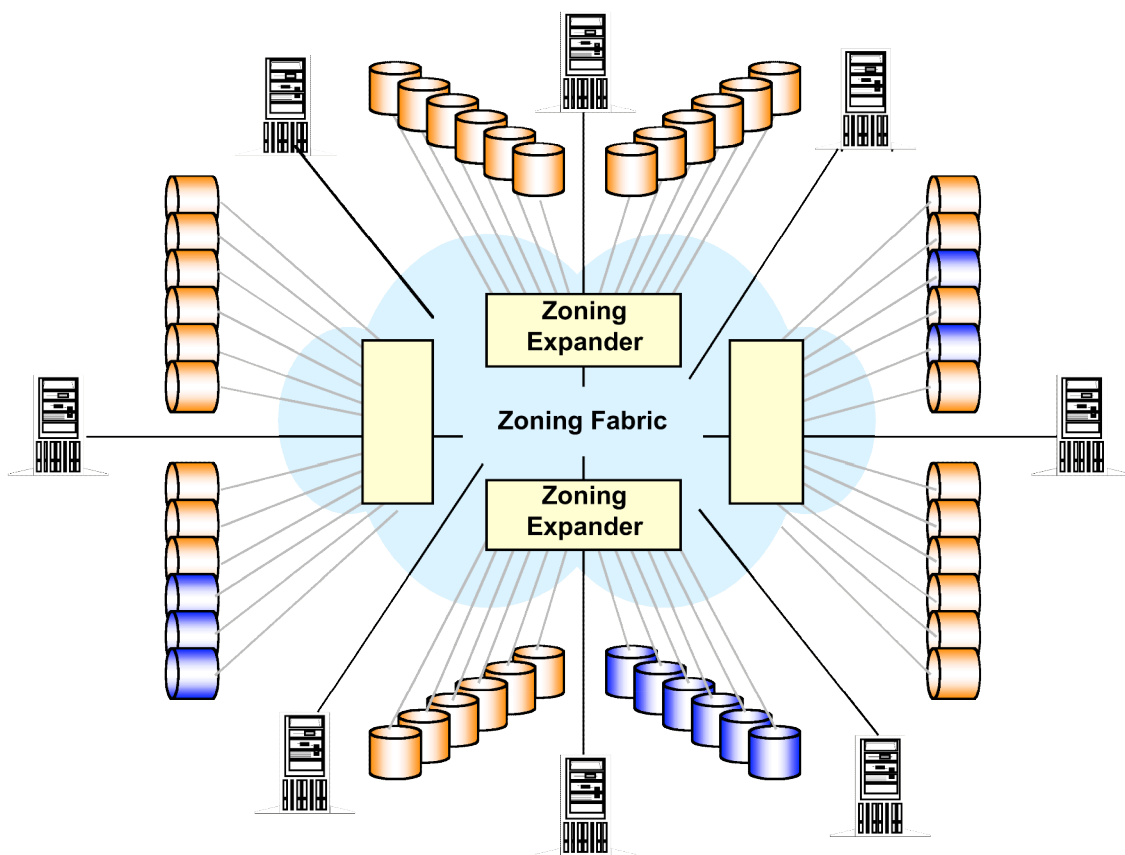
For details, see the SAS-2 Zoning Proposal (05-144r5) on the T10 Technical Committee website, www.t10.org.

2.1 Zoning Domain

Figure 2 shows an example of a SAS-2 zoning service delivery subsystem. The system looks identical to a legacy SAS 1.1 service delivery subsystem, except that it consists of zoning expanders. Zoning expanders are capable of configuring zones and implementing and enforcing access control policies on the network.

The end devices, which are connected to the expanders, can be legacy devices (including expanders). Any of the end devices as well as expanders that are attached to the edge of the zoning fabric will inherit the zone membership of their attachment point, at the boundary of the zoning service delivery subsystem.

Figure 2 Zoning Service Delivery Subsystem Example



Since the zoning function is fully implemented by the expander fabric, the implementation is completely transparent to end devices or legacy devices such as SAS 1.1 HBAs and HDDs that do not have knowledge of zoning, to change their behavior.

2.2 Zone Policy Configuration

2.2.1 Identifying Zones

For a SAS zoning expander to identify a device and its control policies, the system administrator must define the zones by using a unique ID: either the device’s worldwide name that is defined during its manufacture or the port number where a device is attached to the service delivery subsystem. This unique ID is assigned to the expander PHY that is attached to each end device.

Table 1 shows a comparison of worldwide name-based and port-based zoning. Note that each has its intended users. The zoning proposal supports both. The underlying zoning mechanism is based on the most secure method, port-based zoning, which is PHY-based. However, one-to-one mapping can be done between the WWN and the physical port ID using upper-layer management software to present a management API that is WWN-based.

Table 1 Comparison of Worldwide Name and Port Zoning Control Policies

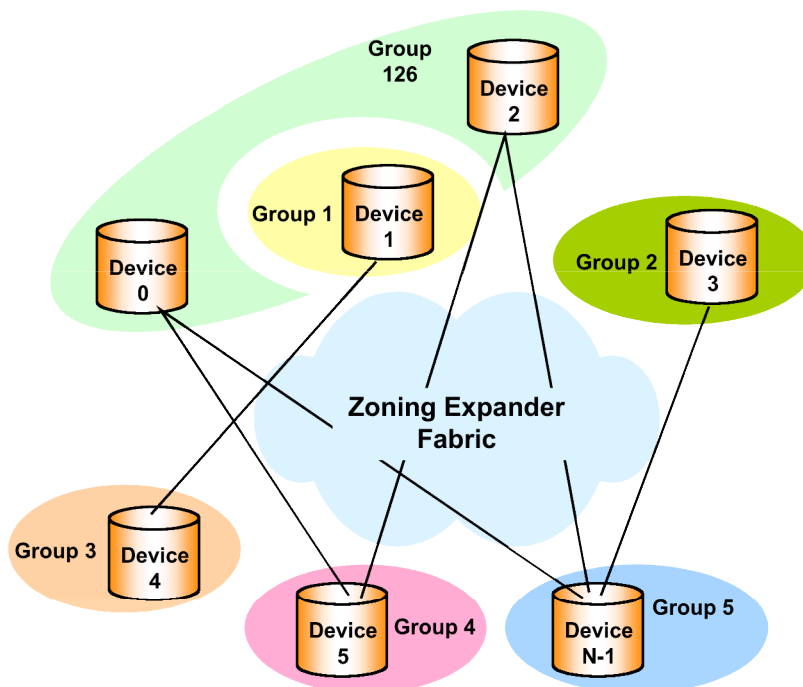
Worldwide Name Zoning	Port Zoning
Uses the unique worldwide name associated with each device to define the zones. A database stores all the WWNs and port numbers.	Uses the physical port IDs of the expanders to define the zones. Access rights are defined for each port in the zoning service subsystem.
Determines which WWNs are allowed to access other WWNs.	Determines which ports are allowed to access other ports.
Devices can be moved between ports without updating the zone configuration.	Zone information must be updated when a device is moved.
This method is susceptible to unauthorized access and spoofing of WWNs.	This is a trusted, hardware implementation. It prevents spoofing.
Easier to use.	More secure.

If a device is relocated on the network, the management software will be able to detect the move and automatically reassign the same policy to a new port.

2.2.2 Defining Policies and Permissions

Figure 3 shows an example zoning topology. In this instance, Device 0 and Device 2 are assigned to the same group (Group 126) since they share common resources (the need to communicate with devices in Group 5 and Group 4).

Figure 3 Zoning Groups Example



The access control policies, otherwise referred to as zone permissions, for each expander group are defined using the zone permission table, which is represented in Figure 4.

Figure 4 Zone Permission Table

		Groups	0	1	2	3	4	5	6	...	126	127
Special Group	{	0	0	0	0	0	0	0	0	...	0	1
			$P[0,0]$...	$P[0,127]$
User Defined Groups	{	1	0			1				...		1
		2	0					1		...		1
		3	0	1						...		1
		4	0							...	1	1
		5	0		1					...	1	1
	
Special Group	{	126	0				1	1		...		1
		127	1	1	1	1	1	1	1	...	1	$P[127,127]$

Group 0 and Group 127 are provided as special groups and cannot be defined by the user. By default, Group 0 cannot access any other groups except for those in Group 127. Group 0 is used, for example, for a new device that is inserted into the zoning service delivery subsystem, and has not yet assigned a group by the system administrator. Group 127 (SMP target) is allowed to access all other groups, including Group 0 and is used for topology discovery and zone management.

Groups 1 to 127 are configured by the system administrator based on the needs of the application. The permission values for each group are reversed for their mirror groups. For example, $P[x,y]$ is equal to $P[y,x]$ to accommodate SSP and STP exchanges, which are bi-directional. Devices in the same group do not always need to access each other and therefore, the permission value for these must be configured. For example, if two servers share the same group of target resources and share the same storage space, but must not access each other, users can assign them to the same group ($P[x,x]$), restrict access, assign all the shared resources to another group (y), and allow permission between $P[x,y]$.

When configuring the zoning permission table, as shown in Figure 4, the rules that must be followed are:

1. When $P[x,y]=0$, Groups x and y are not allowed to access each other. When $P[x,y]=1$, Groups x and y are allowed to access each other.
2. Group 0 is not allowed to access any other group except Group 127.
3. Group 127 is allowed to access all other groups.
4. Permissions are reversible: $P[x,y]$ is equal to $P[y,x]$.
5. Members in the same group $P[x,x]$ may not have permission to access each other.

2.2.3 Propagating Zone Permission Table Updates

Each time a topology change is detected in the SAS subsystem, the zone permission table updates are updated to the service delivery subsystem by the supervisor using the SMP command, CONFIGURE ZONE PERMISSION.

The permission table update is a multi-step process. The supervisor must send the entire permission table to the current supervising expander. The supervising expander propagates the changes to the rest of the expanders in the service delivery subsystem. The use of a supervising expander ensures consistency among the zoning policy tables in all zoning expanders in the service delivery subsystem.

As previously mentioned, the supervising expander is elected based on the largest SAS address in the topology. Each zoning expander has a SUPERVISING PRIORITY attribute, which determines whether an expander is a candidate to be a supervisor. If this attribute is set, when the zoning expander traverses the adjacent expanders in the SAS domain, it is counted with all the other supervisors. The supervisor with the highest election priority value (the largest expander SAS address value) is elected as the supervising expander.

2.3 PHY Zone Configuration

One of the supervisor expander’s responsibilities is to communicate the attributes of the expander PHY port to the service delivery subsystem (expanders). This is the second type of zone configuration that needs to be done. The attributes that need to be configured for the PHY zone include:

- Source Group ID —The GROUP ID attribute is the unique user-defined zone group ID (ranging from 1 to 127).
- Supervisor status —The SUPERVISOR attribute defines whether the device attached to the PHY can be a supervisor.
- Trusted status —The TRUSTED attribute defines if the device is considered trusted thereby defining the boundary of the service delivery subsystem. This is usually an expander PHY-to-expander PHY attribute.
- Source check status — This attribute defines whether the expander can automatically check the source address of any open address frames against the source address that was communicated during the initialization process. The sources check status ensures that a device cannot spoof (pretend to be) a trusted device to gain access to the service delivery subsystem and gain unauthorized data.

The PHY zone updates are atomic and are made from a supervisor to a zoning expander using a single SMP command, CONFIGURE PHY ZONE. The expander self-discovery process propagates the new PHY zone information associated with the attached SAS address. Even in a multi-supervising environment, no coordination is needed at the expander level.

2.4 Impact on Topology Discovery

The SAS 1.1 specification defines an expander route table, which maps the SAS address to an expander PHY. For SAS-2 zoning purposes, this route table is extended to also include the GROUP ID, SUPERVISOR, and TRUSTED bit attributes. During topology discovery, the zoning expanders propagate the group assignment along with the SAS address. Figure 5 is a representation of this table.

Figure 5 Zoning Expander Route Table

	PHY 0	PHY 1	PHY 2	...	PHY n
Expander 0				...	
Expander 1				...	
Expander 2				...	
...
Expander n				...	

Routed SAS Address, Group ID, Supervisor, and Trusted Bit Attributes

All zoning expanders are configured as self-configuring expanders. That is, all the expanders are able to traverse the SAS topology and populate the zoning expander route table. This process also protects the service delivery subsystem from any problems that might occur with the host configuring the zone table, where security is not ensured.

Host topology discovery still occurs, as defined by SAS 1.1, to determine the devices in the service delivery subsystem. The management of the zoning subsystem is done through the REPORT GENERAL, DISCOVER, REPORT ZONE ROUTE TABLE commands, which contain new fields for SAS zoning. However, information filtering is done based on the Group ID of the originating end device. Hence, if an end device is from a particular group, say Group x, the zoning expander will only record the devices attached to Group x even though it may be attached to other groups, limiting what the host can “see”.

2.5 Zoning Operation

2.5.1 OPEN Address Frame Handling

An OPEN address frame (OAF) is the means for establishing a connection between two SAS devices. For SAS zoning, in addition to carrying both the source SAS address and the destination SAS address, the OPEN frame carries two new fields necessary for routing in a SAS-zoned network:

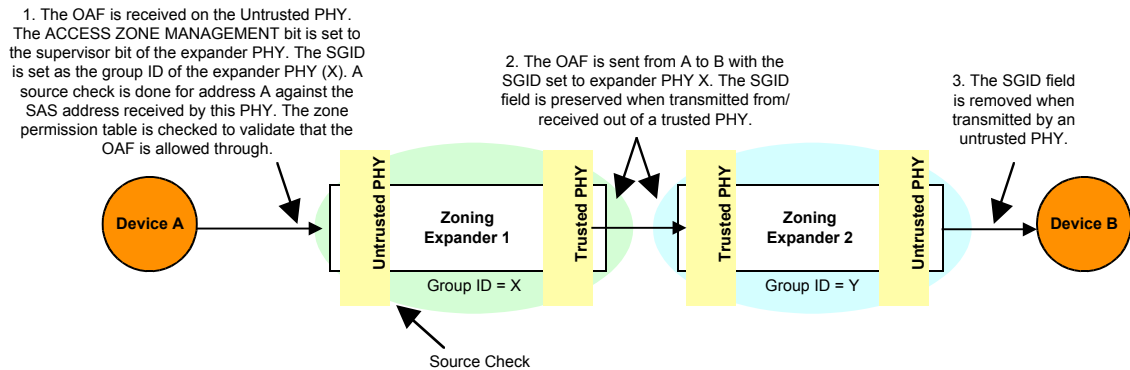
- SOURCE GROUP ID (SGID)
- ACCESS ZONE MANAGEMENT

New OAF Fields Defined

The SGID field is represented in the zoning permission table shown in Figure 4 by P[y]. The expander routes packets according to the destination address in the OPEN frame using one of the three standard SAS methods: table, direct, or subtractive routing. During the routing process, these routing methods check the destination SAS address and map the address to a destination group ID (DGID). (In Figure 4, this is P[x].) Hence, SGID identifies where a frame has come from and DGID identifies where a frame is being sent. By checking the group ID against the zoning permission table, the expander can verify whether the OPEN request is permitted.

The ACCESS ZONE MANAGEMENT field defines whether the OPEN address frame originates from a supervisor device or not. This is important for proper access control and permission checking.

Figure 6 Open Address Frame Handling



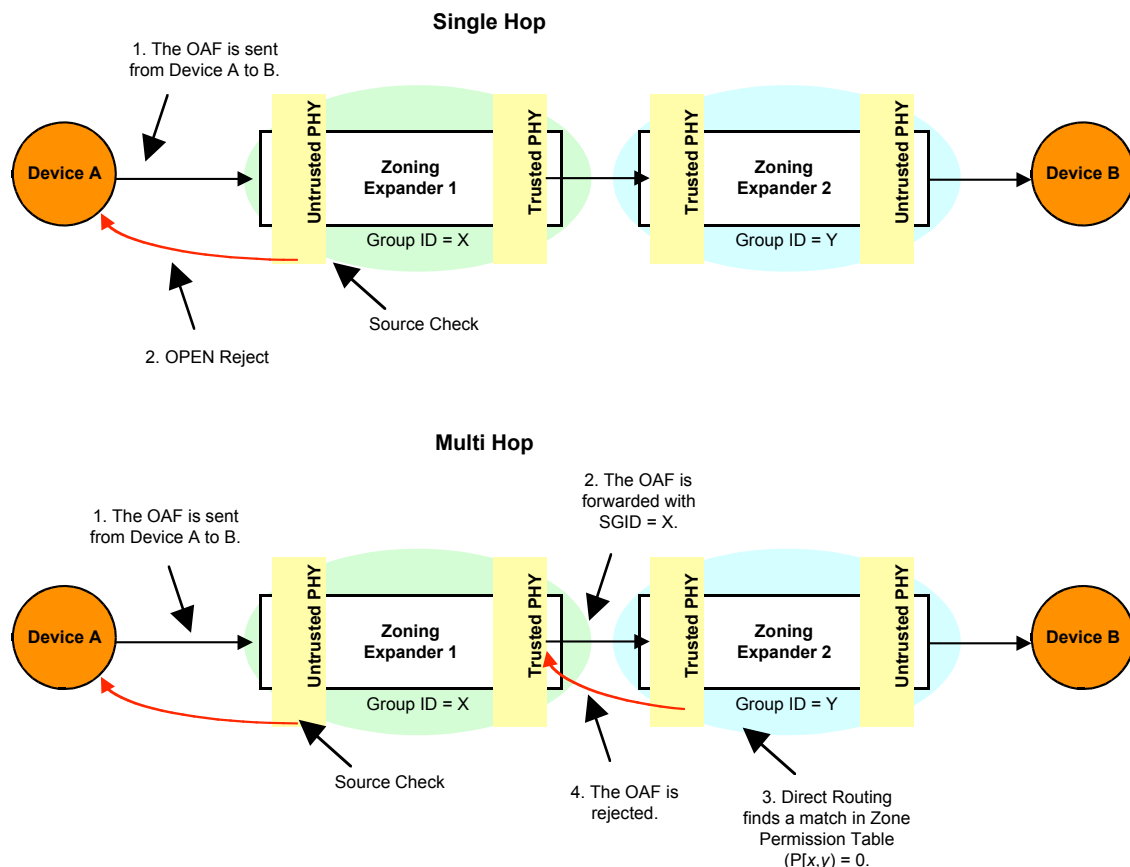
Both the SGID and ACCESS_ZONE_MANAGEMENT fields are only valid for OPEN address frames that are passed among devices inside the zoning service delivery subsystem on trusted expander PHYs. Since end devices can be legacy devices that do not have these fields, when an OPEN frame is received from an untrusted PHY, the ingress PHY of the source zoning expander will set the SGID field to the group ID associated with the PHY and the ACCESS_ZONE_MANAGEMENT field according to the supervisor attribute of the associated PHY. After the frame is routed, the egress PHY of the destination expander will remove these fields before transmitting on to another untrusted PHY. Refer to Figure 6 for details.

Permission Checking

Permission checking for SAS zoning is done either using a single-hop or multi-hop method:

- If the zoning expander topology uses only table routing, then the zoning route tables will contain all of the SAS addresses in the domain, that is, the routing table is flat across the topology. In this case, an illegal OPEN address frame will always be rejected by the first expander. The zone routing table immediately routes the destination address to the DGID that associated with it and accepts or rejects the OPEN frame based on the permission table.
- If subtractive routing is used in the topology, the zone routing table will only contain a subset of the SAS address in the domain. In this case, the OPEN address frame will be propagated from expander to expander via the subtractive port until it reaches an expander that has knowledge of the group ID and corresponding SAS address. In this multi-hop method, an illegal OPEN frame will be subtractively routed as far as the last expander attached to the destination address. Although this approach may consume more inter-expander link bandwidth, it does allow edge expanders to have a smaller route table.

Figure 7 Single-Hop and Multi-Hop Permission Checking



2.5.2 Broadcast Handling

In SAS 1.1, when the status of a device in the topology changes, a Broadcast primitive is generated and sent throughout the topology to inform the host that a change has occurred. The concept of Broadcast limiting restricts the Broadcast primitive to only propagate to the zones it has permission to access. To accomplish this in the SAS-2 zoning specification, the Broadcast frame includes the SGID. The Broadcast frame can only propagate through the service delivery subsystem to groups that the SGID is allowed to access, as defined by the access policy permission table. Broadcast limiting is important for reducing broadcast storms and topology discovery traffic.

Insertion of the SGID into the Broadcast frame is only used for inter-expander communication. If a broadcast message is transmitted from a zoning expander attached to an untrusted PHY (outside the service delivery subsystem), the PHY will remove the SGID before sending the Broadcast on to the untrusted PHY.

3 Applications for SAS Zoning

Serial Attached SCSI was designed to be a feature-rich replacement for parallel SCSI. As such, SAS was architected and specified around existing SCSI Direct Attached Storage (DAS) applications. As SAS is maturing, new applications outside of the DAS market are emerging. Zoning has been introduced to the Serial Attached SCSI standard to support new SAS applications such as SAS switches and SAS storage blades for blade server systems.

3.1 SAS Switches

Since 2004, blade server technology has begun to attract the attention of mainstream customers in the server market. The advantages of a blade server system include increased performance, density and the portfolio of available blade server system products including blades, multiple processor choices, multiple operating systems, new networking options, and new storage options. These advantages have encouraged mainstream customers to deploy blade servers into their data centers to support mission critical business processes.

To continue providing end customers with more options and drive the growth of the blade server market, blade server OEMs are planning to introduce SAS into their blade server designs. By moving to SAS, blade server architectures can support both enterprise class applications with SAS HDDs and provide low-cost storage options with SATA HDDs.

A *blade server* is a modular computing platform consisting of the following elements: blades, switch modules, power supply, management module and fans. A *blade* is a single-board server that contains one to four processors, memory, local disk storage, an on-blade network interface card (NIC), and SAN connectivity, as shown in Figure 8. A *blade chassis* may hold one or more blade cards, one or more Ethernet or SAN switch modules, one to four power supplies, one to two shared management modules and cooling resources. Chassis components communicate across a fully redundant midplane, enabling hot swap functionality of the chassis components and easy serviceability.

Figure 8 Block Diagram of Typical Blade Server Components

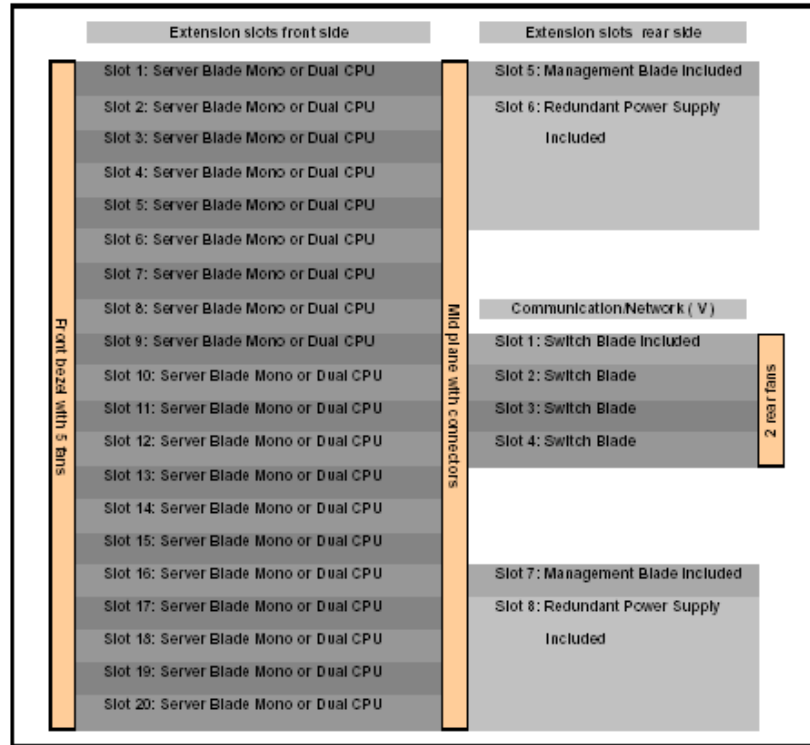
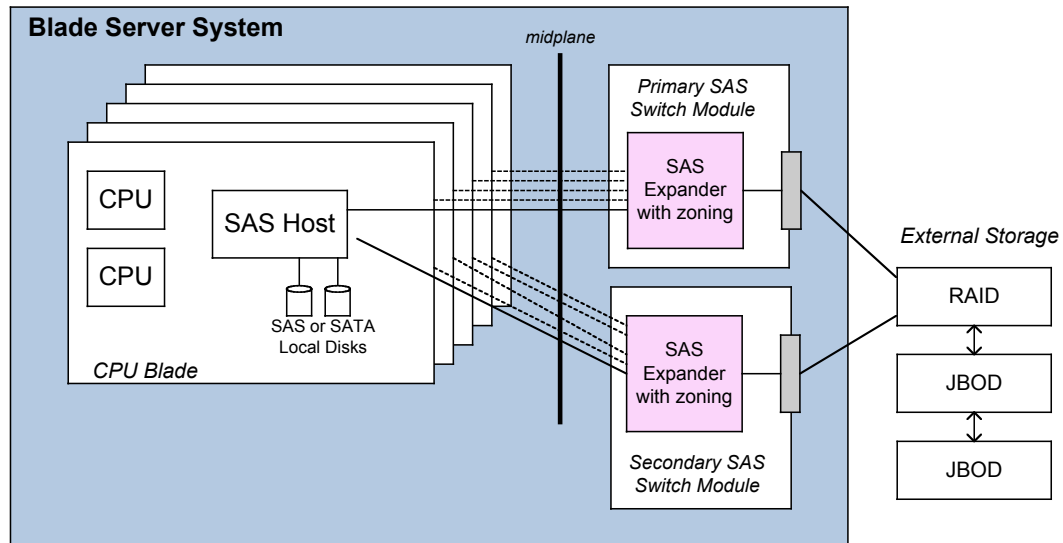


Diagram Courtesy Fujitsu Siemens Computers

The addition of SAS to blade server architecture will occur on the disk-interconnect side of the server. To enable the CPU blades to connect to either SAS or SATA storage, the switch modules or switch blades will require SAS switches rather than Ethernet or Fibre Channel.

The Fibre Channel and Ethernet switches used in blade server applications today provide zoning functionality to ensure that the server blades cannot “see” other servers in the blade architecture. Thus, SAS switches must provide zoning capabilities. The SAS switch design is based on using SAS expanders, as shown in Figure 9.

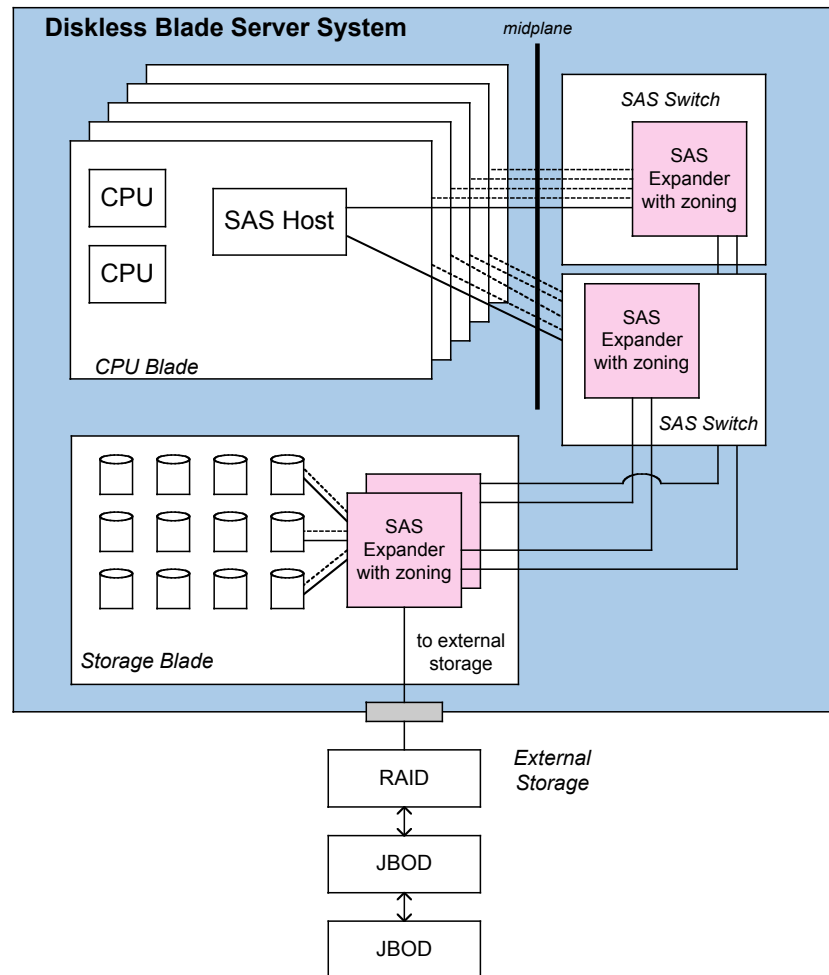
Figure 9 Example SAS Blade Server Architecture



3.2 Storage Blade

A new type of architecture emerging in the blade server market is the concept of disk-less server blades. The hard disk-drives (HDDs) are relocated off of the server/CPU blade and on to a shared location called a *storage blade*, as shown in Figure 10. The storage blade can be thought of as “semi-local” storage that is shared by a small number of blades. The storage blade concept is not truly direct attached storage (DAS) because it is not exclusively owned by a single server and it is not completely a storage area network (SAN) because the storage is not available to all servers on the fabric. The storage blade is a new intermediate level in the storage hierarchy and provides an advantage in price and performance by eliminating the requirement to pre-allocate unused storage and allows the storage to be amortized over multiple blades.

Figure 10 Diskless Blade Server Architecture



The storage blade provides a pool of generic assignable drives to the blade server. A global resource manager can then assign and manage desired boot images, applications or select links to appropriate data volumes. The storage blade provides low latency access to storage for the server blades.

Zoning is key for disk-less blade architectures, because each server blade's protected boot drive will be located on a shared storage blade, as in Figure 10. The zoning will be required to protect each CPU blade and associated storage from unauthorized access/security breaches.

4 Conclusion

SAS zoning provides the traffic segregation, resource flexibility, controlled resource sharing, protection, and topology control functionality required to manage SAS-based systems including blade servers.

References

1. ISO/IEC 14776-151:200x, ANSI INCITS. ***.200x, Project T10/1601-D, Working Draft American National Standard, Information technology – Serial Attached SCSI – 1.1 (SAS-1.1). July 9, 2003. <http://www.t10.org/>.
2. ISO/IEC 14776-150:200x, ANSI INCITS.***.200x, Project 10/1562-D, Working Draft American National Standard, Working Draft American National Standard, Information technology– Serial Attached SCSI (SAS). July 24, 2005. <http://www.t10.org/>.
3. T10/05-144r6, SAS-2 zoning, presented to the T10 Technical Committee. June 13, 2005. <http://www.t10.org/>.

Glossary

Blade	A single-board server that contains one to four processors, memory, local disk storage, an on-blade network interface card (NIC), and SAN connectivity.
Blade server	A modular computing platform consisting of the following elements: blades, switch modules, power supply, management module and fans.
Blade chassis	A chassis that holds blade cards, Ethernet or SAN switch modules, power supplies, shared management modules and cooling resources.
Broadcast message	A message used to notify all SAS ports in a domain of an event.
Expander zoning route table	A structure that provides an association between the destination SAS address and the expander PHY identifier.
Permission table	A structure that provides an association between the unique source and destination group identifiers and that is used to define permission between two groups.
PHY device	A device object that is used to interface to other devices.
PHY zone	The zone attached to a PHY. Three fields are required to be set for proper zoning operation: trusted, group ID, and supervisor. PHY zone updates are atomic.
SAS address	A worldwide unique name assigned to a SAS port, or expander, initiator, or target device.
Service delivery subsystem	A complete or partial SAS domain that provides the services of a normal SAS fabric, including zone management and access control services.
Storage blade	“Semi-local” storage that is shared by a small number of blades.
Subtractive routing	The method the expander uses to route connection requests not resolved using the direct routing method or table routing method to an expander.
Supervising expander	A SAS zoning expander that is automatically designated by an election process based on the highest election priority of all the zoning expanders within a SAS domain. The supervising expander propagates zone permission table updates to all zoning expanders when the zoning supervisors update the zone permission table.
Supervisor	A management entity that is either an expander device inside or an end device attached to the zoning service delivery subsystem. A zoning supervisor is an SMP initiator capable of generating SMP commands for configuration and management.
Topology discovery	The algorithm used by the management application client to configure the SAS domain.
Traffic segregation	The process of separating traffic between hosts and resources, and providing common access privileges to all end devices.
Trusted PHY	A zoning expander attached to a device that can be trusted to generate and receive controlled zoning information such as an OPEN address frame or BROADCAST address frame with zoning details that the service delivery subsystem requires for management and access control. A trusted device can be a zoning expander or a secured end device.
Untrusted PHY	A zoning expander PHY attached to a device that cannot be trusted to generate and receive controlled zoning information. An untrusted PHY is considered to be outside the zoning domain.
Zoning expander	A device that is part of the service delivery subsystem and that facilitates communication between SAS devices. A zoning expander is capable of the zoning function.