SAS Zoning and Access Control (05-186r1)

May 11, 2005

Heng Liao, Steve Gorshe
PMC-Sierra, Inc.
liaoheng@pmc-sierra.com
Steve_gorshe@pmc-sierra.com
604-415-6000x2542
503-431-7440

Thomas Grieff
Hewlett Packard
Thomas.grieff@hp.com
281-514-5581

Outline

- Address issues raised at May 2005 T10 SAS PROT WG
  - Persistent device group re-assignment
  - Synchronized update of Permission Table among multiple supervisors
  - Impact of PHY ZONE configuration and permission table update on current data-path traffic
  - Per Zone Reporting of Expander CHANGE Count
- Proposed solutions
- Appendix: Changes made to the original slidedeck(05-186-r0)

Group Assignment Methods

- Two approaches have been discussed during May T10 WG regarding the group assignment
  - Approach 1: Phy based group assignment (physical security)
    - This approach is more secure as it puts group membership assignment under total control of the Zoning Expanders without relying on the WWN of end device.
  - Approach 2: WWN based group assignment (ease of use)
    - This approach assumes the group assignment is persistent based on WWN. This approach may be easier to use assuming the WWN of the end device can be trusted.
  - Demands exist for approach 1 and approach 2 in different applications.

- How are approach 1 and 2 different?
  - For initial group assignment, approach 1 and approach 2 are equivalent:
    - WWN can be easily mapped to Expander Phy by a Supervisor
  - Difference is in how to handle group reassignment: when a device is removed from the topology, and later added back to the topology, approach 1 does not preserve the old group assignment, while approach 2 preserves persistent group assignment based on WWN.

- Current Zoning proposal (05-144r1) defines PHY ZONE configuration information using approach 1.

- Solution:
  - Provide base mechanism that supports the physical security of approach 1
  - Add Expander and Supervisor functions to handle group reassignment desired by approach 2.

Solution:

Re-assignment of PHY ZONE configuration based on SAS Address

- How to Handle group re-assignment under the following conditions:
  - 1) If a new device is added into an Expander Phy:
    - **Expander Function**: PHY defaults to Group 0, prevents this device from accessing any group except 127 until the Supervisor assigns it to a non-default group
  - 2) If a device is re-attached to the same Expander Phys (or link go down and up between device and expander):
    - **Expander Function**: The expander shall detect the same device has shown up and restore the old group assignment
  - 3) If a device is moved from one place on the fabric to a different place
    - Step 1.1 (**Expander Function**): The device is assigned to group 0 waiting for the Supervisor to re-assign
    - Step 1.2 (**Supervisor Function**): The supervisor detects the WWN of the device that has been moved, and it may decide to do the following:
      - 1.2.1) In case of device redeployment: assign this device to a different group
      - 1.2.2) In case of convenience move: reassign this device to the old group.
      - This function could be implemented as OEM specific value-add features (outside the scope of SAS 2 specification)

- Any physical re-arrangement of things in the zone fabric will cause the need for supervisor involvement.
  - Because the expanders do not have the knowledge of the intention of the physical rearrangement of device, it is appropriate to leave the decision of group re-assignment to the Supervisor. It is dangerous to assume that the expander can determine the intent of the move.

Synchronized Updates Among Multiple Supervisors

- Is there an issue if two Supervisors both want to change the PHY ZONE assignment?
  - No issue – each CONFIGURE PHY ZONE command is an atomic operation sent to one expander only. The self configuring expander topology discovery process handles the propagation of zone route table that can handle simultaneous changes in the fabric.
- Is there an issue if two Supervisors both send ZONE PERMISSION table updates to different expanders in the fabric at the same time?
  - Yes, the current proposal (05-144r1) can detect a conflict between two supervisors based on the proposed GENERATION CODE mechanism. But the proposal does not provide a clean solution to ensure the consistent PERMISSION tables across expanders.
  - How to ensure PERMISSION table updates are consistent across supervisors?

Solution:
Supervising Expander Election (Proxy)

- Elect a single Supervising Expander in a SAS domain based on largest Expander SAS address value as the proxy for zone permission update propagation
  - Each self-configuring expander can detect who is is the supervising expander implicitly through topology discovery process (No need to invent new mechanism for election process), the supervising expander WWN is reported by SMP REPORT GENERAL command
  - CONFIGURE ZONE PERMISSION command must be sent to the supervising expander, and the supervising expander is responsible for propagating the PERMISSION table update to other expanders
  - CONFIGURE ZONE PERMISSION command is rejected if sent to a non-supervising expander.
  - If a supervising expander is in the process of executing/propagating a PERMISSION table update, any new CONFIGURE ZONE PERMISSION commands from any supervisor is rejected
- Atomic operation of permission update is guaranteed across multiple supervisors by election of supervising expander
  - Simplicity – no need to introduce complicated “global lock/unlock” mechanism
  - No single point of failure – if topology changes, new supervising expander will be automatically elected.
  - Fully distributed – topology discovery process in expander/host automatically accomplishes the election without additional effort
  - Can support unlimited number of supervisors simultaneously
  - This schemes relies on the reliable operations of the SMP management entity of the elected supervising expander – same requirement of all expander SMP management entity being healthy as in SAS 1.1

Result of zone change proxy

- The SMP zone permission changes are made atomic by having only one expander making zone changes.
- Requires that the SMP function healthy, but this is not a new requirement for domain health.
- The supervising expander is required to be persistent during one permission update period.

- The Supervisor must re-download the permission table using the latest Supervising Expander address after SAS domain topology change involving one or more expander (adding or losing). This handles corner cases:
  - A Supervising expander dies before it completes the permission update propagation period
  - A New Supervising expander is added during a update period
  - A Non-Supervising expander is added during a update period
  - A Non-Supervising expander is removed during a update period
  - Any expander added to the domain not during the update period
- A Permission redownload from the supervisor puts all expanders into a clean and consistent state after topology changes that affects expanders.

Impact of PHY ZONE Updates on Data Traffic

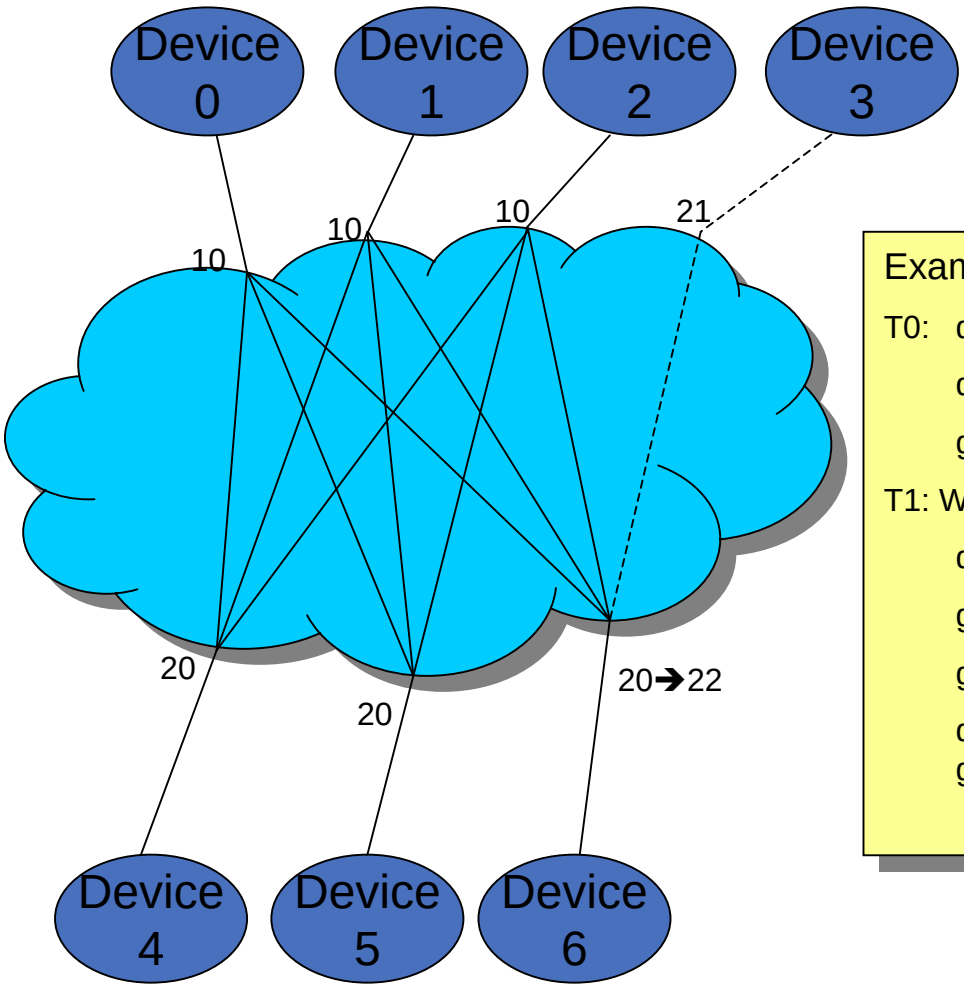
- PHY ZONE configuration update – reassign a device from one group to another
  - Atomic operation to a single expander. No need to stop traffic, or open arbitration
  - The device simply disappears from the current group (no different from phy lost link).
  - Any open to this address will be rejected by the expander (no destination).
  - The device shows up in the new group when the PHY ZONE update is completed and the topology rediscovery is completed.

Impact of ZONE PERMISSION Table Update on Data Traffic

- ZONE PERMISSION table update
  - Update need to propagate to multiple expanders, but no need to stop traffic, or open arbitration
  - If a P(X,Y) is changed at different times at different expanders
    - 0 -> 1: if not all expanders have finished the update. The host should not be able to discover the addresses in group Y (broadcast not generated until propagation complete). So there should be no traffic between X and Y until the update is completed. Even if an OPEN does come from X to Y before an expander has finished the update, the OPEN is rejected, from the host perspective, this only means the new permission has not taken effect until the propagation is completed.
    - 1 ->0: There devices in group X, Y may still attempt to OPEN new connections to each other, an OPEN maybe rejected or not rejected (depending on whether the expanders along the pathway is completed updates), when the update propagation is completed, all open between X and Y will be rejected. Again, from host point of view, the new permission change has not taken effect until the update propagation is completed.
  - The batch update of the permission table takes multiple commands to download the full table. This means there could be a period time the permission is not symmetrical. During this period of time, this may cause new OPEN to be routable from X to Y, but not routable from Y to X. But again, if you look at any group pair X and Y, the analysis above is applicable.
    - The worst that can happen during this period of time is X can send command to Y, but Y can not open a connection back to X (rejected). This means the new permission take some time to fully take effect.
- Conclusion:
  - The expanders do not need to stop current connections, nor stop routing new OPEN during zone permission table update. All ECM function can go on as normal.
  - The permission table update may not fully take effect until the table is propagated to all expanders, but this is what is expected at the system level, and end devices error recovery should be able to handle this any ways.

Example Non-disruptive Zone Change (split existing group)

In this example a new device 3 is added that requires access to device 6. Device 6 is currently in a group with 4 and 5. Since the access privileges are now, not symmetric, group 20 must be split. This can be done without effecting current operation or data flows.



Example split existing group

T0: d0,d1,d2 = g10 – devices 0,1,and 2 are in group 10

d4,d5,d6 = g20 – devices 4,5,and 6 are in group 20

g10←→g20 – 0,1,2 can talk to 4,5,6

T1: Want to add device 3 to talk to device 6 must split group 20

d3=g21 – add new device into new group

g10←→g20,g22 – 0,1,2 can talk to 4,5,6

g21←→g22 – pre-provision permission for new group 22

d6=g22 – takes device 6 out of group 20 and put it into new group22

CHANGE Counters

- PHY CHANGE COUNTER
  - Do not need to be extended for zoning - each PHY belongs a unique group ID
  - Response to DISCOVER command (PHY_CHANGE_COUNT field) based on based on the ZONE PERMISSION TABLE:
    - Report the PHY change counter value, If the SGID of the SMP command is allowed to access this PHY
    - Report change counter value of “0”, if the SGID of the SMP command is not allowed to access this PHY
- EXPANDER CHANGE COUNTER
  - Expander needs to report different change count value based on the SGID of SMP command (only count the changes that affects the group the SMP command has come from)
    - Physically do not need per Group Expander counter
    - The Expander can go through the PHY CHANGE COUNTER one by one, and sum up the PHY CHANGE COUNTER for the PHYs the SMP SGID is allowed to access based on the ZONE PERMISSION table.

Modification to 05-144r1

- Add description on group ID reassignment expander function and supervisor functions
- Add description on supervising expander, election process, and PERMISSION table update process
  - Change CONFIGURE ZONE PERMISSION command format to support this
- Add description on the impact on the PHY ZONE and ZONE PERMISSION TABLE updates on existing expander connections and how expander routes new OPENs.
- Add description on how expanders should report the PHY CHANGE COUNT and EXPANDER CHANGE COUNT based on SGID
- Editorial changes the SMP commands as showed in appendix.

Appendix:

SAS Zoning and Access Control (05-186r0) with markups

April 27, 2005

Heng Liao, Steve Gorshe
PMC-Sierra, Inc.
liaoheng@pmc-sierra.com
Steve_gorshe@pmc-sierra.com
604-415-6000x2542
503-431-7440

Thomas Grieff
Hewlett Packard
Thomas.grieff@hp.com
281-514-5581

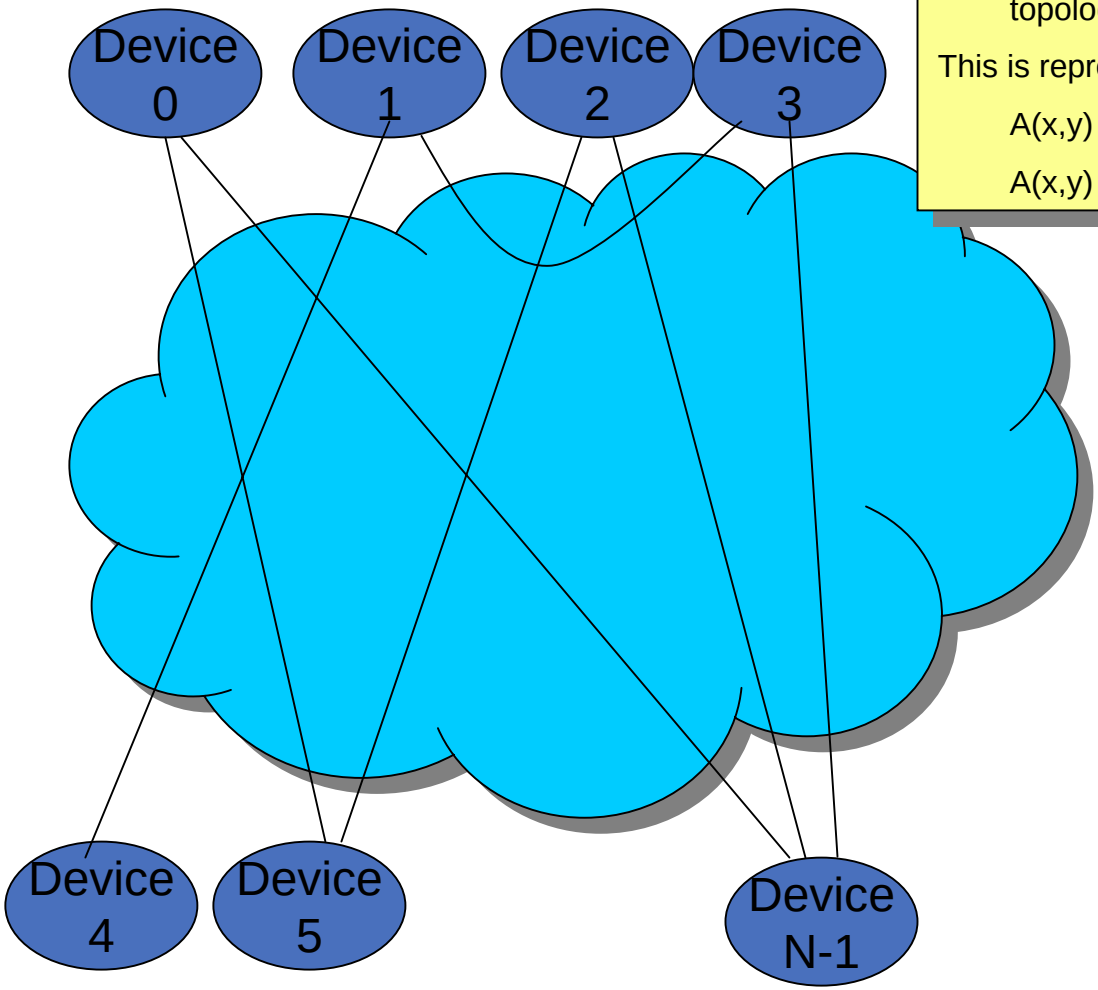
Purpose of Access Control in SAS

- Use SAS topology as a fabric interconnecting multiple hosts to multiple targets
  - Traffic Segregation – similar to FC zoning, or Ethernet VLAN
  - Access Control - allow supervisor to restrict which host can see which target
  - Device sharing – hosts can share targets without seeing each other
  - Identity spoofing prevention – the policy must be created in the SAS fabric without relying on the end devices being honest about who they are
- Zoning transparency
  - Zoning function is implemented by the expanders in the fabric without changing the behavior of end devices
  - Legacy end device can be totally unaware of zoning, see a subset of the SAS domain (restricted by the zoning policy)
  - Legacy expander can be attached to the edge of the zoning fabric

Definitions

- In a SAS physical topology, the devices can be partitioned into two categories:
  - End devices: host/target
  - Expander devices: the interconnected expanders form a SAS fabric whose function is to support connections among end devices
  - Zoning fabric: a topology formed by one or multiple zoning expanders
- Mechanism for zoning:
  - Device Groups: Allow user to partition the End Devices into groups (with common access privileges). Groups are identified by a unique group ID assigned to the expander PHY attached to the end device. This spec supports up to 128 groups.
  - ZONE PERMISSION TABLE: flat table defines the access privilege:
    - P (X,Y) :
      - 1: devices in group X is allowed to communicate to devices in group Y
      - 0: devices in group X is not allowed to communicate to devices in group Y
    - Note that: Permissions are reversible (to allow SSP/STP exchanges).
      - P (X,Y) == Permission (Y,X)
    - Members in the same group may not necessarily get the permission to connect to other members of the group
      - P (X,X) =1: member of Group X can see other members in group X
      - P (X,X) =0: member of Group X can not see other members in group X

GROUP ASSIGNMENT
User's view



1. User wants to control/restrict the access permission among all devices, without worrying about the fabric topology looks like.

This is represented by the device permission matrix NxN :

A(x,y) = 1: device X can see device Y

A(x,y) =0: device X can see device Y

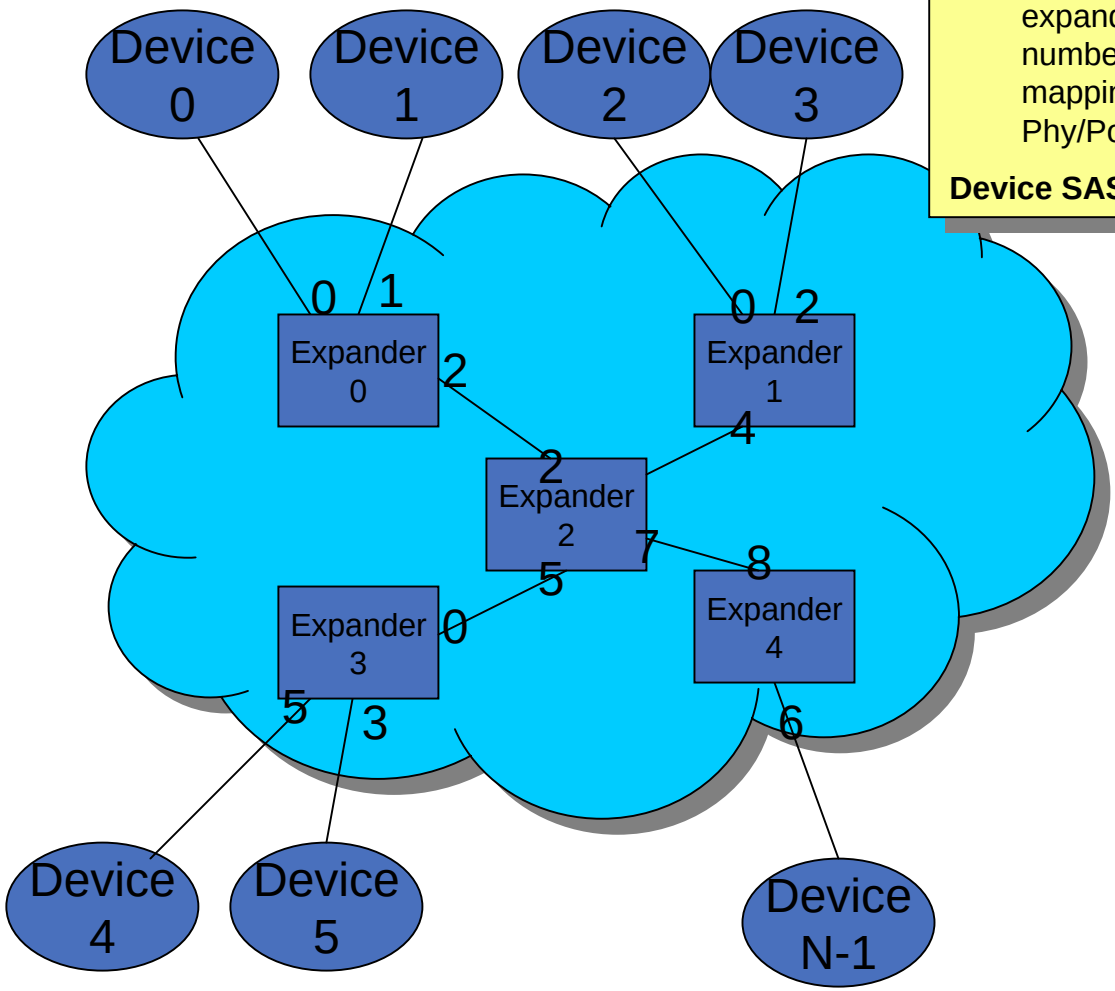
Device Permission Matrix

	0	1	2	3	4	5		N-1
0						1		1
1				1	1			
2						1		1
3		1						1
4		1						
5	1		1					
N-1	1		1	1				

GROUP ASSIGNMENT
Physical topology

1. Servers and Storage devices are all treated the same – as end devices attached to the fabric
2. Each end device is physically attached to one expander (as identified by expander number and phy number). The zone route table provides one to one mapping between device address to the expander and Phy/Port #

Device SAS Address == < Expander Address, Phy #>

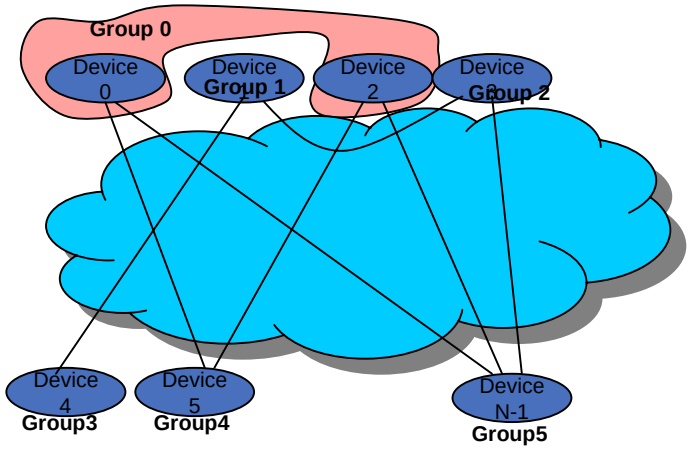


Topology Table

Device 0	Expander 0	Phy 0
Device 1	Expander 0	Phy 1
Device 2	Expander 1	Phy 0
Device 3	Expander 1	Phy 2
Device 4	Expander 3	Phy 5
Device N-1	Expander 4	Phy 6

GROUP ASSIGNMENT

Device grouping



1. Group all the devices with common device access permission together. The purpose is to have an efficient representation of the device permission matrix
2. The rows with common values are grouped together.

   e.g. Row 0 and Row 2 are the same, hence device 0 and device 2 are group together

Group assignment and Group permission table can be computed automatically with a very simple and fast process.

Device Permission Matrix

	0	1	2	3	4	5		N-1
0						1		1
1				1	1			
2						1		1
3		1						1
4		1						
5	1		1					
N-1	1		1	1				

Group assignment

Device 0	Group 0
Device 1	Group 1
Device 2	Group 0
Device 3	Group 2
Device 4	Group 3
Device 5	Group 4
Device N-1	Group 5

ZONE PERMISSION table

	0	1	2	3	4	5
0					1	1
1			1	1		
2		1				1
3		1				
4	1					
5	1		1			

PHY ZONE configuration

- Each zoning expander PHY is associated with the PHY ZONE configuration information:

Table 1. Per-expander phy zoning configuration

Name	Description
Trusted	If set to 0, this phy is on the boundary of the zoning fabric. All message (primitives and frames) that come across this phy shall be mapped to be backwards-compatible to SAS standard without zoning features, except for the new SMP commands defined by the zoning extension. If set to 1, this phy is inside the fabric boundary. The new primitives and frame formats that are defined by the zoning extension are allowed to pass through this phy.
Group ID[6:0]	The GID defines the zoning Group ID in the range from 0..127. GID=0: Group 0 is a special group that is not allowed to communicate with any other group except for group 127. Note that a device belonging to group 0 can still discover all the expanders and communicate with the SMP virtual target in the expanders (i.e. SMP virtual target within the zoning expanders are considered to have GID=127). GID=127: Group 127 is a special group that is allowed to communicate with all other groups. All trusted phys shall be automatically assigned to have GID =127 by the zoning expanders. GID=1..126: User defined groups. The communications amongst the user defined groups are restricted by the zoning permission table.
Supervisor	If Set to 1, the device attached to this phy is allowed to originate SMP commands to set up and change zoning configuration. If set to 0, the device attached to this phy is not allowed to originate SMP commands to change the zoning information.
SOURCE CHECK	The SOURCE CHECK field specifies whether the specified phy shall check the SOURCE SAS address against the SAS address in the IDENTIFY address frame received on the specific phy.

ZONE PERMISSION TABLE

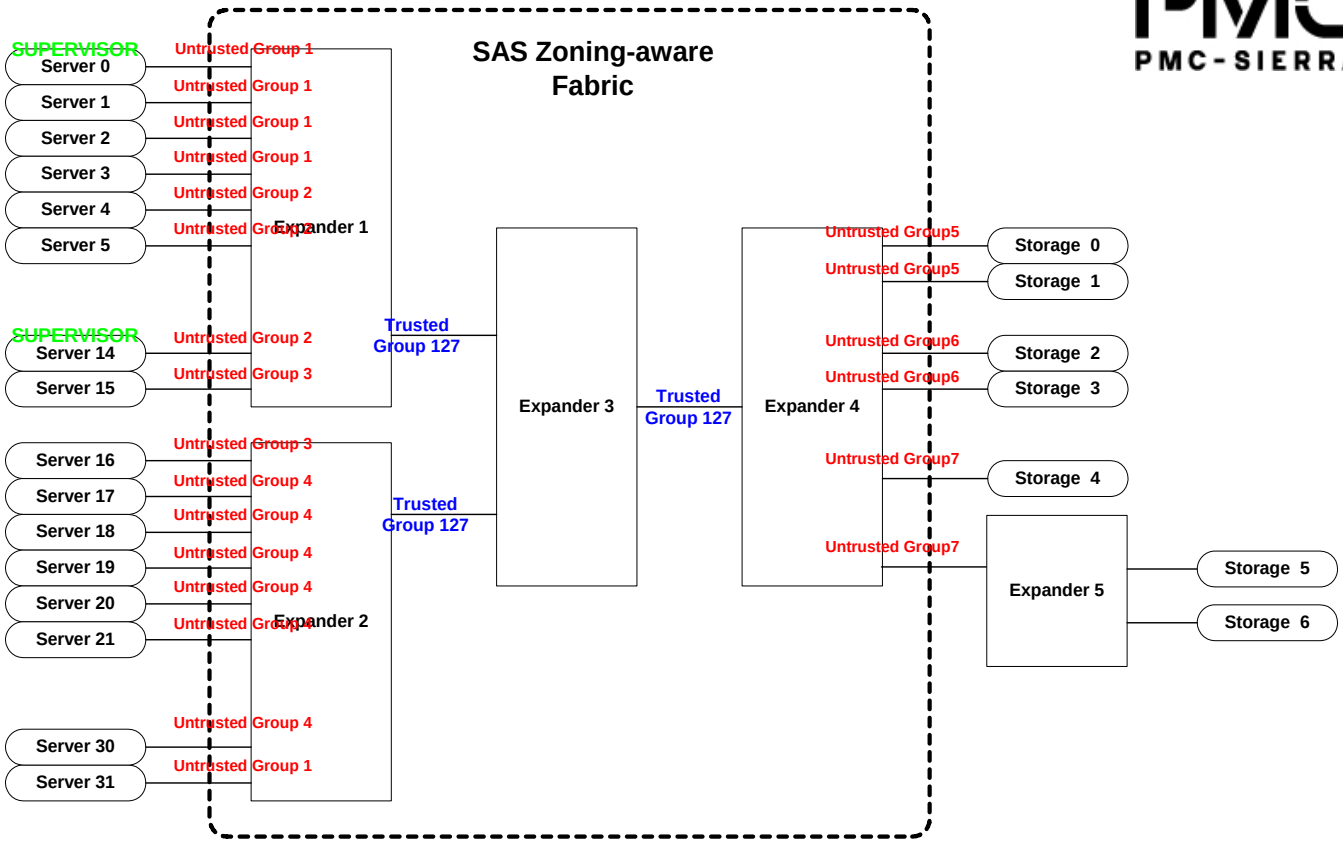
- The ZONE permission table is common across all zoning expanders that defines the access control policy among the groups

Table 2. Zoning permission table

	0	**1**	**2**	**3**	**4**	**5**	**…**	**126**	**127**
0	0	0	0	0	0	0	…	0	1
1	0	P[1,1]	P[1,2]	P[1,3]	…	…	…	P[1,126]	1
2	0	P[2,1]							1
3	0								1
4	0								1
5	0								1
…	…								…
126	0	P[126,1]					…	P[126,126]	1
127	1	1	1	1	1	1	…	1	1

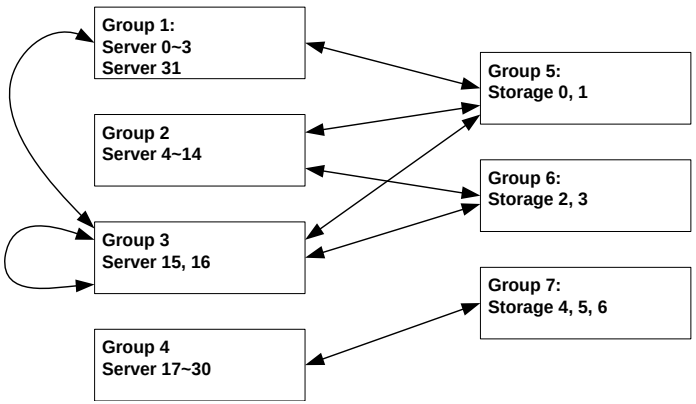
- P[X,Y] =1: means group X has permission to access Group Y
- P[X,Y] =0: means group X has no permission to access group Y
- Special groups:
  - Group 127 is allowed to access all other groups. Therefore, P[0..127, 127] is always set to all 1s, and P[127, 0..127] is also set to all 1s.
  - Group 0 is not allowed to access any other group except for 127. P[0, 0…126] are always set to all zeros. P[0, 127] is always set to 1. P[0..126, 0] are always zero

Example



Example
ZONE PERMISSION TABLE:

	1	2	3	4	5	6	7
1			X		X		
2					X	X	
3	X		X		X	X	
4							X
5	X	X	X				
6		X	X				
7				X			



OPEN Address Frame

— OPEN address frame format

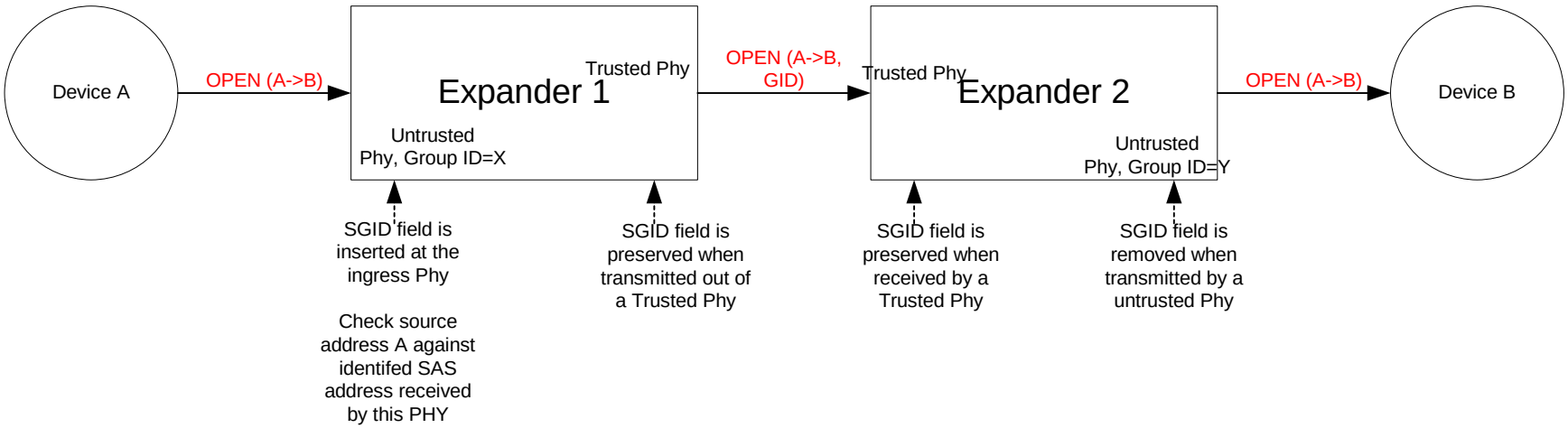
Byte\Bit	7	6	5	4	3	2	1	0
0	INITIATOR PORT	PROTOCOL			ADDRESS FRAME TYPE (1h)			
1	FEATURES				CONNECTION RATE			
2	(MSB)	INITIATOR CONNECTION TAG						
3								(LSB)
4		DESTINATION SAS ADDRESS						
11								
12		SOURCE SAS ADDRESS						
19								
20	ACCESS ZONE MANAGE MENT	SOURCE GROUP ID (SGID)						
21	PATHWAY BLOCK COUNT							
22	(MSB)	ARBITRATION WAIT TIME						
23								(LSB)
24		MORE COMPATIBLE FEATURES						
27								
28	(MSB)	CRC						
31								LSB)

- The ACCESS ZONE MANAGEMENT bit defines whether the OPEN address frame is originated from a supervisor device.
- The SOURCE GROUP ID field defines which source group the OPEN is coming from.

OPEN Frame Handling and Source Address Checking

- When an OAF is received on an
  - Untrusted PHY: Insert ACCESS ZONE MANAGEMENT and SOURCE GROUP ID into the frame based on the PHY ZONE configuration of the ingress PHY
  - Trusted PHY: preserve the ACCESS ZONE MANAGEMENT and SOURCE GROUP ID values
- Source Address Check
  - If the PHY ZONE configuration of the ingress PHY has SOURCE CHECK bit set, the SOURCE SAS ADDRESS in the OAF is checked against the SAS ADDRESS in the IDENTIFY frame received by this PHY
  - This prevents an end point from spoofing source address

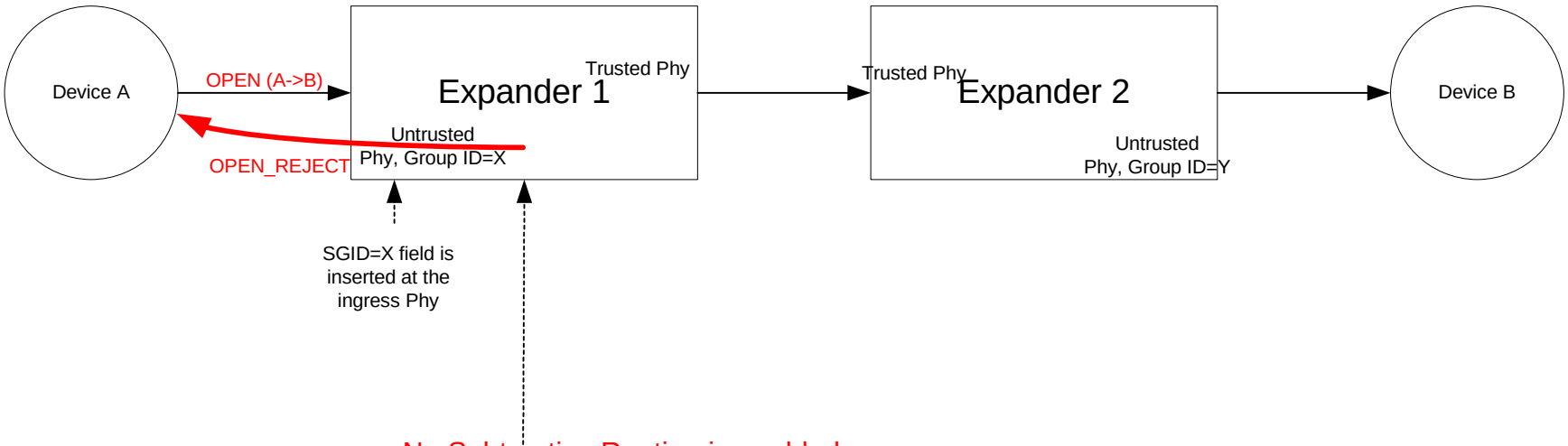
OPEN Handling Example



ZONE Permission check

- The expander ECM routes the packet according to the Destination SAS address in OAF using table, direct or subtractive routing
  - The Destination SAS address is mapped to a target group ID (DGID)
  - The expander checks the PERMISSION TABLE
    - If P(SGID, DGID) =1, the OPEN is allowed to get through
    - If P(SGID, DGID) =0, the OPEN is rejected as ZONE PERMISSION violation
- This flexible scheme supports ZONE permission checks in two ways (implementation specific):
  - Single Hop: if the ZONE ROUTE TABLE contains all SAS address in the domain (flat table), the illegal OPEN requests are rejected at the first expander.
  - Multi Hop: if the ZONE ROUTE TABLE only contains a subset of the SAS address (ex. Subtractive routing used), the illegal OPEN may be routed as far as the destination, the illegal OPEN request shall be rejected by the first expander that has knowledge of the GROUP ID corresponding to the destination SAS address.

ZONE Permission Check Example Single Hop (No Subtractive Routing)



No Subtractive Routing is enabled.
ECM looks up in both direct and zone routing table based on destination address B:

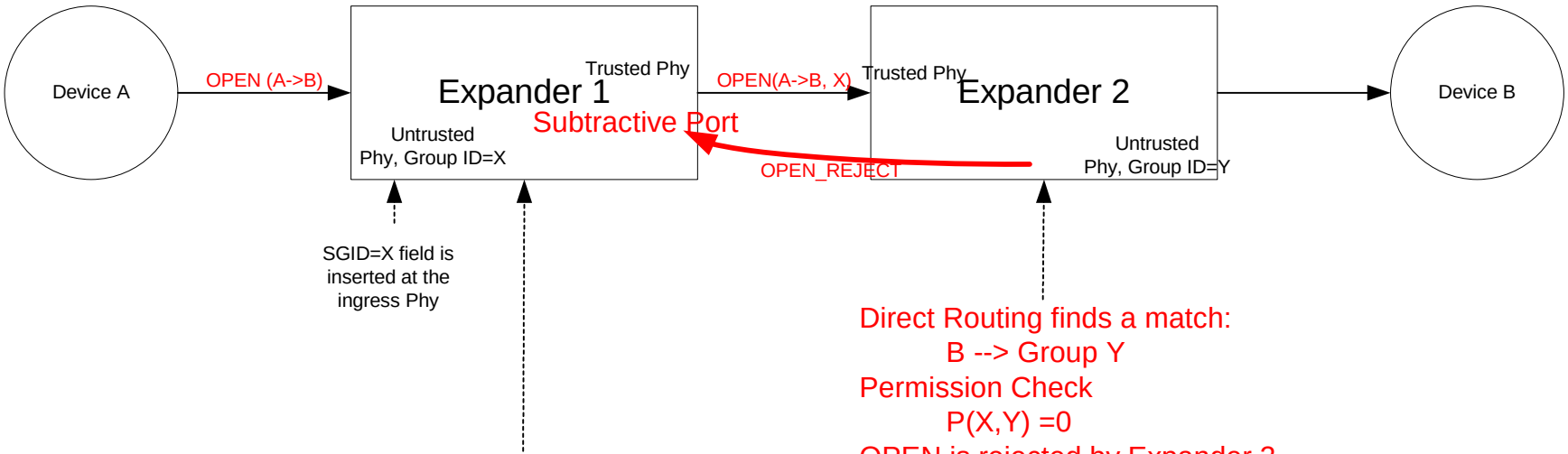
B--> Group Y

Check permission:

P(X,Y) =0

OPEN is rejected by Expander 1

ZONE Permission Check Example Multi-Hop(Subtractive Routing)



Direct Routing finds a match:
B --> Group Y
Permission Check
P(X,Y) =0
OPEN is rejected by Expander 2

Subtractive Routing is enabled.
ECM destination address B not found by direct and table routing method, subtractive routing is used.

OPEN is not rejected by Expander 1

Broadcast limiting

- Broadcast Frames from a source PHY are only propagated to the PHYs that the source PHY can access according to the ZONE PERMISSION Table
  - To limit the propagation of broadcast, the BROADCAST message needs to carry the SOURCE GROUP ID
  - Proposal uses a new address frame to carry this information with CRC protection

— ZONED BROADCAST address frame format

Byte\Bit	7	6	5	4	3	2	1	0
0	Reserved	BROADCAST TYPE			ADDRESS FRAME TYPE (2h)			
1	Reserved							
2								
3	Reserved	SOURCE GROUP ID						
4	Reserved							
27								
28	(MSB)	CRC						
31								(LSB)

Topology Discovery

- The Route table is extended to become ZONE ROUTE TABLE
  - Original ROUTE TABLE maps a SAS address to expander PHY
  - The ZONE ROUTE TABLE maps an SAS address (according to the PHY ZONE configuration at the attached PHY) to:
    - Expander PHY
    - GROUP ID
    - SUPERVISOR
    - TRUSTED
- All Zoning expanders are required to be self configuring expanders:
  - The expander topology discovery process includes the fields above associated with each routed SAS address
  - The DISCOVER command is extended to provide the additional information and the REPORT ZONE ROUTE TABLE command is proposed to replace the current REPORT ROUTE INFO command
- HOST discovery
  - Legacy Host can still use DISCOVER, REPORT ROUTE INFO commands to do topology discovery using legacy algorithm described in SAS 1.1
  - The expanders are responsible for SAS 1.1 compliant response to those commands to ensure the host discovery is done transparently
  - The zoned expander hides the PHYs/SAS address that a host is not permitted to see based on the GROUP ID of the host – such that the host only discovers a partial topology based on the ZONE PERMISSION table

ZONE Management

- SMP Commands are added to support Zone management
  - The Zone expanders only execute these commands if they come from a SUPERVISOR, and that supervisor can be:
    - An end device – the ingress expander PHY supervisor bit defines which end device is supervisor
    - An zone expander device with SMP initiator function
  - ZONE management functions:
    - CONFIGURE PHY ZONE: configure the group assignment and zone boundary for expander PHYs
    - CONFIGURE ZONE PERMISSION: configure the ZONE permission table
      - The configuration procedu
        the same ZONE PERMIS
        determine if zone change
      - A supervisor can send the
        PERMISSION commands
      - A supervisor (end device)
        PERMISSION command
        SUPERVISE THIS bit set)
        responsible for sending th
        command to all expanders in the rest of the topology
    - These commands support “batch” operation to improve efficiency
      - Allow multiple entries to be configured/reported in one command

•The supervising expander is elected to be the zoning expander device with largest SAS WWN in the topology.

•The supervisors shall send ZONE PERMISSION table updates to the supervising Expander using CONFIGURE ZONE PERMISSION command with “PROPAGATE UPDATE” bit set to 1.

•The supervising expander shall propagate the new PERMISSION table to all other expanders in the topology using CONFIGURE ZONE PERMISSION commands with “PROPGATE UPDATE” bit set to 0.

SMP: REPORT GENERAL

REPORT GENERAL response

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (41h)							
1	FUNCTION (10h)							
2	FUNCTION RESULT							
3	Reserved							
4	(MSB)	EXPANDER CHANGE COUNT						
5								(LSB)
6	(MSB)	EXPANDER ROUTE INDEXES						
7								(LSB)
8	Reserved	NUMBER OF ZONES SUPPORTED						
9	NUMBER OF PHYS							
10	Reserved					ZONING SUPPORTED	CONFIGURING	CONFIGURABLE ROUTE TABLE
11	Reserved							
12	ENCLOSURE LOGICAL IDENTIFIER							
19								
20	SUPERVISING EXPANDER SAS ADDRESS							
27								
28	(MSB)	CRC						
31								(LSB)

Remove ZONING SUPPORTED bit as it was deemed redundant information.

- The NUMBER OF ZONES SUPPORTED field indicates the number of zones supported when the ZONING SUPPORTED bit is 1. If this field is zero, the expander does not support zoning. Note that group 0 and group 127 must be supported in all zoning expanders. The remaining zone indexes should range from 1 to (NUMBER OF ZONES –2).
- ~~The ZONING SUPPORTED bit indicates whether the expander device supports the zoning feature.~~
- The SUPERVISING EXPANDER SAS address contains the WWN of the supervising expander that is elected by the zoning expanders in a topology.

SMP: DISCOVER

Position of fields rearranged to be consistent with the PHY ZONE descriptor format

DISCOVER response

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (41h)							
1	FUNCTION (10h)							
2	FUNCTION RESULT							
3	Reserved							
4	Ignored							
7								
8	Reserved							
9	PHY IDENTIFIER							
10	Ignored							
11	Reserved							
12	Ignored	ATTACHED DEVICE TYPE			Ignored			
13	Reserved				NEGOTIATED PHYSICAL LINK RATE			
14					ATTACHED SSP INITIATOR	ATTACHED STP INITIATOR	ATTACHED SMP INITIATOR	ATTACHED SATA HOST
15	ATTACHED SATA PORT SELECTOR	Reserved			ATTACHED SSP TARGET	ATTACHED STP TARGET	ATTACHED SMP TARGET	ATTACHED SATA DEVICE
16	SAS ADDRESS							
23								
24	ATTACHED SAS ADDRESS							
31								
32		ATTACHED PHY IDENTIFIER						
33	Reserved							
39								
40	PROGRAMMED MINIMUM PHYSICAL LINK RATE				HARDWARE MINIMUM PHYSICAL LINK RATE			
41	PROGRAMMAED MAXIMUM PHYSICAL LINK RATE				HARDWARE MAXIMUM PHYSICAL LINK RATE			
42	PHY CHANGE COUNT							
43	VIRTUAL PHY	Reserved			PARTIAL PATHWAY TIMEOUT VALUE			
44	Reserved				ROUTING ATTRIBUTE			
45	Reserved	CONNECTOR TYPE						
46	CONNECTOR ELEMENT INDEX							
47	CONNECTOR PHYSICAL LINK							
48					ZONE VIOLATION	SOURCE CHECK	TRUSTED	SUPERVISOR
49	Reserved	GROUP ID						
50	Vendor Specific							
51								
52	(MSB)	CRC						
55								(LSB)

- The ZONE VIOLATION field is set to 1 if any ZONE violation has occurred causing the specified phy to send OPEN_REJECT(ZONE VIOLIATION). The ZONE VIOLATION shall be cleared if a PHY CONTROL function with operation code of CLEAR ERROR LOG for the specified phy is received from a supervisor.
- The TRUSTED bit reports whether the specified phy is currently configured as trusted phy or untrusted phy by the supervisor.
- The SUPERVISOR bit reports whether the specified phy is currently configured as a zone supervisor phy.
- The SOURCE CHECK bit reports whether the specified phy is doing the source SAS address checking on the specific phy.
- The GROUP ID fields reports the source group ID assignment of the specified phy.

SMP: CONFIGURE PHY ZONE

CONFIGURE PHY ZONE request

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (40h)							
1	FUNCTION (xxh)							
2	Reserved							
3								
4	Ignored							
5								
6	START PHY INDEX							
7	NUMBER OF ZONE PHY ENTRIES							
	PHY ZONE configuration entry list							
8	First PHY ZONE configuration entry descriptor							
9								
…	…							
n-5	Last PHY ZONE configuration entry descriptor							
n-4								
n-3	(MSB) CRC							
n	(LSB)							

- START PHY INDEX field defines the first phy index to be configured.
- The NUMBER OF ZONE PHY ENTRIES field defines how many phy zone entries the request intends to configure. This field has a range of 0 to 255.
- Note that this command configures one or multiple contiguous expander phys starting from START PHY INDEX.
- The PHY ZONE entry descriptor list contains zero or more PHY ZONE entry descriptors

PHY ZONE configuration entry descriptor

Byte\Bit	7	6	5	4	3	2	1	0
0						SOURCE CHECK	TRUSTED	SUPERVISOR
1	Reserved	GROUP ID						

- The GROUP ID field specifies the group ID to be assigned to the specified phy.
- The SUPERVISOR field specifies whether the specified phy is a supervisor.
- The TRUSTED field specifies whether the specified phy is trusted or untrusted.
- The SOURCE CHECK field specifies whether the specified phy shall check the SOURCE SAS address against the SAS address in the IDENTIFY address frame received on the specific phy.

SMP: CONFIGURE ZONE PERMISSION

CONFIGURE ZONE PERMISSION request

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (40h)							
1	FUNCTION (xxh)							
2	Reserved							
3								
4	~~GENERATION CODE~~							
5								
6	SET BATCH	SOURCE GROUP ID						
7	NEW VALUE	TARGET GROUP ID						
8	Ignored							
9							PROPAGATE UPDATE	UPDATE COMPLETE
10	START ZONE ENTRY INDEX							
11	NUMBER OF ZONE PERMISSION ENTRIES							
	ZONE PERMISSION entry list							
12	First ZONE PERMISSION entry descriptor							
27								
…		…						
n-20	Last ZONE PERMISSION entry descriptor							
n-4								
n-3	(MSB)	CRC						
n								(LSB)

ZONE permission entry descriptor

Byte\Bit	7	6	5	4	3	2	1	0
0	(MSB)	ZONE PERMISSION						
15								(LSB)

- The SET BATCH field chooses between the single set mode or batch set mode.
- The NEW_VALUE field is only used in single set mode. It provides the value for permission table between NEW VALUE, SOURCE GROUP ID and the TARGET GROUP ID. For batch set mode, this field should be set to zero. Note that this value is set to both permission table entry [SOURCE GROUP ID bit, TARGET GROUP ID], and permission table entry [TARGET GROUP ID bit, SOURCE GROUP ID].
- The SOURCE GROUP ID field provides the source group ID to be modified by the single set operation.
- The TARGET GROUP ID field provides the target group ID to be modified by the single set operation.
- The PROPAGATE UPDATE bit is set to indicate that the Supervisor is handing the command off to this expander to become the Supervising expander. The Supervising expander is now responsible for propagating the same zone permission table update to other zoning expanders in the fabric. This bit shall only be set if the command is sent by a supervisor to the supervising expander. This bit shall be clear when the command is send by a supervising expander to other expanders in the fabric.
- The UPDATE COMPLETE bit indicates whether the current CONFIGURE ZONE PERMISSION command is the last command of a sequence of CONFIGURE ZONE PERMISSION commands. This may be used as a flag to a supervising expander that a BROADCAST can be generated to the appropriate groups for the changes to zoning.
- ~~The GENERATION CODE field specifies the generation code that must be in effect for the function to be accepted. If the GENERATION CODE field is not set to the current generation code, the SMP target shall return a response of (TBD). The generation code shall be incremented by one each time a CONFIGURE ZONE PERMISSION function with UPDATE COMPLETE set high is completed.~~
- The START ZONE ENTRY INDEX field species the first Zone Permission table entry index to be configure in batch set mode.
- The NUMBER OF ZONE PERMISSION ENTRIES field defines how many zone permission entries the CONFIGURE ZONE PERMISSION request intends to configure starting from START ZONE ENTRY INDEX in batch set mode.
- The PHY ZONE entry descriptor list contains zero or more ZONE PERMISSION entry descriptors in batch set mode..
- Note that n (total number of bytes), is required to be equal to or less than 1032. This limits the number of CONFIGURE ZONE changes to 63.

SMP: REPORT ZONE PERMISSION

REPORT ZONE PERMISSION response

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (41h)							
1	FUNCTION (xxh)							
2	FUNCTION RESULT							
3	Reserved							CONFIGURING
4	GENERATION CODE							
5								
6	START ZONE ENTRY INDEX							
7	NUMBER OF ZONE PERMISSION ENTRIES							
	ZONE PERMISSION entry list							
8	First ZONE PERMISSION entry descriptor							
23								
…		…						
n-19	Last ZONE PERMISSION entry descriptor							
n-4								
n-3	(MSB)	CRC						
n								(LSB)

ZONE permission entry descriptor

Byte\Bit	7	6	5	4	3	2	1	0
0	(MSB)	ZONE PERMISSION						
15								(LSB)

- The GENERATION CODE field indicates the generation of the data returned in the response frame. Each time the zone table changes, the generation code field is incremented. If the management application client detects a different value in the GENERATION CODE field while retrieving one page than it had while retrieving the previous page, it should go back and retrieve all the pages again to obtain a consistent set of information.
- The CONFIGURING field indicates the expander is in the process of zone permission table update and the expander will issue a BROADCAST message when the update is completed.
- The START ZONE ENTRY INDEX field specifies the first Zone Permission table entry index of the first zone permission entry contained in this response frame.
- The NUMBER OF ZONE PERMISSION ENTRIES field defines the number of zone permission entries in the response frame. The response frame contains contiguous permission table entries starting from START ZONE ENTRY INDEX.
- The PHY ZONE entry descriptor list contains zero or more ZONE PERMISSION entry descriptors.
- Note that n (total number of bytes), is required to be equal to or less than 1032. This limits the number of REPORT ZONE changes to 63.

SMP: REPORT ZONE ROUTE TABLE request

REPORT ZONE ROUTE TABLE request

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (40h)							
1	FUNCTION (xxh)							
2	Reserved							
3								
4	NUMBER OF ZONE ROUTE TABLE ENTRIES							
5	PHY IDENTIFIER							
6	(MSB)	START EXPANDER ROUTE INDEX						
7								(LSB)
8		Ignored						
11								
12	(MSB)	CRC						
15								(LSB)

- The NUMBER OF ZONE ROUTE ENTRIES defines how many zone route table entries the REPORT ZONE ROUTE TABLE request intends to read. This command reads the zone route table entries with contiguous expander route index starting from START EXPANDER ROUTE INDEX for PHY IDENTIFIER.
- The PHY IDENTIFIER field specifies the phy for which the expander route entry is being read (see 4.6.7.3).
- The START EXPANDER ROUTE INDEX field specifies the first expander route index for the expander route entry being reported (see 4.6.7.3).

SMP: REPORT ZONE ROUTE TABLE response

REPORT ZONE ROUTE TABLE response

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (41h)							
1	FUNCTION (xxh)							
2	FUNCTION RESULT							
3	Ignored							
4	NUMBER OF ZONE ROUTE TABLE ENTRIES							
5	PHY IDENTIFIER							
6	(MSB)	START EXPANDER ROUTE INDEX						
7								(LSB)
8	GENERATION CODE							
9								
10	Ignored							
11	Reserved						CONFIGURING	END OF ENTRIES
	ZONE ROUTE TABLE entry list							
12	First ZONE ROUTE TABLE entry descriptor							
23								
...	...							
n-15	Last ZONE ROUTE TABLE entry descriptor							
n-4								
n-3	(MSB)	CRC						
n								(LSB)

- The NUMBER OF ZONE ROUTE ENTRIES defines how many zone route table entries the response frame contains
- The PHY IDENTIFIER field specifies the phy for which the expander route entry is being read.
- The CONFIGURING field indicates the expander is in the process of updating its ZONE ROUTE TABLE and the expander will issue a BROADCAST MESSAGE when the ZONE ROUTE TABLE update is completed.
- The START EXPANDER ROUTE INDEX field specifies the first expander route index for the expander route entry being reported.
- The GENERATION CODE field indicates the generation of the data returned in the response frame. Each time the zone table changes, the generation code field is incremented. If the management application client detects a different value in the GENERATION CODE field while retrieving one page than it had while retrieving the previous page, it should go back and retrieve all the pages again to obtain a consistent set of information.
- The END OF ENTRIES field indicates whether the response frame contains the last enabled zoning route table entry of the request PHY.
- The ZONE ROUTE TABLE entry descriptor list contains zero or more ZONE ROUTE TABLE entry descriptors.

Table 20 —ZONE ROUTE TABLE entry descriptor

Byte\Bit	7	6	5	4	3	2	1	0
0	DISABLE EXPANDER ROUTE ENTRY	Reserved						
1	Ignored	ATTACHED DEVICE TYPE			Ignored		TRUSTED	SUPERVISOR
2	Ignored	GROUP ID						
3	Ignored							
4	ROUTED SAS ADDRESS							
11								

- The DISABLE EXPANDER ROUTE ENTRY bit specifies whether this entry is disabled.
- The SUPERVISOR field specifies whether the specified SAS address corresponds to a supervisor.
- The TRUSTED field specifies whether the specified SAS address is trusted or untrusted.
- The ROUTED SAS ADDRESS field contains the routed SAS address for the zone route entry being configured.
- The GROUP ID field contains the GROUP ID for the zone route entry being configured.
- The ZONING EXPANDER field indicates the attached device is a zoning aware expander.

Summary

- The propose SAS zoning scheme provides flexible and efficient function in any SAS physical topology with one or multiple expanders for
  - SAS traffic segregation
  - Access control from any device group to any other device group – more powerful than Ethernet VLAN
  - Broadcast traffic segregation
  - Zoning policy is fully controlled and enforced by the SAS fabric without relying on end devices to be honest
  - Grouping of end devices save the amount of resources required in the expander implementation
- It supports transparent operation with legacy devices:
  - Any SAS 1.1 host or target device will work transparently without knowledge of the “zoning”
  - Any SAS 1.1 expanders can be attached at the edge of a zoning fabric without knowing “zoning”.