

From: Steven Fairchild, HP (Steve.Fairchild@hp.com)
 Date: 3 February 2006
 Subject: SAS BROADCAST (SCSI ASYNCHRONOUS EVENT)

Revision 0; 16 January 2006, initial submission

Revision 1; 2 February 2006, updated at the request of the SAS protocol working group meeting in January 2006.

This proposal defines a new BROADCAST (SCSI ASYNCHRONOUS EVENT) primitive for SAS-2.

BROADCAST (SCSI ASYNCHRONOUS EVENT) is used by a device server to notify an application client that a unit attention condition has been established by one or more logical units accessible through the target port in the device server.

The reasons for a device server generating a unit attention condition defined by SAM-4 include:

“

- a) A hard reset (see 6.3.2), logical unit reset (see 6.3.3), or I_T nexus loss (see 6.3.4) occurs;
- b) A removable medium may have been changed;
- c) The mode parameters associated with this I_T nexus have been changed by a task received on another I_T nexus (i.e., SCSI initiator ports share mode parameters, see SPC-3);
- d) The log parameters associated with this I_T nexus have been changed by a task received on another I_T nexus (i.e., SCSI initiator [ports](#) share log parameters, see SPC-3);
- e) The version or level of microcode has been changed (see SPC-3);
- f) Tasks received on this I_T nexus have been cleared by a task or a task management function associated with another I_T nexus and the TAS bit was set to zero in the Control mode page associated with this I_T nexus (see SPC-3);
- g) INQUIRY data has been changed (see SPC-3);
- h) The logical unit inventory has been changed (see SPC-3);
- i) The mode parameters in effect for the associated I_T nexus have been restored from non-volatile memory (see SPC-3); or
- j) Any other event requiring the attention of the SCSI initiator device.

”

[Note: typo in SAM-4, should be “ports”](#)

If the unit attention condition is the result of aborted tasks, then an application client that is not aware of the unit attention condition may resort to timing out commands and do a reset sequence to the SAS target port of the device server. If more than one application client has commands outstanding to a specific device server, this can result in a “ping pong” affect where application clients are alternately performing reset sequences to SAS target ports which lead to a thrashing behavior on each of the SAS target ports of a SAS target device.

The solution defined in SPC-4 (Control mode page, 0Ah, TAS bit) to handle this situation is:

“A task aborted status (TAS) bit set to zero specifies that aborted tasks shall be terminated by the device server without any response to the application client. A TAS bit set to one specifies that tasks aborted by the actions of an I_T nexus other than the I_T nexus on which the command was received shall be terminated with a TASK ABORTED status (see SAM-3).”

The TAS solution requires the device server to establish a connection with each application client to provide the TASK ABORTED status for each outstanding command.

This presents a problem if the cause of the TASK ABORT status is a hard reset sequence to a misbehaving device server, then the device server may not be able to comply with the TAS solution.

There is also a behavioral conflict, because a hard reset sequence is supposed to have a predictable and low level behavior which is impacted by the requirement of TAS, because the device server is waiting for acknowledgements to the TASK ABORTED status for tasks on one or more target ports, for one or more logical units.

HP solicited feedback from a sampling of drive vendors to determine the level of support for TAS. The results indicated that TAS could not be universally implemented as a solution for application client notification when tasks are aborted. Legal considerations make it impossible for HP to disclose the feedback provided by each drive vendor.

The solution presented here is for the device server's SAS target port(s) to transmit the BROADCAST (SCSI ASYNCHRONOUS EVENT) primitive to notify SAS initiator ports that a unit attention condition has been established by one or more logical units accessible through the SAS target port in the SAS target device.

The drive vendors solicited indicated that they could support the BROADCAST (SCSI ASYNCHRONOUS EVENT) primitive generated at the time a task is aborted by the device server.

During the discussions it was determined that the use of the BROADCAST (SCSI ASYNCHRONOUS EVENT) should be generated based on unit attention conditions rather than specific task abort situations.

Requiring a device server to generate a BROADCAST (SCSI ASYNCHRONOUS EVENT) primitive when it generates a unit attention condition seems more likely to succeed in the potential failing situations, because there is no burden on the device server to establish connections to individual applications clients.

The BROADCAST (SCSI ASYNCHRONOUS EVENT) primitive is only generated outside a connection by the SAS target port of the device server.

For hard resets, the generation of the BROADCAST (SCSI ASYNCHRONOUS EVENT) primitive more closely emulates the behavior of the RST# line in parallel SCSI and can provide a predictable behavior.

In the case of a wide port device server, the BROADCAST (SCSI ASYNCHRONOUS EVENT) would be transmitted on all links of the wide port.

When an application client receives the BROADCAST (SCSI ASYNCHRONOUS EVENT) primitive, it may determine the cause by issuing REQUEST SENSE, TEST UNIT READY or QUERY TASK to each T-L nexus in its queue to see if a unit attention condition has been established.

If the initiator port is attached to a SAS target port through an expander infrastructure, then the application client would need to determine the SAS target port that was the source of the BROADCAST (SCSI ASYNCHRONOUS EVENT) primitive.

The application client may choose to query the expander infrastructure to identify the source of the primitive (i.e. a new SMP command, REPORT PHY BROADCAST COUNTS to query for BROADCAST (SCSI ASYNCHRONOUS EVENT) counts) or may choose to use its outstanding command queue to determine if any of the device server have logical units with outstanding commands pending (i.e. issue REQUEST SENSE, TEST UNIT READY or QUERY TASK to each T-L nexus in its queue to see if a unit attention condition has been established).

In any case, the initiator should notify the upper layers of the driver stack of any outstanding commands that have been aborted. This avoids command recovery trashing and excessive delays waiting for commands to timeout.

Requested changes

Redefine BROADCAST (RESERVED 2) to BROADCAST (SCSI ASYNCHRONOUS EVENT) in table 79, with a description of:

BROADCAST (SCSI ASYNCHRONOUS EVENT)	Notification from a target port that a SCSI asynchronous event has occurred (e.g. unit attention condition).
-------------------------------------	--

Add the following text to section 7.2.5.4 BROADCAST

BROADCAST (SCSI ASYNCHRONOUS EVENT) is transmitted by a SAS target port to notify SAS initiator ports that a unit attention condition has been established by one or more logical units accessible through the SAS target port.

Define a new SMP function that returns the BROADCAST wrapping counters for all phys.

10.4.3.x REPORT PHY BROADCAST COUNTS function

The REPORT PHY BROADCAST COUNTS function returns the BROADCAST primitives received counts from directly attached end devices for the specified phy. This SMP function should be implemented by all SMP target ports in expander devices.

The expander device is not required to increment the fields representing wrapping counters contained in the REPORT PHY BROADCAST COUNTS response again unless a REPORT PHY BROADCAST COUNTS response is transmitted.

NOTE xx - Application clients that use the REPORT PHY BROADCAST COUNTS function should request it often enough to ensure that the counts contained in the REPORT PHY BROADCAST COUNTS response do not increment a multiple of 256 times between requests.

Table xxx defines the request format.

Table xxx – REPORT PHY BROADCAST COUNTS request

Byte/Bit	7	6	5	4	3	2	1	0	
0	SMP FRAME TYPE (40h)								
1	FUNCTION (xxh)								
2	Reserved								
3	REQUEST LENGTH (02h)								
4	Reserved								
8	Reserved								
9	PHY IDENTIFIER								
10	Reserved								
11	Reserved								
12	(MSB)	CRC							
15								(LSB)	

The SMP FRAME TYPE field shall be set to 40h.

The FUNCTION field shall be set to xxh.

The REQUEST LENGTH field shall be set to 02h.

The PHY IDENTIFIER field specifies the phy (see 4.2.7) for the broadcast counters being requested.

The CRC field is defined in 10.4.3.1

Table yyy defines the response format.

Table yyy – REPORT PHY BROADCAST COUNTS response

Byte/Bit	7	6	5	4	3	2	1	0	
0	SMP FRAME TYPE (41h)								
1	FUNCTION (xxh)								
2	FUNCTION RESULT								
3	RESPONSE LENGTH (07h)								
4	Reserved								
8	Reserved								
9	PHY IDENTIFIER								
10	Reserved								
11	Reserved								
12	BROADCAST RESERVED 4 COUNT VALID	BROADCAST RESERVED 3 COUNT VALID	BROADCAST SCSI ASYNC EVENT COUNT VALID	BROADCAST EXPANDER COUNT VALID	BROADCAST SES COUNT VALID	BROADCAST CHANGE RESERVED 1 COUNT VALID	BROADCAST CHANGE RESERVED 0 COUNT VALID	BROADCAST CHANGE COUNT VALID	
13	Reserved								
15	Reserved								
16	BROADCAST CHANGE COUNT								
17	BROADCAST CHANGE RESERVED 0 COUNT								
18	BROADCAST CHANGE RESERVED 1 COUNT								
19	BROADCAST SES COUNT								
20	BROADCAST EXPANDER COUNT								
21	BROADCAST SCSI ASYNCHRONOUS EVENT COUNT								
22	BROADCAST RESERVED 3 COUNT								
23	BROADCAST RESERVED 4 COUNT								
24	(MSB)	CRC							
27							(LSB)		

The SMP FRAME TYPE field shall be set to 41h.

The FUNCTION field shall be set to xxh.

The FUNCTION RESULT field is defined in 10.4.3.2.

The RESPONSE LENGTH field shall be set to 07h.

The PHY IDENTIFIER field specifies the phy (see 4.2.7) for which the broadcast counters is being returned.

A BROADCAST CHANGE COUNT VALID field set to one indicates that the BROADCAST CHANGE COUNT field is supported. A BROADCAST CHANGE COUNT VALID field set to zero indicates that the BROADCAST CHANGE COUNT field is not supported.

A BROADCAST CHANGE RESERVED 0 COUNT VALID field set to one indicates that the BROADCAST CHANGE RESERVED 0 COUNT field is supported. A BROADCAST CHANGE RESERVED 0 COUNT VALID field set to zero indicates that the BROADCAST CHANGE RESERVED 0 COUNT field is not supported.

A BROADCAST CHANGE RESERVED 1 COUNT VALID field set to one indicates that the BROADCAST CHANGE RESERVED 1 COUNT field is supported. A BROADCAST CHANGE RESERVED 1 COUNT VALID field set to zero indicates that the BROADCAST CHANGE RESERVED 1 COUNT field is not supported.

A BROADCAST SES COUNT VALID field set to one indicates that the BROADCAST SES COUNT field is supported. A BROADCAST SES COUNT VALID field set to zero indicates that the BROADCAST SES COUNT field is not supported.

A BROADCAST EXPANDER COUNT VALID field set to one indicates that the BROADCAST EXPANDER COUNT field is supported. A BROADCAST EXPANDER COUNT VALID field set to zero indicates that the BROADCAST EXPANDER COUNT field is not supported.

A BROADCAST SCSI ASYNC EVENT COUNT VALID field set to one indicates that the BROADCAST SCSI ASYNCHRONOUS EVENT COUNT field is supported. A BROADCAST SCSI ASYNCHRONOUS EVENT COUNT VALID field set to zero indicates that the BROADCAST SCSI ASYNCHRONOUS EVENT COUNT field is not supported.

A BROADCAST RESERVED 3 COUNT VALID field set to one indicates that the BROADCAST RESERVED 3 COUNT field is supported. A BROADCAST RESERVED 3 COUNT VALID field set to zero indicates that the BROADCAST RESERVED 3 COUNT field is not supported.

A BROADCAST RESERVED 3 COUNT VALID field set to one indicates that the BROADCAST RESERVED 3 COUNT field is supported. A BROADCAST RESERVED 3 COUNT VALID field set to zero indicates that the BROADCAST RESERVED 3 COUNT field is not supported.

The BROADCAST CHANGE COUNT field indicates the value of a wrapping counter counting the number of BROADCAST (CHANGE)s received from an end device attached to the phy specified by the PHY IDENTIFIER field. Expander devices shall support this field. Other devices shall not support this field. This field shall be set to zero at power on. The expander device shall increment this field at least once when it receives a BROADCAST (CHANGE) from an attached end device. The expander device shall not increment the count for a BROADCAST (CHANGE) received from an attached expander device.

The BROADCAST CHANGE RESERVED 0 COUNT field indicates the value of a wrapping counter counting the number of BROADCAST (RESERVED CHANGE 0)s received from an end device attached to the phy specified by the PHY IDENTIFIER field. Expander devices shall support this field. Other devices shall not support this field. This field shall be set to zero at power on. The expander device shall increment this field at least once when it receives a BROADCAST (RESERVED CHANGE 0) from an attached end device. The expander device shall not increment the count for a BROADCAST (RESERVED CHANGE 0) received from an attached expander device.

The BROADCAST CHANGE RESERVED 1 COUNT field indicates the value of a wrapping counter counting the number of BROADCAST (RESERVED CHANGE 1)s received from an end device attached to the phy specified by the PHY IDENTIFIER field. Expander devices shall support this field. Other devices shall not support this field. This field shall be set to zero at power on. The expander device shall increment this field at least once when it receives a BROADCAST (RESERVED CHANGE 1) from an attached end device. The expander device shall not increment the count for a BROADCAST (RESERVED CHANGE 1) received from an attached expander device.

The BROADCAST SES COUNT field indicates the value of a wrapping counter counting the number of BROADCAST (SES)s received from an end device attached to the phy specified by the PHY IDENTIFIER field. Expander devices shall support this field. Other devices shall not support this field. This field shall be set to zero at power on. The expander device shall increment this field at least once when it receives a BROADCAST (SES) from an attached end device. The expander device shall not increment the count for a BROADCAST (SES) received from an attached expander device.

The BROADCAST EXPANDER COUNT field indicates the value of a wrapping counter counting the number of BROADCAST (EXPANDER)s received from an end device attached to the phy specified by the PHY IDENTIFIER field. Expander devices shall support this field. Other devices shall not support this field. This field shall be set to zero at power on. The expander device shall increment this field at least once when it receives a BROADCAST (EXPANDER) from an attached end device. The expander device shall not increment the count for a BROADCAST (EXPANDER) received from an attached expander device.

The BROADCAST SCSI ASYNCHRONOUS EVENT COUNT field indicates the value of a wrapping counter counting the number of BROADCAST (SCSI ASYNCHRONOUS EVENT)s received from an end device attached to the phy specified by the PHY IDENTIFIER field. Expander devices shall support this field. Other devices shall not support this field. This field shall be set to zero at power on. The expander device shall increment this field at least once when it receives a BROADCAST (SCSI ASYNCHRONOUS EVENT) from an attached end device. The expander device shall not increment the count for a BROADCAST (SCSI ASYNCHRONOUS EVENT) received from an attached expander device.

The BROADCAST RESERVED 3 COUNT field indicates the value of a wrapping counter counting the number of BROADCAST (RESERVED 3)s received from an end device attached to the phy specified by the PHY IDENTIFIER field. Expander devices shall support this field. Other devices shall not support this field. This

field shall be set to zero at power on. The expander device shall increment this field at least once when it receives a BROADCAST (RESERVED 3) from an attached end device. The expander device shall not increment the count for a BROADCAST (RESERVED 3) received from an attached expander device.

The BROADCAST RESERVED 4 COUNT field indicates the value of a wrapping counter counting the number of BROADCAST (RESERVED 4)s received from an end device to the phy specified by the PHY IDENTIFIER field. Expander devices shall support this field. Other devices shall not support this field. This field shall be set to zero at power on. The expander device shall increment this field at least once when it receives a BROADCAST (RESERVED 4) from an attached end device. The expander device shall not increment the count for a BROADCAST (RESERVED 4) received from an attached expander device.

The CRC field is defined in 10.4.3.1.