

To: T10 Technical Committee  
 From: Rob Elliott, HP (elliott@hp.com)  
 Date: 11 September 2005  
 Subject: 04-172r1 SAS-2 More counters

### Revision history

Revision 0 (21 June 2004) First revision

Revision 1 (11 September 2005) Targeted at SAS-2 rather than SAS-1.1. Incorporated comments from July 2004 SAS protocol WG - use variable-length descriptor format rather than fixed format for SMP (making it the same as the log page format), add a mechanism to clear the maximum/peak values, define a BROADCAST for expanders to send when a counter wraps or when a maximum value recorder is cleared

### Related documents

sas1r09e - Serial Attached SCSI 1.1 revision 9e

### Overview

Additional standard per-phy counters or maximum value registers, for both errors and non-errors, are desired to help testing and diagnosing problems, particularly in large configurations.

***These counters should not be construed as required; they are all optional. A reasonable implementation would be to implement a small number of counters (e.g. 4) with vendor-specific means to configure which events are actually counted by those counters (implemented as a MUX with the various measurable events feeding each counter). If a system is suspected to be having a problem with arbitration, then devices might be loaded with firmware to count AIPs and OPEN\_REJECTs to help diagnose the problem. This is similar to the performance counters in Intel IA-64 and AMD AMD64 processors, which feature 4 counters with a 8-bit field to select the implementation-specific event counted by each.***

---



---

[Editor's Note 1: rather than make it vendor-specific which counters are reported, this could be made software-selectable like with the processors.](#)

---



---

Proposed for both SSP targets (reported via the SSP log page) and expanders (reported via SMP functions):

- a) Elasticity buffer overflow - not enough ALIGNS/NOTIFYs coming in, causing loss of a dword. This is important for checking ALIGN/NOTIFY insertion rates and performing clock frequency ppm tests.
- b) Receive address frame error - incoming address frame has CRC or other problem (too few or too many dwords) so is ignored.
- c) ERROR primitive received - phy receives an ERROR primitive (which is not counted as an invalid dword and may or may not have other effects)
- d) Receive BREAK (after transmitting OPEN or during a connection)
- e) Transmit BREAK - possibly differentiate the reasons: Open Timeout timer expires, Close Timeout Timer expires
- f) Break Timeout timer expires. This causes the phy to assume the link is idle again (unless it chooses to start a new link reset sequence).

Proposed for both SSP target ports (reported via log page) and STP target ports (reported via a new SMP function):

- a) Receive an abandon class OPEN\_REJECT
- b) Receive a retry class OPEN\_REJECT
- c) Transmit an abandon class OPEN\_REJECT
- d) Transmit a retry class OPEN\_REJECT
- e) Maximum ARBITRATION WAIT TIME value used in outgoing OPENS
- f) Maximum PATHWAY BLOCKED COUNT value used in outgoing OPENS
- g) Maximum arbitration time. How long did it take to get an OPEN\_ACCEPT or OPEN\_REJECT?

Proposed for SSP targets (reported via the SSP log page):

- a) Receive SSP frame error - SSP target port receives a bad SSP frame and generates NAK (CRC ERROR)
- b) Transmit SSP frame error - SSP target port transmits an SSP frame and receives a NAK or an ACK/NAK timeout
- c) Receives CREDIT\_BLOCKED
- d) Transmits CREDIT\_BLOCKED

Proposed for SMP target ports (reported via SMP functions):

- a) Receive address frame error
- b) Receive SMP frame error - SMP target port receives a bad SMP frame and generates a BREAK
- c) Transmit BREAK
- d) Receive BREAK
- e) Break Timeout. Sent a BREAK and got no reply (a very bad sign).
- f) Transmit an abandon class OPEN\_REJECT
- g) Transmit a retry class OPEN\_REJECT
- h) Connection count. Can use this to tell if some phys in a wide link are not being used as much as others (or at all).
- i) Maximum connection time. Detect if there are any bus hogs.

Proposed counters for STP target ports (reported via SMP function):

- a) STP flow control buffer overflow - received too many dwords after a HOLD, causing a dword to be lost. The cable might be too long, allowing too many dwords in flight.

In SMP, the counters wrap rather than saturate; a bit in the request can be used to query the size of the counter. This facilitates easier multi-initiator usage.

In the log page, the counters are all 32 bits. They wrap rather than saturate for better multi-initiator use.

Wrapping versions of the four existing counters (invalid dword, running disparity error, loss of dword synchronization, and phy reset problem) are also defined.

### **Other possibilities not in this proposal**

In a multiple initiator environment, several of the items would be more useful if recorded on a per destination address basis - maximum AWT used, maximum PBC used, and received CREDIT\_BLOCKED counter. This is not proposed due to complexity. Maximum # RRDYs received per connection to the destination might also be useful.

### **Suggested changes**

#### **7.2.2/3 Primitive summary/encoding**

Change BROADCAST (RESERVED 1) to BROADCAST (EXPANDER)

#### **7.2.3.4 BROADCAST**

BROADCASTs are used to notify all SAS ports in a domain of an event.

The versions of BROADCAST representing different reasons are defined in table 1.

**Table 1 — BROADCAST primitives**

Primitive	Description
BROADCAST (CHANGE)	Notification of a configuration change.
BROADCAST (RESERVED CHANGE 0)	Reserved. Processed the same as BROADCAST (CHANGE) by SAS ports (i.e, SAS initiator ports and SAS target ports).
BROADCAST (RESERVED CHANGE 1)	Reserved. Processed the same as BROADCAST (CHANGE) by SAS ports (i.e., SAS initiator ports and SAS target ports).
BROADCAST (SES)	Notification of an asynchronous event from a logical unit with a peripheral device type set to 0Dh (i.e., enclosure services device) (see SPC-3 and SES-2) in the SAS domain.
BROADCAST ( <del>RESERVED 1-EXPANDER</del> )	<del>Reserved</del> <a href="#">Notification of a phy event information peak value being reached or clearing of peak values.</a>
BROADCAST (RESERVED 2)	Reserved.
BROADCAST (RESERVED 3)	Reserved.
BROADCAST (RESERVED 4)	Reserved.

When an expander port receives a BROADCAST it shall transmit the same BROADCAST on at least one phy in all other expander ports. BROADCAST shall only be sent outside of connections after the phy reset sequence has completed.

An expander device is not required to queue multiple identical BROADCASTs for the same expander port. If a second identical BROADCAST is requested before the first BROADCAST has been transmitted, the second BROADCAST may be ignored.

BROADCAST (CHANGE) is sent by an expander device to notify SAS initiator ports and other expander devices that a configuration change has occurred. BROADCAST (CHANGE) may also be transmitted by SAS initiator ports. BROADCAST (CHANGE) shall be ignored by SAS target ports.

BROADCAST (SES) is sent by a SAS target port to notify SAS initiator ports that an asynchronous event has occurred in an enclosure, and SSP initiator ports should poll all the logical units with peripheral device types set to 0Dh (i.e., enclosure services devices)(see SPC-3 and SES-2) in the SAS domain. BROADCAST (SES) shall be ignored by SAS target ports.

[BROADCAST \(EXPANDER\) is sent by an expander device to notify SAS initiator ports that one or more of its phy event information peak detectors has reached its maximum value, or that the clear peaks function has been performed.](#)

BROADCAST (RESERVED CHANGE 0) and BROADCAST (RESERVED CHANGE 1) shall be processed the same as BROADCAST (CHANGE) by SAS ports. ~~BROADCAST (RESERVED 1)~~, BROADCAST (RESERVED 2), BROADCAST (RESERVED 3), and BROADCAST (RESERVED 4) shall be ignored by SAS ports.

See 7.11 for details on SAS domain changes. See 10.4.3.3 for details on counting BROADCAST (CHANGE) generation in an expander device.

## 10.2.7 SCSI log parameters

### 10.2.7.1 Protocol-Specific log page

The Protocol Specific log page for SAS defined in table 2 is used to report errors that have occurred on the SAS target device's phy(s).

**Table 2 — Protocol-Specific log page for SAS**

Byte\Bit	7	6	5	4	3	2	1	0
0	PAGE CODE (18h)							
1	Reserved							
2	(MSB)	PAGE LENGTH (m - 3)						(LSB)
3								
Protocol-specific log parameters								
4	First protocol-specific log parameter							
...	...							
	n <sup>th</sup> protocol-specific log parameter							
m								

The PAGE CODE field shall be set to 18h.

The PAGE LENGTH field shall be set to the total length in bytes of the log parameters.

Table 3 defines the format for a SAS log parameter.

**Table 3 — Protocol-Specific log parameter format for SAS**

Byte\Bit	7	6	5	4	3	2	1	0
0	(MSB)	PARAMETER CODE (relative target port identifier)						(LSB)
1								

Table 3 — Protocol-Specific log parameter format for SAS

Byte\Bit	7	6	5	4	3	2	1	0
2	DU	DS	TSD	ETC	TMC		LBIN	LP
3	PARAMETER LENGTH (y - 3)							
4	Reserved				PROTOCOL IDENTIFIER (6h)			
5	Reserved							
6	Reserved							
7	NUMBER OF PHYS							
<b>SAS phy log descriptors</b>								
8	First SAS phy log descriptor							
	...							
	Last SAS phy log descriptor							
y								

The PARAMETER CODE field contains the relative target port identifier (see SPC-3) of the SSP target port that this log parameter describes.

Table 4 defines the values for the log parameter control bits for this log parameter.

Table 4 — Parameter control bits for SAS log parameters

Bit	Value	Description
DU	0	The value is provided by the device server.
DS	0	The device server supports saving of the parameter.
TSD	0	The device server manages saving of the parameter.
ETC	0	No threshold comparison is made on this value.
TMC	any	This field is ignored when the ETC bit is 0.
LBIN	1	The parameter is in binary format.
LP	1	The parameter is a list parameter.

The PARAMETER LENGTH field is set to the length of the log parameter minus three.

The PROTOCOL IDENTIFIER field is set to 6h.

The NUMBER OF PHYS field contains the number of SAS phy log descriptors that follow.

Table 5 defines the SAS phy log descriptor. Each SAS phy log descriptor is the same length.

**Table 5 — SAS phy log descriptor (part 1 of 2)**

Byte\Bit	7	6	5	4	3	2	1	0
0	Reserved							
1	PHY IDENTIFIER							
2	Reserved							
3	Reserved							
4	Reserved	ATTACHED DEVICE TYPE			Reserved			
5	Reserved				NEGOTIATED PHYSICAL LINK RATE			
6	Reserved				ATTACHED SSP INITIATOR PORT	ATTACHED STP INITIATOR PORT	ATTACHED SMP INITIATOR PORT	Reserved
7	Reserved				ATTACHED SSP TARGET PORT	ATTACHED STP TARGET PORT	ATTACHED SMP TARGET PORT	Reserved
8	SAS ADDRESS							
15	SAS ADDRESS							
16	ATTACHED SAS ADDRESS							
23	ATTACHED SAS ADDRESS							
24	ATTACHED PHY IDENTIFIER							
25	Reserved							
31	Reserved							

**Table 5 — SAS phy log descriptor (part 2 of 2)**

Byte\Bit	7	6	5	4	3	2	1	0
32	(MSB)	INVALID DWORD COUNT						(LSB)
35		RUNNING DISPARITY ERROR COUNT						(LSB)
36	(MSB)	LOSS OF DWORD SYNCHRONIZATION COUNT						(LSB)
39		PHY RESET PROBLEM COUNT						(LSB)
40	(MSB)	<u>Reserved</u>						(LSB)
43		<u>NUMBER OF PHY EVENT DESCRIPTORS</u>						(LSB)
44	(MSB)	<u>Phy event descriptor(s)</u>						(LSB)
47								(LSB)
<u>48</u>								
<u>51</u>								
<u>52</u>								
<u>53</u>								
<u>m</u>								

NOTE 1 Logical units compliant with SAS and SAS-1.1 only support a 48 byte SAS phy log descriptor. To determine the size of each SAS phy log descriptor, use:

$$\text{Phy log descriptor length} = (\text{parameter length} - 4) / \text{number of phys}.$$

The PHY IDENTIFIER field, ATTACHED DEVICE TYPE field, NEGOTIATED PHYSICAL LINK RATE field, ATTACHED SSP INITIATOR PORT bit, ATTACHED STP INITIATOR PORT bit, ATTACHED SMP INITIATOR PORT bit, ATTACHED SSP TARGET PORT bit, ATTACHED STP TARGET PORT bit, ATTACHED SMP TARGET PORT bit, SAS ADDRESS field, ATTACHED SAS ADDRESS field, and attached PHY IDENTIFIER field are defined in the SMP DISCOVER function (see 10.4.3.5).

The INVALID DWORD COUNT field, RUNNING DISPARITY ERROR COUNT field, LOSS OF DWORD SYNCHRONIZATION field, and PHY RESET PROBLEM COUNT field are each defined in the SMP REPORT PHY ERROR LOG response data (see 10.4.3.6).

For the INVALID DWORD COUNT field, RUNNING DISPARITY ERROR COUNT field, LOSS OF DWORD SYNCHRONIZATION COUNT field, and PHY RESET PROBLEM COUNT field, the phy may maintain any size counter but should maintain a 32-bit counter. If it reaches its maximum value, the count shall stop and a value of FFFFFFFFh shall be returned.

The NUMBER OF PHY EVENT DESCRIPTORS field indicates how many phy event descriptors follow. Each phy event descriptor is 8 bytes long and follows the format defined in table 9 (see 10.4.3.x).

**10.4.3 SMP functions**

**10.4.3.1 SMP function request frame format**

...

The FUNCTION field specifies which SMP function is being requested and is defined in table 6. If the value in the FUNCTION field is not supported by the SMP target port, it shall return a function result of UNKNOWN SMP FUNCTION as described in table 148.

Table 6 — SMP functions

Code	SMP function	Description	Request frame size (in bytes)	Response frame size (in bytes)	Reference
00h	REPORT GENERAL	Return general information about the device	8	32	10.4.3.3
01h	REPORT MANUFACTURER INFORMATION	Return vendor and product identification	8	64	10.4.3.4
02h - 0Fh	Reserved for general SMP input functions				
10h	DISCOVER	Return information about the specified phy	16	56	10.4.3.5
11h	REPORT PHY ERROR LOG	Return error logging information about the specified phy	16	32	10.4.3.6
12h	REPORT PHY SATA	Return information about a phy currently attached to a SATA device	16	60	10.4.3.7
13h	REPORT ROUTE INFORMATION	Return route table information	16	44	10.4.3.8
<a href="#">14h</a>	<a href="#">REPORT PHY EVENT INFORMATION</a>	<a href="#">Return phy event information for the specified phy</a>	<a href="#">16</a>	<a href="#">variable</a>	<a href="#">10.4.3.xx</a>
<del>14h</del> <a href="#">15h</a> - 1Fh	Reserved for phy-based SMP input functions				
20h - 3Fh	Reserved for SMP input functions				
40h - 7Fh	Vendor specific				
80h - 8Fh	Reserved for general SMP output functions				
90h	CONFIGURE ROUTE INFORMATION	Change route table information	44	8	10.4.3.9
91h	PHY CONTROL	Request actions by the specified phy	44	8	10.4.3.10
92h - 9Fh	Reserved for phy-based SMP output functions				
A0h - BFh	Reserved for SMP output functions				
C0h - FFh	Vendor specific				

...



**10.4.3.x REPORT PHY EVENT INFORMATION function**

The REPORT PHY EVENT INFORMATION function returns information about the specified phy. This SMP function may implemented by any SMP target port.

Table 6 defines the request format.

**Table 7 — REPORT PHY EVENT INFORMATION request**

Byte\Bit	7	6	5	4	3	2	1	0	
0	SMP FRAME TYPE (40h)								
1	FUNCTION (14h)								
2	Reserved								
3	Reserved						CLEAR PEAKS	REPORT MAXIMUMS	
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 14h.

A CLEAR PEAKS field set to one specifies that the fields in the response frame containing peak values be set to zero. A CLEAR MAXIMUMS field set to zero specifies no change to the fields in the response frame containing peak values.

A REPORT MAXIMUMS field set to one specifies that the maximum value of each field in the response frame be returned. A REPORT MAXIMUMS field set to zero specifies that the current value of each field in the response frame be returned.

The PHY IDENTIFIER field specifies the phy (see 4.2.7) for which information shall be reported.

The CRC field is defined in 10.4.3.1.

Table 7 defines the response format.

**Table 8 — REPORT PHY EVENT INFORMATION response**

Byte\Bit	7	6	5	4	3	2	1	0	
0	SMP FRAME TYPE (41h)								
1	FUNCTION (14h)								
2	FUNCTION RESULT								
3	Reserved							REPORT MAXIMUMS	
4	Reserved								
8	Reserved								
9	PHY IDENTIFIER								
10	Reserved								
14	Reserved								
15	NUMBER OF PHY EVENT DESCRIPTORS								
53	Phy event descriptor(s)								
n - 4	Phy event descriptor(s)								
n - 3	(MSB)	CRC							
n								(LSB)	

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

The FUNCTION field shall be set to 14h.

The FUNCTION RESULT field is defined in 10.4.3.2.

The REPORT MAXIMUMS field indicates the value of the REPORT MAXIMUMS field in the request.

The PHY IDENTIFIER field indicates the phy (see 4.2.7) for which information is being reported.

The NUMBER OF PHY EVENT DESCRIPTORS field indicates how many phy event descriptors follow.

Each phy event descriptor is 8 bytes long and follows the format defined in table 5.

**Table 9 — Phy event descriptor**

Byte\Bit	7	6	5	4	3	2	1	0	
0	Reserved								
2	Reserved								
3	PHY EVENT CODE								
4	(MSB)	PHY EVENT INFORMATION							
7								(LSB)	

The PHY EVENT CODE field, defined in table 6, indicates the type of phy event information being reported in the PHY EVENT INFORMATION field.

**Table 10** — PHY EVENT INFORMATION field (part 1 of 3)

Code	Name	Description
00h	Invalid dword count	Number of invalid dwords (see 3.1.100) that have been received outside of phy reset sequences (i.e., between when the SP state machine (see 6.8) sends a Phy Layer Ready (SAS) confirmation or Phy Layer Ready (SATA) confirmation and when it sends a Phy Layer Not Ready confirmation to the link layer)
01h	Running disparity error count	Number of dwords containing running disparity errors (see 6.2) that have been received outside of phy reset sequences
02h	Loss of dword synchronization count	Number of times the phy has restarted the link reset sequence because it lost dword synchronization (see 6.9)(i.e., the SP state machine transitioned from SP15:SAS_PHY_Ready or SP22:SATA_PHY_Ready to SP0:OOB_COMINIT (see 6.8))
03h	Phy reset problem count	Number of times the phy did not obtain dword synchronization during the final SAS speed negotiation window (see 6.7.4.2)
04h	Elasticity buffer overflow count	Number of times the phy's receive elasticity buffer (see 7.3) has overflowed (e.g., because it did receive a sufficient number of ALIGNs and/or NOTIFYs)
05h	Received ERROR count	Number of times the phy received an ERROR primitive
06h - 1Fh	Reserved for phy layer-based phy event information	
20h	Received address frame error count	Number of times the phy detected an invalid address frame (e.g., because of a CRC error)
21h	Received OPEN_REJECT abandon count	Number of times the phy transmitted an OPEN address frame and received an abandon-class OPEN_REJECT (see 7.2.5.11). In expander devices, forwarded OPEN_REJECTs shall not be counted.
22h	Received OPEN_REJECT retry count	Number of times the phy transmitted an OPEN address frame and received a retry-class OPEN_REJECT (see 7.2.5.11). In expander devices, forwarded OPEN_REJECTs shall not be counted.
23h	Transmitted OPEN_REJECT abandon count	Number of times the phy received an OPEN address frame and transmitted an abandon-class OPEN_REJECT (see 7.2.5.11) In expander devices, forwarded OPEN_REJECTs shall not be counted.
24h	Transmitted OPEN_REJECT retry count	Number of times the phy received an OPEN address frame and transmitted a retry-class OPEN_REJECT (see 7.2.5.11). In expander devices, forwarded OPEN_REJECTs shall not be counted.
25h	Received AIP (WAITING ON PARTIAL) count	Number of times the phy received an AIP (WAITING ON PARTIAL) or AIP (RESERVED WAITING ON PARTIAL). In expander devices, forwarded AIPs shall be counted.

Table 10 — PHY EVENT INFORMATION field (part 2 of 3)

Code	Name	Description
26h	Received AIP (WAITING ON CONNECTION) count	Number of times the phy received an AIP (WAITING ON CONNECTION).In expander devices, forwarded AIPs shall be counted.
27h	Received BREAK count	Number of times the phy received a BREAK that was not a response to a BREAK that it transmitted
28h	Transmitted BREAK count	Number of times the phy transmitted a BREAK that was not a response to a BREAK it received (e.g., a Close Timeout was detected by the SL state machine interfacing to the SMP target port).
29h	Break Timeout count	Number of times the phy transmitted a BREAK and did not receive a BREAK in response (e.g., as detected by the XL state machine and/or the SL state machine interfacing to the SMP target port).
2Ah	Connection count	Number of connections in which the phy was involved
2Bh	Peak transmitted pathway blocked count	Peak value of a PATHWAY BLOCKED field in an OPEN address frame transmitted by the phy. Since the maximum value is FFh, only byte 3 is used.
2Ch	Peak transmitted arbitration wait time	Peak value of an ARBITRATION WAIT TIME field in an OPEN address frame transmitted by the phy. Since the maximum value is FFFFh, only bytes 2 and 3 are used.
2Dh	Peak arbitration time	Peak time in microseconds after transmitting an OPEN address frame that the phy has waited for connection response (e.g., OPEN_ACCEPT or OPEN_REJECT).
2Eh	Peak connection time	The peak duration, in microseconds, of any connection in which the phy was involved.
2Fh - 3Fh	Reserved for SAS arbitration-related phy information	
40h	Received SSP frame error count	Number of times the phy was used in a connection involving the SSP target port, detected an invalid frame, and transmitted a NAK (CRC ERROR) (e.g., because of a CRC error)
41h	Transmitted SSP frame error count	Number of times the phy was used in a connection involving the SSP target port, transmitted a frame, and received a NAK or an ACK/NAK timeout
42h	Received CREDIT_BLOCKED count	Number of times the phy received a CREDIT_BLOCKED

**Table 10** — PHY EVENT INFORMATION field (part 3 of 3)

Code	Name	Description
43h	Transmitted CREDIT_BLOCKED count	Number of times the phy transmitted a CREDIT_BLOCKED
44h	Transmitted SSP data dword count	Number of data dwords transmitted inside SSP frames (i.e., between SOF and EOF)
45h	Received SSP data dword count	Number of times the phy received inside SSP frames CREDIT_BLOCKED (i.e., between SOF and EOF)
46h - 4Fh	Reserved for SSP-related phy event information	
50h	Transmitted SATA data dword count	Number of data dwords transmitted inside STP or SATA frames (i.e., between SATA_SOF and SATA_EOF)
51h	Received SATA data dword count	Number of times the phy received inside STP or SATA frames (i.e., between SATA_SOF and SATA_EOF)
52h	SATA flow control buffer overflow count	Number of times the phy's STP flow control buffer (see 7.17.3) has overflowed (e.g., because it received more data dwords than allowed after transmitting HOLD during an STP connection). This count should be maintained in the phy transmitting the HOLD and receiving the data dwords, but may be maintained in the phy receiving the HOLD and transmitting the data dwords.
53h - 5Fh	Reserved for STP and SATA-related phy event information	
60h	Receive SMP frame error count	Number of times the phy was used for to access the SMP target port and the SMP target port detected an invalid frame and transmitted a BREAK (e.g., because of a CRC error).
61h - 6Fh	Reserved for STP-related phy event information	
70h - EFh	Reserved	
F0h - FFh	Vendor specific	

The CRC field is defined in 10.4.3.2.

---



---

[Editor's Note 3: end of all-new section](#)

---



---