

Date: December 14, 2006

To: T10 Committee (SCSI)

From: George Penokie (IBM)

Subject: SAS-2: Expander Notification of Temporary Shutdown

## **1 Overview**

Within the Service Delivery Subsystem in the SAS topology we have intelligent devices that will have to have their software updated. The software is updated by issuing WRITE BUFFER commands to a SES device within the expander device. When this occurs there will be a period of time while the flash is being written when the expander will not permit any traffic between it's phys. This period of time can, depending on the implementation, take longer than a minute. After the code is written the expander will cause a reset sequence to occur on all it's phys before allowing traffic to resume.

Needless to say, this will cause timeouts at every layer of SAS for any operations outstanding at the time of the code download. This is a problem that can be handled within the current SAS and SCSI definitions by the application client/initiator device that issued the WRITE BUFFER command. However, any other application client/initiator device will not be happy, as any outstanding operations will without warning just disappear.

This proposal defines a NOTIFY (GOING OFFLINE) that expanders would use to notify all devices it is connected to that it will be temporarily (as in seconds) be shutting down all traffic. The NOTIFY is being proposed instead of a BROADCAST mainly to allow the information to easily and quickly pass between zoning expanders.

## **2 SAS-2 changes**

### **7.2.5.11 NOTIFY**

#### **7.2.5.11.1 NOTIFY overview**

NOTIFY may be transmitted in place of any ALIGN (see 7.2.5.2) being transmitted for physical link rate tolerance management (see 7.3) and rate matching (see 7.13). Substitution of a NOTIFY for an ALIGN may or may not affect the ALIGN rotation (i.e., the NOTIFY may take the place of one of the ALIGNs in the rotation through ALIGN (0), ALIGN (1), ALIGN (2), and ALIGN (3), or it may delay the rotation). A specific NOTIFY shall not be transmitted in more than three consecutive dwords until at least three other dwords have been transmitted.

NOTIFYs are deletable primitives (see 7.3).

The forwarding of NOTIFY through expander devices is as specified in table 1.

~~NOTIFY shall not be forwarded through expander devices. Expander devices shall substitute an ALIGN for a NOTIFY if necessary.~~

SAS target devices are not required to detect every transmitted NOTIFY.

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

**Table 1 — NOTIFY primitives**

Primitive	<a href="#">Forwarded through expander device</a>	Description	Reference
NOTIFY (ENABLE SPINUP)	<a href="#">no</a> <sup>a</sup>	Specify to a SAS target device that it may temporarily consume additional power while transitioning into the active or idle power condition state.	7.2.5.11.2
NOTIFY (POWER LOSS EXPECTED)	<a href="#">no</a> <sup>a</sup>	Specify to a SAS target device that power loss may occur within the time specified by the POWER LOSS TIMEOUT field in the Protocol-Specific Port mode page Shared Port Control subpage (see 10.2.7.2.4).	7.2.5.11.3
<a href="#">NOTIFY (GOING OFFLINE)</a>	<a href="#">yes</a>	<a href="#">Specify to a SAS device that an expander device is going to temporarily disable communication between SAS devices.</a>	7.2.5.11.4
NOTIFY (RESERVED 1)	<a href="#">no</a> <sup>a</sup>	Reserved.	

<sup>a</sup> [Expander devices shall substitute an ALIGN for a NOTIFY if necessary.](#)

NOTIFY (RESERVED 1) ~~and NOTIFY (RESERVED 2)~~ shall be ignored by all devices.

#### 7.2.5.11.2 NOTIFY (ENABLE SPINUP)

...

#### 7.2.5.11.3 NOTIFY (POWER LOSS EXPECTED)

...

#### [7.2.5.11.4 NOTIFY \(GOING OFFLINE\)](#)

[NOTIFY \(GOING OFFLINE\) is transmitted by an expander port to specify that the expander is going to disable all phy to phy communication for a period of time. An application client may determine the length of time the expander device is going to be offline by requesting write buffer command information using a REPORT SUPPORTED OPERATION CODES command \(see SPC-4\).](#)

[NOTIFY \(GOING OFFLINE\) shall be transmitted at least three times on each expander port. After the last NOTIFY \(GOING OFFLINE\) is transmitted the expander device shall wait the time specified in the TIME TO OFFLINE field in the CONFIGURE GENERAL function \(see 10.4.3.15\) before disabling phy to phy communication.](#)

[When the expander device receives NOTIFY \(GOING OFFLINE\) the expander shall transmit NOTIFY \(GOING OFFLINE\) on all expander ports except the expander port from which the NOTIFY \(GOING OFFLINE\) was received.](#)

[The expander device shall only transmit a NOTIFY \(GOING OFFLINE\) on one phy of each expander port.](#)

[After the operation that caused the phy to phy communication to be disabled is complete the expander device shall initiate a link reset sequence on all phys.](#)

**10.4.3.3 REPORT GENERAL function**

The REPORT GENERAL function returns general information about the SAS device (e.g., a SAS device contained in an expander device). This SMP function shall be implemented by all management device servers.

Table 2 defines the request format.

**Table 2 — REPORT GENERAL request**

Byte\Bit	7	6	5	4	3	2	1	0	
0	SMP FRAME TYPE (40h)								
1	FUNCTION (00h)								
2	Reserved								
3	REQUEST LENGTH (00h)								
4	(MSB)	CRC						(LSB)	
7									

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

The FUNCTION field shall be set to 00h.

The REQUEST LENGTH field shall be set to 00h.

The CRC field is defined in 10.4.3.1.

Table 3 defines the response format.

**Table 3 — REPORT GENERAL response (part 1 of 2)**

Byte\Bit	7	6	5	4	3	2	1	0	
0	SMP FRAME TYPE (41h)								
1	FUNCTION (00h)								
2	FUNCTION RESULT								
3	RESPONSE LENGTH (0Ch)								
4	(MSB)	EXPANDER CHANGE COUNT						(LSB)	
5									
6	(MSB)	EXPANDER ROUTE INDEXES						(LSB)	
7									
8	Reserved								
9	NUMBER OF PHYS								
10	TABLE TO TABLE SUPPORTED	Reserved			ZONE ADDRESS RESOLVED SUPPORTED	CONFIGURES OTHERS	CONFIGURING	EXTERNALLY CONFIGURABLE ROUTE TABLE	

Table 3 — REPORT GENERAL response (part 2 of 2)

Byte/Bit	7	6	5	4	3	2	1	0
11	TIME TO OFFLINE Reserved							
12	ENCLOSURE LOGICAL IDENTIFIER							
19	ENCLOSURE LOGICAL IDENTIFIER							
20	Reserved							
29	Reserved							
30	(MSB)	STP BUS INACTIVITY TIME LIMIT						(LSB)
31	STP BUS INACTIVITY TIME LIMIT							
32	(MSB)	STP MAXIMUM CONNECT TIME LIMIT						(LSB)
33	STP MAXIMUM CONNECT TIME LIMIT							
34	(MSB)	STP SMP I_T NEXUS LOSS TIME						(LSB)
35	STP SMP I_T NEXUS LOSS TIME							
36	Reserved			ZONE LOCKED	PHYSICAL PRESENCE SUPPORTED	PHYSICAL PRESENCE ASSERTED	ZONING SUPPORTED	ZONING ENABLED
37	Reserved							
38	(MSB)	MAXIMUM NUMBER OF ROUTED SAS ADDRESSES						(LSB)
39	MAXIMUM NUMBER OF ROUTED SAS ADDRESSES							
40	ACTIVE ZONE MANAGER SAS ADDRESS							
47	ACTIVE ZONE MANAGER SAS ADDRESS							
48	(MSB)	ZONE LOCK INACTIVITY TIME LIMIT						(LSB)
49	ZONE LOCK INACTIVITY TIME LIMIT							
50	Reserved							
51	Reserved							
52	(MSB)	CRC						(LSB)
55	CRC							

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

The FUNCTION field shall be set to 00h.

The FUNCTION RESULT field is defined in 10.4.3.2.

The RESPONSE LENGTH field shall be set to 0Ch. For compatibility with previous versions of this standard, a RESPONSE LENGTH field set to 00h indicates that there are 6 dwords before the CRC field.

The EXPANDER CHANGE COUNT field counts the number of Broadcast (Change)s originated by an expander device (see 7.11). Management device servers in expander devices shall support this field. management

device servers in other device types (e.g., end devices) shall set this field to 0000h. This field shall be set to at least 0001h at power on. If the expander device has originated Broadcast (Change) for any reason described in 7.11 since transmitting a REPORT GENERAL response, it shall increment this field at least once from the value in the previous REPORT GENERAL response. It shall not increment this field when forwarding a Broadcast (Change). This field shall wrap to at least 0001h after the maximum value (i.e., FFFFh) has been reached.

NOTE 1 - Application clients that use the EXPANDER CHANGE COUNT field should read it often enough to ensure that it does not increment a multiple of 65 536 times between reading the field.

NOTE 2 - Management device servers in expander devices compliant with previous versions of this standard may return an EXPANDER CHANGE COUNT field set to 0000h.

The EXPANDER ROUTE INDEXES field contains the maximum number of expander route indexes per phy for the expander device (see 4.6.7.3). Management device servers in externally configurable expander devices containing phy-based expander route tables shall support this field. Management device servers in other device types (e.g., end devices, externally configurable expander devices with expander-based expander route tables, and self-configuring expander devices) shall set the EXPANDER ROUTE INDEXES field to zero. Not all phys in an externally configurable expander device are required to support the maximum number indicated by this field.

The NUMBER OF PHYS field contains the number of phys in the device, including any virtual phys and any vacant phys.

A TABLE TO TABLE SUPPORTED bit set to one indicates the expander device is a self-configuring expander device that supports its table routing phys being attached to table routing phys in other expander devices. The TABLE TO TABLE SUPPORTED bit shall only be set to one if the EXTERNALLY CONFIGURABLE ROUTE TABLE bit is set to zero. A TABLE TO TABLE SUPPORTED bit set to zero indicates the expander device is not a self-configuring expander device that supports its table routing phys being attached to table routing phys in other expander devices.

A ZONE ADDRESS RESOLVED SUPPORTED bit set to one indicates that the zoning expander device supports address resolved zoning. A ZONE ADDRESS RESOLVED SUPPORTED bit set to zero indicates that the zoning expander device does not support address resolved zoning (see 4.9.3.1).

A CONFIGURES OTHERS bit set to one indicates that the expander device is a self-configuring expander device that performs the configuration subprocess defined in 4.8. A CONFIGURES OTHERS bit set to zero indicates the expander device may or may not perform the configuration subprocess. Self-configuring expander devices compliant with this standard shall set the CONFIGURES OTHERS bit to one.

NOTE 3 - If the CONFIGURES OTHERS bit is set to zero, the expander device may configure all externally configurable expander devices in the SAS domain.

A CONFIGURING bit set to one indicates that either:

- a) the management device server is in a self-configuring expander device, the self-configuring expander device's management application client is currently performing the discover process (see 4.7), and it has identified at least one change to its expander routing table; or
- b) the zoning expander device is locked and the zoning expander shadow values differ from the zoning expander active values.

A CONFIGURING bit set to zero indicates that the management device server is not in a self-configuring expander device currently performing the discover process and changing its expander routing table. Changes in this bit from one to zero result in a Broadcast (Change) being originated (see 7.11). Management device servers in self-configuring expander devices shall support this bit. Management device servers in externally configurable expander devices and in other device types shall set the CONFIGURING bit to zero.

An EXTERNALLY CONFIGURABLE ROUTE TABLE bit set to one indicates that the management device server is in an externally configurable expander device that has a phy-based expander route table that is required to be configured with the SMP CONFIGURE ROUTE INFORMATION function (see 4.6.7.3). An EXTERNALLY CONFIGURABLE ROUTE TABLE bit set to zero indicates that the management device server is not in an externally

configurable expander device (e.g., it is in an end device, in a self-configuring expander device, or in an expander device with no phys with table routing attributes).

[The TIME TO OFFLINE field contains the time the expander waits from transmitting a NOTIFY \(GOING OFFLINE\) to disabling phy to phy communication which is set by the CONFIGURE GENERAL function \(see 10.4.3.15\).](#)

The ENCLOSURE LOGICAL IDENTIFIER field identifies the enclosure, if any, in which the device is located, and is defined in SES-2. The ENCLOSURE LOGICAL IDENTIFIER field shall be set to the same value reported by the enclosure services process, if any, for the enclosure. An ENCLOSURE LOGICAL IDENTIFIER field set to zero indicates no enclosure information is available.

The STP BUS INACTIVITY TIME LIMIT field contains the bus inactivity time limit for STP connections which is set by the CONFIGURE GENERAL function (see 10.4.3.15).

The STP MAXIMUM CONNECT TIME LIMIT field contains the maximum connect time limit for STP connections which is set by the CONFIGURE GENERAL function (see 10.4.3.15).

The STP SMP I\_T NEXUS LOSS TIME field contains the time that an STP target port and an SMP initiator port retry certain connection requests which is set by the CONFIGURE GENERAL function (see 10.4.3.15).

A ZONE LOCKED bit set to one indicates that the zoning expander device is locked (see 4.9.6.2). A ZONE LOCKED bit set to zero indicates that the zoning expander device is not locked.

A PHYSICAL PRESENCE SUPPORTED bit set to one indicates that the expander device supports physical presence as a mechanism for allowing zoning to be enabled or disabled from phys in zone groups without access to zone group 2. A PHYSICAL PRESENCE SUPPORTED bit set to zero indicates that the expander device does not support physical presence as a mechanism for allowing zoning to be enabled or disabled.

A PHYSICAL PRESENCE ASSERTED bit set to one indicates that the expander device is currently detecting physical presence. A PHYSICAL PRESENCE ASSERTED bit set to zero indicates that the expander device is not currently detecting physical presence. The PHYSICAL PRESENCE ASSERTED bit shall be set to zero if the PHYSICAL PRESENCE SUPPORTED bit is set to zero.

A ZONING SUPPORTED bit set to one indicates that zoning is supported by the expander device (i.e., it is a zoning expander device). A ZONING SUPPORTED bit set to zero indicates that zoning is not supported by the expander device.

A ZONING ENABLED bit set to one indicates that zoning is enabled in the expander device. A ZONING ENABLED bit set to zero indicates that zoning is disabled in the expander device. The ZONING ENABLED bit shall be set to zero if the ZONING SUPPORTED bit is set to zero.

The MAXIMUM NUMBER OF ROUTED SAS ADDRESSES field contains the number of routed SAS addresses in an expander-based expander route table (see 4.6.7.3 and 4.9.3.4). Management device servers in expander devices containing expander-based expander route tables shall support this field. Management device servers in other device types (e.g., end devices and expander devices with phy-based expander route tables) shall set this field to 0000h.

The ACTIVE ZONE MANAGER SAS ADDRESS field indicates the SAS address of the zone manager that last locked the zoning expander device. If the zoning expander device is currently being configured by a vendor-specific sideband method then the ACTIVE ZONE MANAGER SAS ADDRESS field shall be set to zero. This field shall be set to zero at power on.

The ZONE LOCK INACTIVITY TIME LIMIT field indicates the minimum time between any SMP ZONE LOCK requests, SMP zone configuration function requests, or SMP ZONE ACTIVATE requests from the active zone manager that the locked expander device allows and is set in the SMP ZONE LOCK request (see 10.4.3.18).

The CRC field is defined in 10.4.3.2.

#### **10.4.3.15 CONFIGURE GENERAL function**

The CONFIGURE GENERAL function requests actions by the device containing the management device server. This SMP function may be implemented by any management device server. In zoning expander

devices, if zoning is enabled then this function shall only be processed from SMP initiator ports that have access to zone group 2 (see 4.9.3.2).

Table 4 defines the request format.

**Table 4 — CONFIGURE GENERAL request**

Byte\Bit	7	6	5	4	3	2	1	0	
0	SMP FRAME TYPE (40h)								
1	FUNCTION (80h)								
2	Reserved								
3	REQUEST LENGTH (03h)								
4	(MSB)	EXPECTED EXPANDER CHANGE COUNT						(LSB)	
5									
6	Reserved								
7									
8	Reserved					UPDATE STP SMP I_T NEXUS LOSS TIME	UPDATE STP MAXIMUM CONNECT TIME LIMIT	UPDATE STP BUS INACTIVITY TIME LIMIT	
9	<del>TIME TO OFFLINE</del> Reserved								
10	(MSB)	STP BUS INACTIVITY TIME LIMIT						(LSB)	
11									
12	(MSB)	STP MAXIMUM CONNECT TIME LIMIT						(LSB)	
13									
14	(MSB)	STP SMP I_T NEXUS LOSS TIME						(LSB)	
15									
16	(MSB)	CRC						(LSB)	
19									

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

The FUNCTION field shall be set to 80h.

The REQUEST LENGTH field shall be set to 03h.

If the management device server is not in an expander device or the EXPECTED EXPANDER CHANGE COUNT field is set to 0000h, the EXPECTED EXPANDER CHANGE COUNT field shall be ignored. If the management device server is in an expander device and the EXPECTED EXPANDER CHANGE COUNT field is not set to 0000h, then:

- a) if the EXPECTED EXPANDER CHANGE COUNT field contains the current expander change count (i.e., the value of the EXPANDER CHANGE COUNT field that would be returned by an SMP REPORT GENERAL response at this time), the management device server shall process the function; and

- b) If the EXPECTED EXPANDER CHANGE COUNT field does not contain the current expander change count, the management device server shall return a function result of INVALID EXPANDER CHANGE COUNT in the response frame.

An UPDATE STP BUS INACTIVITY TIME LIMIT bit set to one specifies that the STP BUS INACTIVITY TIME LIMIT field shall be honored. An UPDATE STP BUS INACTIVITY TIME LIMIT bit set to zero specifies that the STP BUS INACTIVITY TIME LIMIT field shall be ignored.

An UPDATE STP MAXIMUM CONNECT TIME LIMIT bit set to one specifies that the STP MAXIMUM CONNECT TIME LIMIT field shall be honored. An UPDATE STP MAXIMUM CONNECT TIME LIMIT bit set to zero specifies that the STP MAXIMUM CONNECT TIME LIMIT field shall be ignored.

An UPDATE STP SMP I\_T NEXUS LOSS TIME bit set to one specifies that the STP SMP I\_T NEXUS LOSS TIME field shall be honored. An UPDATE STP SMP I\_T NEXUS LOSS TIME bit set to zero specifies that the STP SMP I\_T NEXUS LOSS TIME field shall be ignored.

The TIME TO OFFLINE field contains the minimum period that an expander device shall wait from the transmission of a NOTIFY (GOING OFFLINE) to disabling phy to phy communication (see 7.2.5.11.4). This timer value shall be in 10 ms increments. When this timer is exceeded, the expander device may disable phy to phy communication. A value of zero in this field specifies that the time from transmission of a NOTIFY (GOING OFFLINE) to disabling phy to phy communication is vendor specific. This value is reported in the TIME TO OFFLINE field in the SMP REPORT GENERAL response (see 10.4.3.3). The expander device should set the default value for the TIME TO OFFLINE field to 1 s (i.e., 64h).

The STP BUS INACTIVITY TIME LIMIT field contains the maximum period that an STP target port is permitted to maintain a connection (see 4.1.12) without transferring a frame to the STP initiator port. This value shall be the number of 100  $\mu$ s increments between frames that the STP target port transmits during a connection. When this number is exceeded, the STP target port shall close the connection. A value of zero in this field specifies that there is no bus inactivity time limit. This value is reported in the STP BUS INACTIVITY TIME LIMIT field in the SMP REPORT GENERAL response (see 10.4.3.3). The bus inactivity time limit is enforced by the port layer (see 8.2.3).

The STP MAXIMUM CONNECT TIME LIMIT field contains the maximum duration of a connection (see 4.1.12). This value shall be the number of 100  $\mu$ s increments that an STP target port transmits during a connection after which the STP target port shall connection at the next opportunity (e.g., a value of one in this field means that the time is less than or equal to 100  $\mu$ s and a value of two in this field means that the time is less than or equal to 200  $\mu$ s). If an STP target port is transferring a frame when the maximum connection time limit is exceeded, the STP target port shall complete transfer of the frame before closing the connection. A value of zero in this field specifies that there is no maximum connection time limit. This value is reported in the STP MAXIMUM CONNECT TIME LIMIT in the SMP REPORT GENERAL response (see 10.4.3.3). The maximum connection time limit is enforced by the port layer (see 8.2.3).

The STP SMP I\_T NEXUS LOSS TIME field contains the time that an STP target port or SMP initiator port shall retry connection requests that are rejected with responses indicating the destination port may no longer be present (see 8.2.2) before recognizing an I\_T nexus loss (see 4.5). Table 5 defines the values of the STP SMP I\_T NEXUS LOSS TIME field. This value is enforced by the port layer (see 8.2.2).

**Table 5 — STP SMP I\_T NEXUS LOSS TIME field**

Code	Description
0000h	Vendor-specific amount of time.
0001h to FFFFh	Time in milliseconds.
FFFFh	The port shall never recognize an I_T nexus loss (i.e., it shall retry the connection requests forever).

NOTE 4 - The default value of the STP SMP I\_T NEXUS LOSS TIME field should be non-zero. It is recommended that this value be 2 000 ms.



## T10/07-008 revision 0

NOTE 5 - An STP initiator port should retry connection requests for the time indicated by the STP SMP\_I\_T NEXUS LOSS field in the SMP REPORT GENERAL response for the STP target port to which it is trying to establish a connection.

The CRC field is defined in 10.4.3.1.

Table 4 defines the response format.

**Table 6 — CONFIGURE GENERAL response**

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (41h)							
1	FUNCTION (80h)							
2	FUNCTION RESULT							
3	RESPONSE LENGTH (00h)							
4	(MSB)	CRC						(LSB)
7								

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

The FUNCTION field shall be set to 80h.

The FUNCTION RESULT field is defined in 10.4.3.2.

The RESPONSE LENGTH field shall be set to 00h.

The CRC field is defined in 10.4.3.2.