Sept. 19, 2002

To: T10 Technical Committee

From : Jim Reif
        Hewlett-Packard
        20555 SH 249
        Houston, TX 77070
        jim.reif@hp.com

Re: SAS Change Notification Proposal

Overview

The current rules of the SAS specification regarding how Expanders schedule CHANGE primitive sequences are as follows:

1) The CHANGE primitive sequence shall only be sent outside of a connection.
2) The expander device shall transmit CHANGE to one physical link attached to each expander port.
3) The expander device should not transmit CHANGE to more than one physical link per expander port.
4) CHANGE shall not be transmitted to any phy that is part of the expander port that is the cause for sending CHANGE.

Rules 2 and 3, while on the surface appear to be rather simple, turn out to require a substantial amount of gates and intelligence within expanders. The problem is associated with how an expander makes the determination of which of its PHYs are members of a wide port. The expander requires recursive 64-bit comparisons between each of the SAS Addresses that were exchanged on each of its ports. Depending on how an expander is connected to other SAS devices in a topology, it could have multiple wide ports, with each wide port having a different width. The process involved in detecting these different configurations is cumbersome and expensive in hardware implementations.

Since the intent of the CHANGE primitive was to give initiators and self-configuring expanders notification of changes in topologies, rule 2, which requires CHANGE to be sent on *each* port of an expander does not provide any real benefit for Target devices attached to expander ports. Also, since initiators and self-configuring expanders have to build a topology map and are aware of their own wide ports and those of all expanders, they should be given the capability via an SMP request to control which pathways in the topology are used to send CHANGE primitives on. This would give initiators and self-configuring expanders the responsibility of guaranteeing that only one physical link per wide expander port is used for change notification, and it would relieve the expander from the burden of wide port determination.

It is therefore recommended that rule 2 be changed as follows:

2) The expander device shall transmit CHANGE to one physical link attached to each expander port *that has been enabled for change notification by an Initiator or self-configuring expander via the SMP PHY CONTROL(Enable Change Notification) request.*

This will require two new PHY operation codes for the SMP PHY CONTROL request, one for enabling a PHY for change notification and one for disabling change notification. There probably should also be a means for reading whether or not a PHY is currently enabled or disabled for change notification as well, so a reserved bit in the DISCOVER Response should be used indicate this status. If there are multiple expanders between an initiator or a self-configuring expander and other end devices, they are responsible for sending one SMP PHY CONTROL(Enable Change Notification) request to each expander in the pathway to the end devices.

A possible new rule for expanders with respect to this new SMP request might be the following:

An expander shall not queue up any hot-plug or unplug events detected on its links until at least one of its PHYs has been enabled for change notification.

This rule would prevent initiators or self-configuring expanders from having to rediscover the topology if any hot-plug events occurred while they are doing topology discovery but before they had enabled CHANGE notification.

Remaining issues:
1) **How do we guarantee that multiple initiators or self-configuring expanders enable the same single physical link of a wide port between expanders? Should we make them always enable the link with the smallest-numbered phy identifier?**
2) **Should initiators be responsible for enabling the appropriate phys of self-configuring expanders, or should self-configuring expanders enable their own phys by some means other than SMP PHY CONTROL?**

Required changes to SAS Specification:

1) Change the following paragraph from**:**

4.8.4.3 CHANGE primitive processing
The Expander Function shall process CHANGE primitive requests specified by each link via the Request.Change expander primitive. The Expander Function is responsible to schedule the transmission of a CHANGE primitive sequence on each port (using the Indicate.Change expander primitive) other than the one on which it received the CHANGE primitive sequence. The Expander Function shall ensure that the CHANGE primitive sequence is only scheduled for transmission on a single link of a wide port.

To:

4.8.4.3 CHANGE primitive processing
The Expander Function shall process CHANGE primitive requests specified by each link via the Request.Change expander primitive. The Expander Function is responsible to schedule the transmission of a CHANGE primitive sequence on each port (using the Indicate.Change expander primitive) other than the one on which it received the CHANGE primitive sequence **that has been enabled for change notification.**

2) Change the following paragraphs from:

7.11 Domain changes
SAS initiator ports scan the domain with SMP (see 7.17) to search for expander devices and target ports after power on or receiving a CHANGE primitive sequence.

The CHANGE primitive sequence shall only be sent outside of a connection. The expander device shall transmit CHANGE to one physical link attached to each expander port. The expander device should not transmit CHANGE to more than one physical link per expander port.

To:

7.11 Domain changes
A SAS initiator shall issue one SMP PHY CONTROL(Enable Change Notification) request to each non-self-configuring expander that resides between itself and other SAS end devices.

A self-configuring expander shall issue one SMP PHY CONTROL(Enable Change Notification) request to each non-self-configuring expander that resides between itself and SAS end devices.

SAS initiator ports **and self-configuring expanders** scan the domain with SMP (see 7.17) to search for expander devices and target ports after power on or receiving a CHANGE primitive sequence.

The CHANGE primitive sequence shall only be sent outside of a connection. The expander device shall transmit CHANGE to one physical link attached to each expander port **that has been enabled for change notification via the SMP PHY CONTROL(Enable Change Notification) request if the expander device is not self-configuring, or via some other mechanism if the expander device is self-configuring.**
The expander device should not transmit CHANGE to more than one physical link per expander port, **provided that an initiator or self-configuring expander properly enabled only one physical link per expander port for change notification.**

An expander device shall not queue up any hot-plug or unplug events detected on its links until at least one of its PHYs has been enabled for change notification.

Expander devices shall default at power-up time to have all of their PHYs disabled for change notification.

3) Change Table 103 from:

Table 103. Phy operation

PHY OPERATION	Operation	Description
00h	NOP	No operation.
01h	LINK RESET	Perform a link reset sequence on the specified phy.
02h	HARD_RESET	Perform a link reset sequence on the specified phy. Before the IDENTIFY address frame, transmit HARD_RESET repeatedly. After the next link reset sequence recognized, do not transmit HARD_RESET; transmit an IDENTIFY address frame again.
03h	ENABLE	Enable the specified phy. Perform a link reset sequence.
04h	DISABLE	Disable the specified phy. Stop transmitting and receiving.
05h	NEA LOOPBACK	Set the specified phy to near-end analog loopback mode (see **Error! Reference source not found.**).
06h	CLEAR ERROR LOG	Clear the error log counters for the specified phy.
All others		Reserved.

To:

Table 103. Phy operation

PHY OPERATION	Operation	Description
00h	NOP	No operation.
01h	LINK RESET	Perform a link reset sequence on the specified phy.
02h	HARD_RESET	Perform a link reset sequence on the specified phy. Before the IDENTIFY address frame, transmit HARD_RESET repeatedly. After the next link reset sequence recognized, do not transmit HARD_RESET; transmit an IDENTIFY address frame again.
03h	ENABLE	Enable the specified phy. Perform a link reset sequence.
04h	DISABLE	Disable the specified phy. Stop transmitting and receiving.
05h	NEA LOOPBACK	Set the specified phy to near-end analog loopback mode (see **Error! Reference source not found.**).
06h	CLEAR ERROR LOG	Clear the error log counters for the specified phy.
07h	ENABLE FOR CHANGE NOTIFICATION	Enables the specified PHY for transmitting CHANGE primitive sequences.
08h	DISABLE FOR CHANGE NOTIFICATION	Disables the specified PHY from transmitting CHANGE primitive sequences.
All others		Reserved.

4) Change the SMP DISCOVER Response to the following (redefine Bit 7 of Byte 13):

Table 88. DISCOVER response

Byte	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (41h)							
1	FUNCTION (00h)							
2	FUNCTION RESULT							
3	Reserved							
4 7	Ignored							
8	Reserved							
9	PHY IDENTIFIER							
10	Ignored							
11	Reserved							
12	~~ROUTE ENTRY DISABLED~~ Ignored	ATTACHED DEVICE TYPE			ROUTING METHOD			
13	**ENABLED FOR CHANGE NOTIFICATON**	Reserved			PHYSICAL LINK RATE			
14	Reserved				ATTACHED SSP INITIATOR	ATTACHED STP INITIATOR	ATTACHED SMP INITIATOR	Reserved
15	Reserved				ATTACHED SSP TARGET	ATTACHED STP TARGET	ATTACHED SMP TARGET	ATTACHED SATA TARGET
16 23	(MSB) ATTACHED SAS ADDRESS (LSB)							
24 31	(MSB) SAS ADDRESS (LSB)							
32	PROGRAMMED MINIMUM PHYSICAL LINK RATE				HARDWARE MINIMUM PHYSICAL LINK RATE			
33	PROGRAMMED MAXIMUM PHYSICAL LINK RATE				HARDWARE MAXIMUM PHYSICAL LINK RATE			
34 35	Ignored							
36 39	Reserved							
40 43	(MSB) CRC (LSB)							

ENABLED FOR CHANGE NOTIFICATON IF SET, INDICATES THAT THE SPECIFIED PHY HAS BEEN ENABLED FOR CHANGE NOTIFICATION.