

To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 3 November 2003
Subject: 03-345r1 SAS-1.1 Wide STP port livelock prevention

Revision history

Revision 0 (8 October 2003) First revision

Revision 1 (3 November 2003) Incorporated comments from November SAS WG.

Related documents

sas1r01 - Serial Attached SCSI 1.1 revision 1

Overview

If a wide STP initiator port and a wide STP target port try to open connections to each other at the same time, the connection requests might pass on different physical links.

Four results are possible:

- a) initiator accepts, target accepts. Two parallel connections are established. The STP/SATA link layer does not know how to use both connections, so this must be prohibited (7.17.3 prohibits it, but doesn't explain how to ensure it does not happen)
- b) initiator accepts, target rejects. Works.
- c) initiator rejects, target accepts. Works.
- d) initiator rejects, target rejects. No forward progress is made; livelock results when both sides keep trying.

Rules are needed to ensure that either b) or c) is the result, not a) or d).

SSP has similar issues, although the SSP link layer can benefit from multiple connections if it chooses. For SSP, an SSP initiator port is not allowed to reject the incoming request because it has an outgoing one to avoid livelocks (8.2.2.3.5).

Since the X_RDY crossing rules mean that the host has to defer to the target after the connections are established, there is some symmetry in b) (target rejects) where the target's request wins just like its X_RDY is going to win.

However, there could be a performance advantage with c) (initiator rejects). The target could reply to the incoming request with OPEN ACCEPT and immediately follow it with X_RDY. The initiator has to wait for the OPEN_ACCEPT to arrive before sending its X_RDY back. It can drop its X_RDY sooner and replace it with R_RDY (because it saw an incoming X_RDY) without incurring a roundtrip delay.

Since wide STP initiator port designs are well underway today, but wide STP target ports are probably not, placing the new requirement on the (future) STP target ports has the least impact (but loses the performance advantage); therefore b) is proposed.

Suggested changes

7.17.3 Affiliations

Coherent access to the SATA task file registers shall be provided for each STP initiator port. STP target ports that do not track all commands by the STP initiator ports' SAS addresses shall implement affiliations to provide coherency. STP target ports that track all commands by the STP initiator ports' SAS addresses shall not implement affiliations.

An affiliation is a state entered by an STP target port where it refuses to accept connection requests from STP initiator ports other than the one that has established an affiliation.

An STP target port that supports affiliations shall establish an affiliation whenever it accepts a connection request. When an affiliation is established, the STP target port shall reject all subsequent connection requests from other STP initiator ports with OPEN_REJECT (STP RESOURCES BUSY).

An STP target port shall maintain an affiliation until any of the following occurs:

- a) Power on;
- b) the SAS target device receives an SMP PHY CONTROL request specifying the phy with the affiliation and specifying a phy operation of HARD RESET (see 10.4.3.10) from any SMP initiator port;
- c) the SAS target device receives an SMP PHY CONTROL request specifying the phy with the affiliation and specifying a phy operation of CLEAR AFFILIATION from the same SAS initiator port that has the affiliation;
- d) A connection to the phy with the affiliation is closed with CLOSE (CLEAR AFFILIATION); or
- e) the STP target port is part of a STP/SATA bridge and a link reset sequence is begun on the SATA physical link.

An affiliation established when the command is transmitted shall be maintained until all frames for the command have been delivered. An STP initiator port implementing command queuing shall maintain an affiliation while any commands are outstanding. This avoids confusing the SATA device, which only knows about one SATA host. STP initiator ports may keep affiliations for longer tenures, but this is discouraged.

An STP target port that implements affiliations shall implement one affiliation per STP target port. Multiple phys on the same STP target port shall use the same affiliation. Support for affiliations is indicated in the SMP REPORT SATA PHY function response (see 10.4.3.7).

~~Only one connection between a wide STP initiator port and a wide STP target port shall be allowed at one time. The STP target port shall reject a second connection request from the same STP initiator port with OPEN_REJECT (STP RESOURCES BUSY).~~

7.17.4 Opening an STP connection

If no STP connection exists when the SATA host port in an STP/SATA bridge receives a SATA_X_RDY from the attached SATA device port, the STP target port in the STP/SATA bridge shall establish an STP connection to the appropriate STP initiator port before it transmits a SATA_R_RDY to the SATA device.

Wide STP initiator ports shall not request more than one connection at a time to an STP target port. Wide STP target ports shall not request more than one connection at a time to an STP initiator port.

While a wide STP target port is waiting for a response to a connection request or has established a connection to an STP initiator port, it shall:

- a) reject incoming connection requests from that STP initiator port with OPEN_REJECT (RETRY); and
- b) if affiliations are supported, reject incoming connection requests from other STP initiator ports with OPEN_REJECT (STP RESOURCES BUSY).

While a wide STP initiator port is waiting for a response to a connection request to an STP target port, it shall not reject an incoming connection request from that STP target port because of its outgoing connection request. It may reject incoming connection requests for other reasons (see 7.2.5.11).

7.17.5 Closing an STP connection

Either STP port (i.e., either the STP initiator port or the STP target port) may originate closing an STP connection. An STP port shall not originate closing an STP connection after sending a SATA_X_RDY or SATA_R_RDY until after both sending and receiving SATA_SYNC. An STP port shall transmit CLOSE after receiving a CLOSE if it has not already transmitted CLOSE.

When an STP initiator port closes an STP connection, it shall transmit a CLOSE (NORMAL) or CLOSE (CLEAR AFFILIATION). When an STP target port closes an STP connection, it shall transmit a CLOSE (NORMAL).

An STP initiator port may issue CLOSE (CLEAR AFFILIATION) in place of a CLOSE (NORMAL) to cause the STP target port to clear the affiliation (see 7.17.3) along with closing the connection. If an STP target port receives CLOSE (CLEAR AFFILIATION), the STP target port shall clear the affiliation for the STP initiator port that sent the CLOSE (CLEAR AFFILIATION).

See 7.12.7 for additional details on closing connections.