To: T10 Technical Committee
From: Bob Sheffield (Robert.L.Sheffield@Intel.com), Intel Corporation
Date: October 24, 2002
Subject: T10/02-435r1, SAS STP Buffering

Revision History

Revision 0 (24 October 2002) first revision
Revision 1 (25 October 2002) Incorporate feedback (discover latency & buffering, and insert ALIGNs to compensate).

Related Documents

SAS-r02b – Serial Attached SCSI revision 02b

Overview

The requirement for an expander that supports attachment of a SATA target to provide additional buffering to compensate for the latency it adds in propogating SATA flow-control primitives is not adequately bounded to establish a proper basis for interoperability. Only the expander device that actually has a SATA disk attached is aware of the need to buffer data. But when the SAS domain consists of multiple levels of expander devices, including fan-out expander devices and edge expander device sets composed of edge expander devices, the expander device with the SATA target attached has no means to determine the added latency between the STP initiator and SATA target, and therefore has no means to determine the amount of buffering required. Even if there were a means to determine dynamically the amount of buffering required, an expander must be built with a fixed amount of buffering capability and cannot dynamically allocate more based on the topology it finds itself in.

This proposal provides a mechanism to establish interoperability, albiet with a performance penalty in some cases, regardless of the expander topology by establishing the following additions to the SAS protocol:

- The addition of SMP commands to allow STP Initiators to discover the Dword propogation latency introduced by each expander and the size in Dwords of the STP HOLD/HOLDA overflow buffers on each port of each edge expander or edge expander set.
- Requires an STP initiator to avoid overflowing the STP HOLD/HOLDA buffers by inserting sufficient ALIGN characters into outbound data frames to establish an outbound effective transfer rate so that the expander's buffer will not overflow before the STP initiator recognizes the HOLD primitive and the returned HOLDA completes the trip back to the STP Target port.

Suggested Changes

In section 10.3.1.3 Table 100 – Report SATA CAPABILITIES Response, change the definition of byte-11 to:

Dwords Buffering per SATA-capable Target Port

Add the definition below the table:

Dwords Buffering per SATA-capable Target Port: The size in Dwords of the internal buffers (for each SATA-capable Target Port) used to buffer Target-bound data during propogation of a HOLD primitive from the target to the STP initiator, and propogation of the HOLDA back from the STP Initiator back to the SATA-capable Target Port.

In section 10.3.1.2 Table 98 – Report General response, modify the reserved byte in Byte 11 to read:

Expander Dword Propogation Latency

Add the description after the table:

Expander Dword Propogation Latency: The average time in microseconds a Dword is in transit through the expander as it travels from one link to another.

Add a new section, 17.x, following 17.15 as follows:

17.x STP Initiator ALIGN Insertion for Flow Control

During Discovery, an STP initiator shall determine the amount of buffering per SATA-capable port provided in each edge expander that supports attachment of SATA devices. Also an STP Initiator shall discover the latency introduced by each expander providing a pathway to SATA devices. When sending data to a SATA target, an STP Initiator shall insert sufficient ALIGN primitives into the data stream to make certain that the edge-expander's buffer will not overflow in the time it takes for a HOLD primitive sent by a SATA target to make the trip to the STP initiator, and for the HOLDA sent in reply to make the return trip back to the SATA Target Port.