To: T10 Technical Committee
From: Bob Sheffield (Robert.L.Sheffield@Intel.com), Intel Corporation
Date: November 5, 2002
Subject: T10/02-435r3, 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).
Revision 2 (November 4, 2002) Revise to have all expanders in pathway account for STP buffering requirements.

## Related Documents

SAS-r02c – Serial Attached SCSI revision 02c

## Overview

Serial Attached SCSI revision 02c requires an expander device that supports attachment of a SATA target to provide additional buffering to compensate for the HOLD/HOLDA turnaround time at each stage through the pathway between an STP initiator port and an STP target device, however it does not specify the allocation of that requirement across pathways involving multiple expander devices. Also the current text is erroneous in the assertion that each expander device must add buffering to account for latency through the expander device. This is not necessary.

To support scalable expander device topologies, each expander device in the pathway must participate in the SATA_HOLD/SATA_HOLDA hand-shake protocol and provide buffering to compensate for HOLD/HOLDA turnaround delay at each stage through the pathway. This proposal incorporates the following elements to establish a basis for STP flow control to work properly in across pathways involving multiple expander devices:

- Requires that each expander device along the pathway of an STP connection participate in the SATA_HOLD/SATA_HOLDA handshake in both directions in a symetrical fashion,
- Requires that each expander device along the pathway of an STP connection provide dword buffering to compensate for the HOLD/HOLDA turnaround delay at each stage through the pathway.
- Replaces the 64 dword buffering requirement for STP initiators with the same 20 dword requriement that applies to SATA target devices in a symmetrical fashion.
- The same 20 dword turnaround delay constraint applies at each link interface along the pathway. This results in a requirement for each expander along the pathway to accept an additional 20 dwords from the upstream phy after sending SATA_HOLD.

**Suggested Changes**

Add definitions for downstream phy and upstream phy to section 3.1 Definitions as follows:

downstream phy: This refers to the next phy along the pathway of a connection between two SAS devices in the same direction as the direction frame transmission. The term is relevant only in contexts where the primary direction of frame transmission is clear.

upstream phy: This refers to the next phy along the pathway of a connection between two SAS devices in the direction opposite the direction of frame transmission. The term is relevant only in contexts where the primary direction of frame transmission is clear.

Replace clause 7.17.2 with the following:

Each expander device along the pathway of an STP connection between an STP initiator port and a SATA target device shall follow the SATA SATA_HOLD / SATA_HOLDA protocol to implement flow control across every link in the pathway along the connection. The requirement applies to transfers in either direction. If an expander device receives a SATA_HOLD from the downstream phy receiving a stream of dwords from the expander device the expander device shall stop sending data to the downstream phy and shall transmit SATA_HOLDA to the downstream phy. The number of dwords transmitted to the downstream phy between the time the downstream phy sends SATA_HOLD and the time the expander sends SATA_HOLDA to the downstream phy shall be no more than 20.

Upon receiving a SATA_HOLD from the downstream phy, an expander device shall propogate the SATA_HOLD to the next upstream phy along the pathway of the connection, and shall accept up to 20 more dwords from the upstream phy before receiving SATA_HOLDA from the upstream phy. If the expander device issues SATA_HOLDA to the downstream phy before receiving SATA_HOLDA from the next upstream phy along the pathway of the connection, the expander device shall maintain dwords received from the upstream phy during that interval in an internal buffer until the downstream phy is again ready to receive dwords. The expander may delay sending of SATA_HOLDA to the downstream phy to minimize the number of dwords it must store in an internal buffer if it can do so without exceeding the 20 dword SATA_HOLDA turnaround time on the link to the downstream phy.

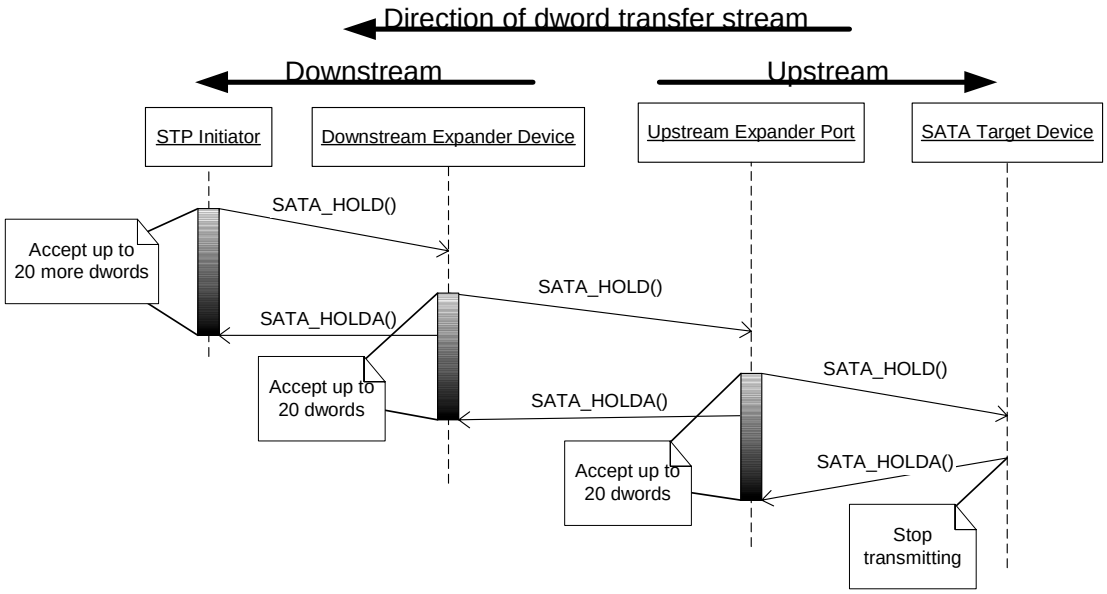


**Figure xx –** Expander device buffering & **HOLD/HOLDA** handshake (transfer in)

Figure xx shows the propogation of the SATA_HOLD primitive upstream through the pathway of an STP connection, the SATA_HOLDA reply of each expander along the pathway, and the interval during which each expander must accept incomming dwords. In this case the time to propogate the SATA_HOLD primitive to the next upstream phy is hidden behind the processing time to turnaround SATA_HOLDA to the downstream phy, so the expander need only accept up to 20 additional dwords to account for the SATA_HOLDA turnaround time of the next upstream phy.

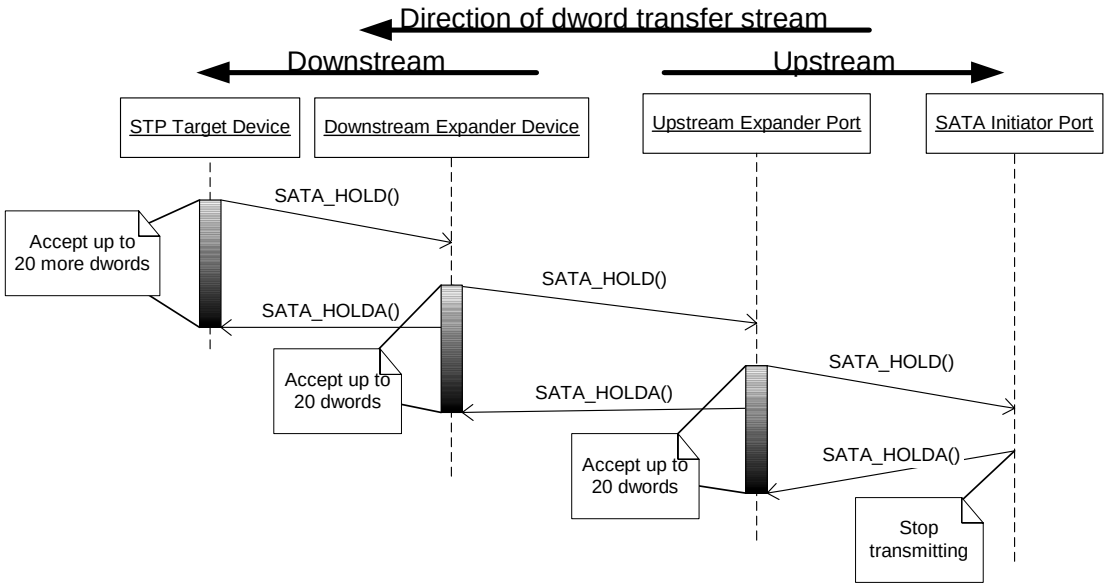


**Figure xx** – Expander device buffering & **HOLD/HOLDA** handshake (transfer out)

Figure xx shows the SATA_HOLD/SATA_HOLDA handshake and dword buffering for two stages of expander devices along the pathway of an STP connection transferring data from a SATA initiator port to a SATA target device. The flow-control protocol is completely symmetrical as compared to the transfer-in shown in the previous diagram.

Expander devices may use and shall recognize the SATA alternative of transmitting a SATA CONTINUE followed by scrambled data in leiu of repeated SATA_HOLD and SATA_HOLDA transmissions as defined by SATA.