To: T10 Technical Committee
From: Bob Sheffield (Robert.L.Sheffield@Intel.com), Intel Corporation
Date: November 4, 2002
Subject: T10/02-435r2, 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 20 dword buffering requirement for each expander along the pathway.

Suggested Changes

Replace clause 7.17.2 with the following:

Each expander device along the pathway of a STP connection between an STP initiator and a SATA target shall participate in the SATA_HOLD/SATA_HOLDA handshake according to the SATA 20 dword maximum turnaround delay requirement across every link along the pathway between initiator and target. Furthermore, each expander device along the pathway in an STP connection shall provide dword buffering to compensate for HOLD/HOLDA turnaround delay. The requirement to participate in the flow-control handshake and to provide buffering for incomming dwords on all phys along the pathway applies for transfers in both directions. An STP initiator must be able to accept up to 20 additional dwords after issuing the SATA_HOLD primitive to give the expander device enough time to suspend the data transfer and issue SATA_HOLDA. Expander devices providing pathways for STP connections shall be capable of buffering at least 20 dwords received on transfers in between the time it forwards SATA_HOLD to the next SAS phy along the pathway and the time it receives a SATA_HOLDA in response from the next SAS phy along the pathway.

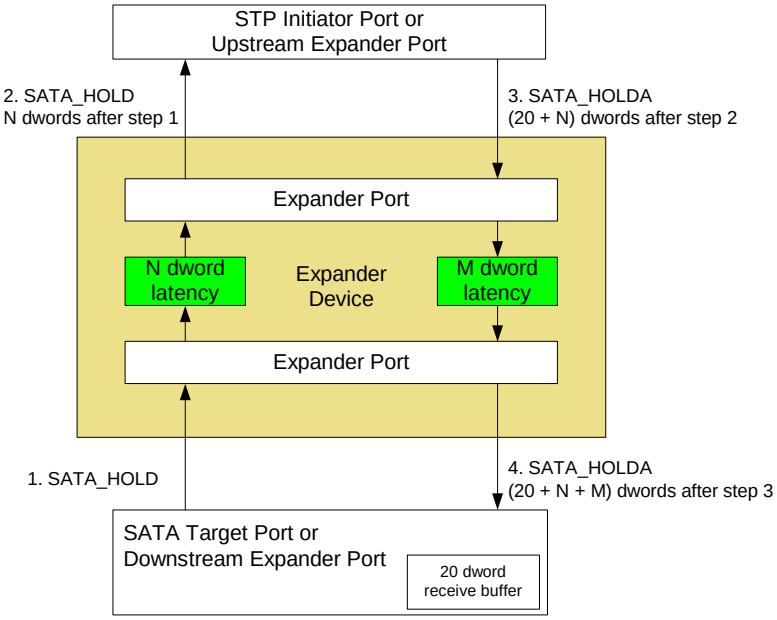

Figure 79 HOLD/HOLDA latency through an expander device

Figure 79 shows how an expander device that does not implement its own SATA_HOLD/SATA_HOLDA handshake with a SATA target port may cause the 20 dword receive buffer in the SATA target device to overflow on transfers to the SATA target device. Outbound data continues to stream through the expander device to the SATA target device while the expander device propogates the SATA_HOLD upstream, during the time the SATA initiator port takes to turnaround the SATA_HOLDA response, and the SATA_HOLDA to make the return trip to the SATA target device. The dwords transmitted during the added latency above and beyond that allocated for the SATA_HOLDA turnaround time

at the STP initiator port may overflow the 20-dword buffer in the SATA target device.

To solve this problem, each expander device shall generate its own SATA_HOLDA in response to a SATA_HOLD primitive received from a downstream phy along the STP connection pathway, and shall propogate the SATA_HOLD primitive to the next upstream phy along the STP connection pathway. After transmitting the SATA_HOLDA to the downstream phy, the expander device shall be capable of accepting at least 20 more dwords of data from the upstream phy before the upstream suspends the transfer and replies with SATA_HOLDA. If the expander introduces latency between the transmission of SATA_HOLDA to the downstream phy and forwarding SATA_HOLD to the upstream device, the expander shall be capable of buffering the number of additional dwords (beyond the 20 dwords for SATA_HOLDA turnaround) transferred during this delay.

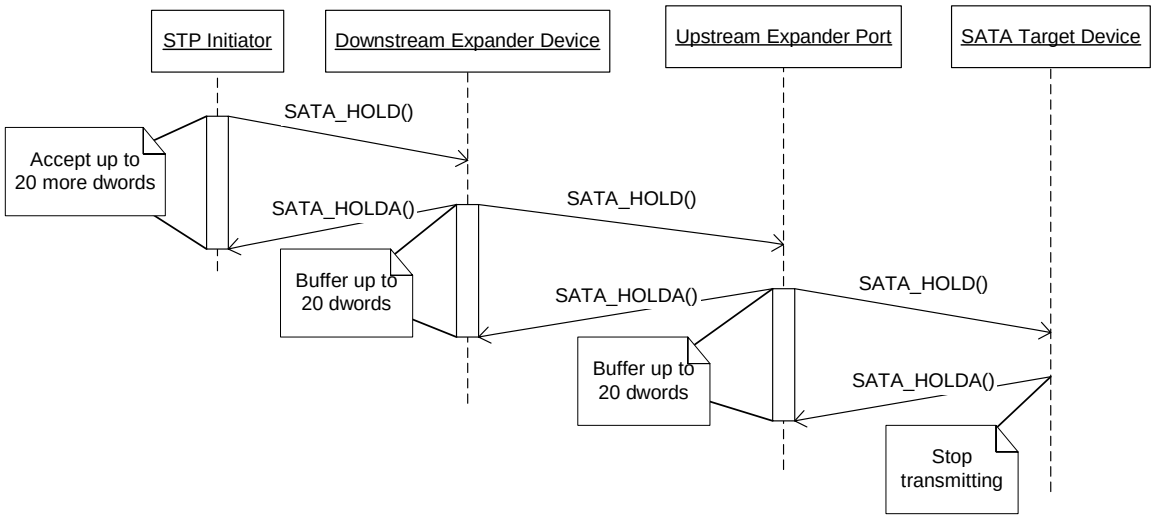

Figure xx – Expander device buffering & HOLD/HOLDA handshake (transfer in)

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 buffer 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 buffer up to 20 dwords to account for the SATA_HOLDA turnaround time of the next upstream phy.

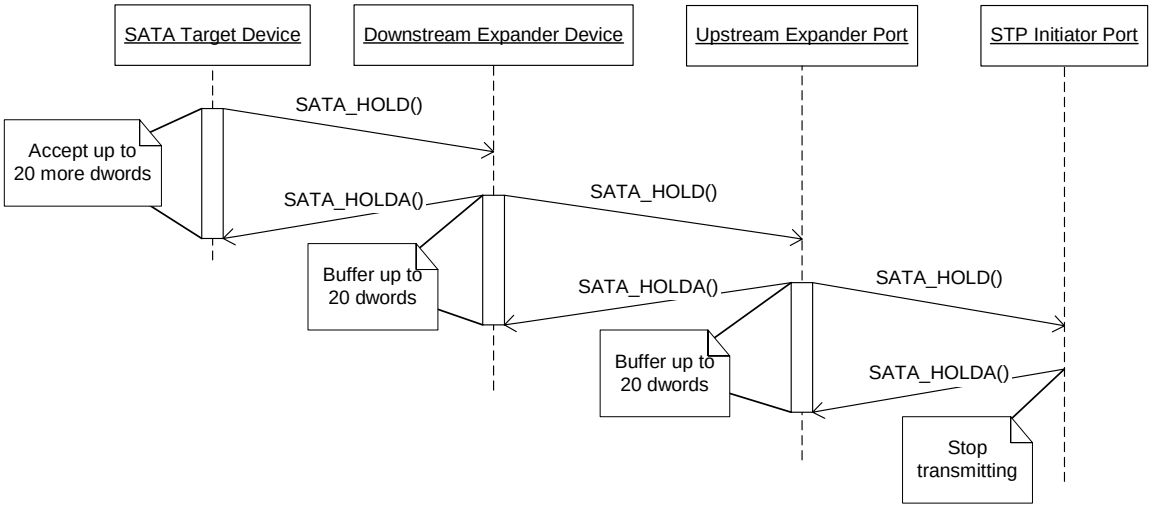

Figure xx – Expander device buffering & HOLD/HOLDA handshake (transfer out)

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.

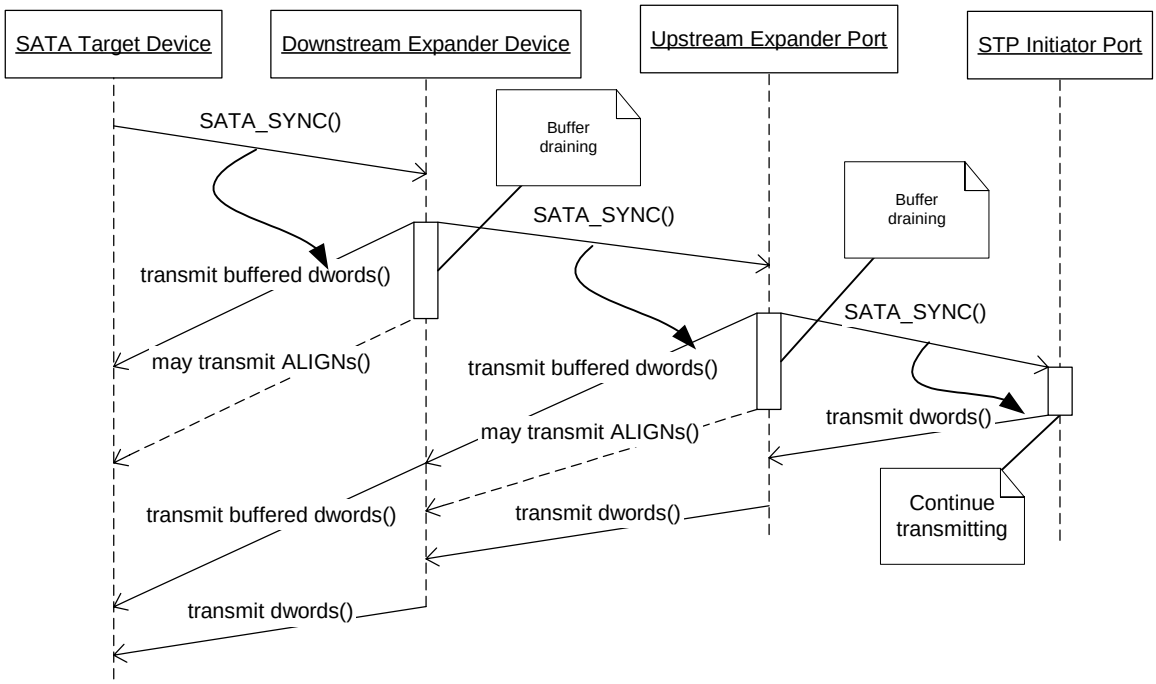

Figure xx - Continuing a transfer following SATA_HOLD/SATA_HOLDA

Figure xx shows the continuation of a transfer previously suspended via SATA_HOLD. Initially the upstream path shall be a continuous stream of SATA_HOLDA transmissions across every link from the SATA target device to the STP initiator port, and the downstream path shall be a continuous stream of SATA_HOLD transmissions across every link of the pathway from the STP initiator port to the SATA target device. When the SATA target device is again ready to accept data, it shall quit transmitting SATA_HOLD and shall transmit something else (such as SATA_SYNC). Each expander device shall continue to transmit SATA_HOLD to the upstream phy and SATA_HOLDA to the downstream phy until it receives something other than SATA_HOLD from the upstream phy. Upon receiving something other than SATA_HOLD from the upstream phy, an expander device shall begin transmitting data from its internal buffer to the downstream phy and shall forward the received dwords towards the upstream phy. If the expander device buffer becomes empty before receiving data from the upstream phy, the expander device shall fill the downstream dword stream with ALIGNs until it receives data from the upstream phy to forward to the downstream phy.