To: T10 Technical Committee
From: Bob Sheffield (Robert.L.Sheffield@intel.com)
Date: 24 August 2006
Subject: 06-386r0 SAS2: Store-and-Forward Expander Devices

Revision history
Revision 0 (24 August 2006) First revision

Related documents
SAS-2-r05a - Serial Attached SCSI - 2 (SAS-2) revision 05a

Overview

There is general consensus that because servers have a growing appetite for increased storage transfer bandwidth while it will be some years later before disk drives will need increased link speeds to maintain pace with the maximum sustained media transfer rates, a reasonable way for SAS-2 to satisfy the needs both for SAS initiator ports and SAS target ports is to provide some means of aggregating bandwidth through the service delivery subsystem. Such a solution would involve data transferred back and forth between expanders and target devices at 3Gb per second, and data transferred back and forth between expanders and initiator devices at 6Gb per second, without sacrificing link utilization on either end.

This proposal provides changes to the SAS-2 draft standard that define the means for an expander device to use store-and-forward (S&F) buffers and manage partial connections between S&F buffers and 3Gbps SAS target ports, and between S&F buffers and 6Gbps SAS initiator ports to achieve high link utilization with both SAS ports. On transfers from a 3Gb SAS port to a 6Gb SAS port, the frames transferred to the expander during a single connection at the 3Gb port would be transferred to a SAS initiator through multiple connections to the 6Gb port. Likewise, frames destined to a 3Gb SAS port received from a 6Gb SAS port would be transferred to the expander device through multiple connections with the 6Gb SAS port, and the same frames would be transferred to the 3Gb SAS port in a single connection.

An S&F expander device responds to an SSP or STP connection request from a SAS initiator port by allocating an S&F buffer pair (i.e., consisting of an S&F in buffer, and S&F out buffer, and an S&F buffer context) that acts as a proxy for the SAS target port, establishing a 6Gbps partial connection with the SAS initiator port. The same S&F buffer pair acts as a proxy for the SAS initiator port and establishes 3Gbps partial connections with the S&F target port identified in the connection request received from the SAS initiator port. Partial connections are opened and closed with the SAS initiator port and the SAS target port as needed depending on the flow of frames through the S&F buffer. S&F buffers are not used for SMP connections.

> NOTE 1 - The normative text is written to provide bandwidth aggregation in configurations where either the SAS target port or the SAS initator port supports the higher link rate; however, the discussion in the overview of this proposal is written with respect to the more likely case where SAS initiator ports support 6Gbps or 3Gbps link rates, and SAS target ports support either 3Gbps or 1.5 Gbps link rates.

In topologies with multiple levels of S&F-capable expander devices, an S&F expander device may appear at any level in the topology, but no more than one such S&F expander device along the pathway between a SAS initiator port and a SAS target port shall perform an S&F function for the scope of that I_T nexus.

This proposal defines the rules for opening and closing STP and SSP partial connections between S&F buffer pairs in an S&F expander device and SAS ports, and actions to be taken to handle error conditions affecting the partial connections and the I_T nexus. SMP partial connections are not needed, nor are they discussed in this proposal.

This proposal provides extensions to SMP commands to discover, enable, disable, and otherwise manage S&F expander capabilities.

This proposal does not define:

a) support for transport-layer retries in conjunction with S&F services;
b) elements required to support a SATA port multiplier in conjunction with S&F services; or
c) methods to limit the size of a SATA DATA FIS to limit S&F expander device buffering requirements.

S&F Buffer Pairs (still part of overview)

An S&F expander device allocates S&F buffer pairs. An S&F buffer pair is associated with a specific I_T nexus (identified by the initiator SAS address and target SAS address in an OPEN address frame), and consists of a buffer context, an S&F in buffer (for SSP frames or SATA frames received from a SAS target port and transmitted to a SAS initiator port), and an S&F out buffer (for SSP or SATA frames received from a SAS initiator port and transmitted to a SAS target port).

An S&F buffer context consists of:

a) the SAS identifier of the SAS initiator port associated with the I_T nexus of the S&F buffer pair;
b) the SAS identifier of the SAS target port associated with the I_T nexus of the S&F buffer pair;
c) a state machine variable indicating the protocol (i.e., SSP or STP) applicable to the S&F buffer pair;
d) the value of the CONNECTION RATE field from the last OPEN address frame received from the SAS initiator port, which is used to set the CONNECTION RATE field in OPEN address frames originated by the S&F buffer to establish partial connections to the SAS initiator port;
e) the value of the REMOTE CONNECTION RATE field from the last OPEN address frame received from the SAS initiator port, which is used to set the CONNECTION RATE field in OPEN address frames originated by the S&F buffer to establish partial connections to the SAS target port; and
f) links to S&F in and S&F out buffers that may be associated with the S&F buffer context .

S&F in buffers and S&F out buffers should be large enough to hold more than one SSP or SATA frame to achieve effective bandwidth aggregation. Systems may be able to limit the maximum size of a SATA frame to avoid excessive S&F expander memory requirements.

Editor's Note 1: The ATA SET MULTIPLE MODE command may be used to limit the maximum size of a DATA FIS transmitted or received by a SATA device for an ATA READ MULTIPLE command, an ATA READ MULTIPLE EXT command, an ATA WRITE MULTIPLE command, an ATA WRITE MULTIPLE EXT command, and an ATA WRITE MULTIPLE FUA EXT command. Other means to limit the maximum size of a DATA FIS would be vendor-specific, or would require SATA protocol extensions not currently defined in SATA 2.5.

In some cases, an S&F buffer context may exist without the associated S&F in buffer and S&F out buffer. For example, if an SSP target device is processing a READ command, but neither the SSP initiator port nor the SSP target port has any frames to transmit for several milliseconds, the S&F expander need not maintain active in and out buffers for that I_T nexus. However when the target reestablishes a partial connection to the S&F buffer, the S&F expander device allocates a buffer for the context and extends R_RDY credit to the SSP taget port. When the SSP target port transmits enough frames to fill the S&F in buffer, the S&F buffer needs to establish a partial connection with the SSP initiator port to start draining the S&F in buffer. The S&F buffer needs to know the connection rate to request in the OPEN address frame. This information is not available from the OPEN address frame received from the target since the target uses a connection rate of 3Gbps or less. One of the values maintained in the S&F buffer context is the connection rate from the last OPEN address frame received from the SSP initiator port for that I_T nexus.

It is vendor-specific how long an S&F expander device maintains the S&F buffer context, and the S&F expander device is not required to analyze the transport-layer protocol to determine when the target has discarded the context of a given tag. An S&F expander device may use a least-recently used (LRU) algorithm to manage a limited pool of S&F buffer contexts available in the S&F expander device. If an S&F expander device receives an OPEN address frame from a target device and it does not have an S&F buffer context for the I_T nexus, the S&F expander shall create an S&F buffer context for the connection request and shall forward the OPEN address frame to the SAS initiator port with a connection rate equal to the negotiated physical link rate of the immediate phy to which the OPEN address frame is to be forwarded. If the partial connection request is rejected because of the connection rate, OPEN_REJECT (CONNECTION RATE NOT SUPPORTED), the S&F buffer shall retry the partial connection request using a lower connection rate.

NOTE 2 - Subsequent descriptions discuss different behavior for a SAS initiator port compared to a SAS target port, assuming the target port is running at a slower link rate. But the normative text later in the proposal is symmetric so that if the initiator port has a lower link rate, it works as well, with the roles reversed.

SSP Connection Management (still part of overview)

Connection management for SSP data-out transfers works as follows:

a) An S&F expander device intercepts connection requests from SAS initiator ports.
b) If there is an S&F context for the specified I_T nexus, the S&F expander device accepts the connection request, the S&F buffer acts as a proxy for the SAS target port, grants R_RDY frame credit and receives frames from the S&F initiator.
c) If the S&F expander has no context for the I_T nexus, it creates a context and forwards the connection request to the SAS target port. If the connection request to the SAS target port is rejected, the S&F buffer likewise rejects the connection request from the SAS initiator port and discards the context for the I_T nexus.
d) If the SAS target port accepts the connection request, S&F in and out buffers are allocated for the context for the I_T nexus and SSP frames begin to flow from the SSP initiator port, into the S&F out buffer, and to the SAS target port.
e) The S&F buffer may close the connection with the SAS target port while waiting for frames from the SSP initiator port, but normally the connection to the SAS target port will be open while there is a connection open anytime there is a connection with the SAS initiator. It is expected, however, that often the S&F buffer will be connected to the SAS target port while not connected to the SAS initiator port.
f) When frames are available in the S&F buffer to transmit to a SAS target device, the S&F buffer, forwards frames to the SAS target port.
g) If the S&F buffer fills to the point where it cannot accept another frame, it withholds R_RDY frame credit to the SAS initiator port and transmits CREDIT BLOCKED to the SAS initiator port.
h) In response, the SAS initator port initiates closing the connection with the normal exchange of DONE primitives (DONE (CREDIT TIMEOUT) should be used) and exchanging of CLOSE primitives.
i) The S&F buffer continues transmitting frames to the SAS target port.
j) The SAS initiator port uses the time to establish a partial connection with and transmit frames to a different SAS target port through a different S&F buffer in the same or a different S&F expander device.
k) After transmitting a quantity of data destined to a different SAS target port sufficient to fill another S&F buffer, the S&F initiator port retries the connection request to the first SAS target port. If the S&F buffer for that I_T nexus has an SSP frame or less remaining to transmit to the SAS target port, the S&F buffer accepts the connection request from the SAS initiator port and resumes filling the S&F buffer with SSP frames received from the SAS initiator port.
l) The process continues until the SAS initiator port has no more frames to transmit to the target and closes the partial connection with the S&F buffer. When the S&F buffer has no more frames to transmit to the SAS target port, it closes the partial connection to the SAS target port, and may or may not discard the context for the I_T nexus. Preserving the context of the I_T nexus will facilitate the process if and when the SAS target port reconnects to transmit frames, through the S&F buffer, to the SAS initiator port.

Connection management for SSP data-in transfers work as follows:

a) An S&F expander device establishes an S&F buffer pair and context for the I_T nexus when the SAS initiator port establishes a connection to transmit a COMMAND frame as described above.
b) Both partial connections are closed while the SAS target fetches data, and the S&F expander device may discard the actual S&F in and out buffers, but maintains the context of the I_T nexus.
c) When the SAS target port originates a connection request to transfer SSP frames, the S&F expander device accepts the connection request on behalf of the SAS initiator port, attaches an S&F in buffer to the context of the I_T nexus and thereby establishes a partial connection with the SAS target port.
d) The SAS target port transmits SSP frames which the S&F expander device accumulates in the S&F in buffer.
e) When the S&F buffer is full, it originates a connection request to the SAS initiator port and begins transferring frames to the SAS initiator port.
f) If a connection cannot be established with the SAS initiator port, the S&F buffer transmits CREDIT BLOCKED to the SAS target port and the partial connection to the SAS target port closes until the S&F buffer can reconnect to the SAS initiator port and begin draining frames from the S&F in buffer.

g) Meanwhile the SAS target port attempts to reestablish a connection by sending connection requests, and the S&F buffer rejects the connection requests with OPEN_REJECT (RETRY) until there is space to receive at least one frame in the S&F in buffer.
h) When the S&F buffer no longer has a full frame to transmit to the SAS initiator port, the S&F buffer closes the partial connection to the SAS initiator port exchanging DONE (NORMAL) and then CLOSE with the SAS initiator port.
i) If both the S&F in buffer and S&F out buffer are empty and there are no active or pending partial connections with either the SAS initiator port or the SAS target port, the S&F expander may discard the buffer space associated with the context for the I_T nexus.
j) The S&F expander may maintain the context for the I_T nexus for a vendor-specific period of time of inactivity before discarding it (i.e., the S&F expander need not monitor the application-level state of individual I_T_L_Q nexuses to determine when a SAS target is no longer processing tasks for a given SAS initiator port).

S&F buffer pairs in S&F expander devices operate in full-duplex mode. COMMAND, DATA out, and TASK frames may be flowing through the S&F out buffer at the same time XFER_RDY, DATA in, and RESPONSE frames are flowing through the S&F in buffer.

An S&F buffer pair originates a partial connection request to a SAS initiator port when the S&F in buffer is full. An S&F buffer pair originates a partial connection request to a SAS target port as soon as it has a frame available in the S&F out buffer to transmit to a target.

An S&F buffer pair originates closing a partial connection to a SAS initiator only when the S&F in buffer is empty and the S&F out buffer is full. An S&F buffer pair originates closing a connection to a SAS target port only when the S&F in buffer is full and the S&F out buffer is empty.

STP Connection Management (still part of overview)

STP S&F buffer pairs are used only to accumulate SATA frames containing DATA FISes. Otherwise, protocol between an STP initiator port and an STP target port is interlocked as in unbuffered STP protocol. This avoids a race condition where an STP initiator port and an STP target port both transmit X_RDY at the same time.

Connection management for STP data-out transfers works as follows:

a) An S&F expander device intercepts STP connection requests from SAS initiator ports.
b) If there is an S&F context for the specified I_T nexus and there is a partial connection between the S&F expander device and the STP target port, the S&F expander device accepts the connection request from the STP initiator port, and the S&F buffer acts as a proxy for the STP target port.
c) If the S&F expander has no context for the I_T nexus, it creates a context and forwards the connection request to the STP target port. If the connection request to the STP target port is rejected, the S&F buffer rejects the connection request from the STP initiator port and discards the context for the I_T nexus.
d) If the STP target port accepts the connection request, S&F in and out buffers are allocated for the context for the I_T nexus, the S&F buffer receives SATA_X_RDY from the STP initiator port, forwards the SATA X_RDY to the STP target port, receives SATA_R_RDY from the STP target port, forwards the SATA_R_RDY to the STP initiator port, and begins receiving one or more SATA frames from the STP initiator port into the S&F out buffer.
e) If the last SATA frame the S&F buffer receives from the STP initiator port does not contain a DATA FIS, then the S&F expander shall not return SATA_R_OK until:
   1) the last SATA frame (i.e., the one that does not contain a DATA FIS) is transmitted to the STP target port; and
   2) the S&F buffer has received SATA_R_OK from the STP target port.

   If the S&F buffer receives a SATA_X_RDY from the STP target port before transmitting the last SATA frame in the S&F out buffer to the STP target port, then:
   1) the S&F buffer shall return SATA_R_ERR to the STP initiator port instead of the SATA_R_OK; and
   2) discard remaining SATA frames in the S&F out buffer.

   The S&F buffer shall then forward the SATA_X_RDY received from the STP target port to the STP initiator port and continue processing according to the STP data-in transfer process (see below).
f) If the S&F buffer receives a SATA frame from the STP initiator port that contains a DATA FIS and there is enough space in the S&F out buffer to receive another SATA frame with the maximum

payload, the S&F buffer may accept another frame by originating a SATA_R_RDY transmitted to the STP initiator port in response to receiving a SATA_X_RDY from the STP initiator port.

g) As SATA frames are received from the STP initiator port, the S&F buffer forwards those frames to the STP target port. STP data frames accumulate in the S&F out buffer.
h) The S&F buffer may close the connection with the STP target port while waiting for a SATA frame frame from the STP initiator port.
i) The S&F buffer shall not close the partial connection with the STP target port if it has transmitted an SOF for a SATA frame frame for which it has not yet received and EOF from the STP iniatiator port. Normally the connection to the STP target port remains open while there is a connection with the STP initiator port. It is expected, however, that often the S&F buffer will be connected to the STP target port while not connected to the STP initiator port.
j) When a SATA frame or part of a SATA frame frame is available in the S&F out buffer to transmit to an STP target port, the S&F buffer establishes a partial connection with the STP target port if one is not already established, and forwards dwords to the STP target port.
k) If the S&F buffer receives EOF for a SATA frame and the S&F buffer is full to the point where it cannot accept another complete SATA frame, after SATA_SYNC has been exchanged in both directions the S&F buffer shall transmit CLOSE to the STP initiator port.
l) The STP initator port completes closing the connection by transmitting CLOSE.

NOTE 3 - An S&F aware STP initiator port should honor closing the connection even if it has a SATA frame frame ready to transmit.

m) The S&F buffer continues transmitting the remaining SATA frame(s) to the STP target port.
n) The STP initiator port uses the time to establish a partial connection with and transmit frames to a different STP target port through a different S&F buffer in the same or a different S&F expander device.
o) After transmitting one or more complete SATA frames destined to different STP target port(s) through different S&F buffer(s), the STP initiator port originates a new connection request to the first STP target port. If the S&F buffer for that I_T nexus has one or fewer SATA frames remaining to transmit to the STP target port, the S&F buffer accepts the connection request from the STP initiator port and resumes filling the S&F buffer with dwords for the next SATA frame received from the STP initiator port.
p) Before the S&F out buffer fills to the point of overflow, it shall transmit SATA_HOLD to the STP initiator port, and the partial connection with the STP initiator port remains open. The S&F buffer pair resumes the transfer by transmitting SATA_R_IP to the STP initaitor port.
q) The process continues until the STP initiator port has no more frames to transmit to the target and closes the partial connection with the S&F buffer. When the S&F buffer has no more SATA frames to transmit to the STP target port, it closes the partial connection to the SAS target port, and may or may not discard the context for the I_T nexus. Preserving the context of the I_T nexus will facilitate the process if and when the STP target port reconnects to transmit SATA frames, through the S&F buffer, to the STP initiator port.

Connection management for STP data-in transfers work as follows:

a) An S&F expander device establishes an S&F buffer pair and context for the I_T nexus when the STP initiator port establishes a connection to transmit a SATA frame frame containing an ATA COMMAND FIS as described above.
b) Both partial connections should be closed while the STP target fetches data, and the S&F expander device may discard the empty S&F in and out buffers, but should maintain the S&F buffer context for the I_T nexus.
c) When the STP target port originates a connection request to transmit a SATA frame or other SATA protocol dwords, the S&F expander device accepts the connection request on behalf of the STP initiator port, attaches an S&F in buffer to the context of the I_T nexus and thereby establishes a partial connection with the STP target port.
d) The STP target port transmits SATA_X_RDY to request transmission of a SATA frame.
e) The S&F buffer pair responds by transmitting SATA_R_RDY to the STP target port; however, the S&F buffer pair shall not transmit SATA_R_RDY unless there is enough space in the S&F in buffer to accept the largest possible SATA frame that does not contain a DATA FIS.

NOTE 4 - This is because SATA only permits use of HOLD for flow control during transmissoin of a SATA frame containing a DATA FIS. Without this restriction the S&F in buffer could overflow.

f) The STP target port begins transferring dwords for a SATA frame, which the S&F expander device accumulates in the S&F in buffer.
g) When the S&F buffer is full, it originates a connection request to the STP initiator port, transmits SATA_X_RDY, receives SATA_R_RDY, and begins transferring the SATA frame to the STP initiator port.
h) If a connection cannot be established with the STP initiator port and the S&F in buffer is full, the S&F buffer transmits SATA_HOLD to the SAS target port until it successfuly establishes a connection with the STP initiator port to begin transmission of the SATA frame to the STP initiator port.

NOTE 5 - This should only happen on a SATA DATA frame because, per step (e), the S&F buffer allowed enough space to receive any other frame before transmitting SATA_R_RDY.

i) When the S&F buffer has transmitted EOF for a SATA frame and no longer has a full SATA frame available to transmit to the STP initiator port, the S&F buffer originates closing of the partial connection to the STP initiator port (i.e., sends CLOSE (NORMAL)).
j) If both the S&F in buffer and S&F out buffer are empty and there are no active or pending partial connections with either the STP initiator port or the STP target port, the S&F expander may discard the buffer space associated with the context for the I_T nexus.
k) The S&F expander may maintain the context for the I_T nexus for a vendor-specific period of time of inactivity before discarding it (i.e., the S&F expander need not monitor the application-level state of individual ATA command contexts to determine when an STP target is no longer processing commands for a given STP initiator port).

Resolving an NCQ Collision (still part of overview)

When using NCQ, an STP initiator may transmit a SATA frame containing a Register - Host to Device FIS (i.e., a command FIS, type 27h), while at the same time a device transmits a SATA frame containing a data FIS (i.e., FIS type 46h). SATA defines this as a collision which must be resolved by sending a SATA_R_ERR primitive instead of a SATA_R_OK primitive in response to the Register - Host to Device FIS.

If the S&F buffer receives any FIS other than a DATA FIS (i.e., FIS type 46h) from an STP initiator port, the S&F buffer shall not return R_OK until the FIS has been delivered to the STP target port. If the STP target port transfers a FIS to the S&F buffer before the non-data FIS has been delivered to the STP target port, the S&F buffer shall discard the non-data FIS received from the STP initiator port and return SATA_R_ERR to the STP initiator port instead of R_OK.

Connection Rates (still part of overview)

An S&F capable expander device activates it's S&F capability for an I_T nexus only if it receives an OPEN address frame from a SAS initiator port with a non-zero value in the REMOTE CONNECTION RATE field. If the REMOTE CONNECTION RATE field is zero, the S&F expander device disables its S&F function for the specified I_T nexus. This means legacy (non-S&F-aware) SAS initiator ports are not able to use the bandwidth aggregation capabilities in S&F expander devices.

A SAS initiator port requesting S&F service for a connection to a SAS target port sets the REMOTE CONNECTION RATE field to the negoitiated physical link rate supported by a potential pathway from the S&F expander device that provides the S&F service to the SAS target port, and sets the CONNECTION RATE field to a negotiated physical link rate supported by a potential pathway from the SAS initiator port to the S&F expander device that provides the S&F service. The connection rate specified in the CONNECTION RATE field is power-of-two multiple of the connection rate specified in the REMOTE CONNECTION RATE field.

When the S&F expander device acting as a proxy for the SAS initiator port forwards a partial connection request to the SAS target port, it replaces the CONNECTION RATE field in the OPEN address frame with the value received in the REMOTE CONNECTION RATE field.

The REMOTE CONNECTION RATE field in the OPEN address frame from the SAS target port will be set to zero. If the S&F expander device has an active context for the I_T nexus, it shall enable the S&F service for the

connection, and accept the connection request, thereby establishing a partial connection between the S&F buffer and the SAS target port. The S&F buffer pair accumulates frames received from the SAS target port in the S&F in buffer. When it is full, the S&F buffer pair uses information stored in the S&F buffer context to set the value in the CONNECTION RATE field in the OPEN address frame sent to the SAS initiator port. When an S&F expander device originates or forwards an OPEN address frame to a SAS initiator port, it sets the REMOTE CONNECTION RATE field to a value of 1h (DISABLED). If an S&F expander device receives an OPEN address frame from a SAS target port with the REMOTE CONNECTION RATE field set to 1h, it shall disable its S&F function for that I_T nexus. This assures that only the S&F expander device closest to the SAS target port provides the S&F services.

Cascaded S&F Expander Devices (still part of overview)

S&F expander devices may be cascaded and intermixed with non-S&F expander devices. Only one S&F expander device along a potential pathway from a SAS initiator port to a SAS target port shall provide an S&F buffer pair for the I_T nexus. An SMP initiator port may disable the S&F function of an S&F expander device on either a device basis or on a phy-by-phy basis. If an S&F expander device has it's S&F function disabled for a phy in the potential pathway from a SAS initiator port to a SAS target port, it shall handle connections as a an expander device without S&F capability. From this point forward in the discussion, references to an S&F expander device mean an S&F expander device which does not have its S&F function disabled (for the I_T nexus of interest) by an SMP command received from an SMP initiator port.

When an S&F expander device receives an OPEN address frame with the INITIATOR PORT bit set to one, it shall return AIP (SFT TAKEN) to indicate that, unless subsequently overridden, that S&F expander device will provide the S&F buffer pair for the I_T nexus identified by the SOURCE SAS ADDRESS and DESTINATION SAS ADDRESS in the OPEN address frame. SFT is an acronym for store-and-forward token, and AIP (SFT TAKEN) is the primitive from SAS-1.1 referred to as AIP (RESERVED WAITING ON PARTIAL). AIP (SFT TAKEN) is transmitted and detected as an extended primitive sequence.

If an S&F expander device originates or forwards an OPEN address frame along a potential pathway towards a SAS target port and receives AIP (SFT TAKEN) in response, that S&F expander shall not provide an S&F buffer pair for the I_T nexus (i.e., it was overridden by an S&F expander device closer to the SAS target port).

If an S&F expander receives an OPEN address frame originating from a SAS initiator port and the S&F expander does not have an S&F buffer context for the specified I_T nexus, the S&F expander shall not return OPEN ACCEPT until it has received OPEN ACCEPT from the phy to which it forwarded the OPEN address frame.

NOTE 6 - An S&F expander device should never originate a connection request destined to a SAS target port and receive an AIP (SFT TAKEN) in reply because it wouldn't have established an S&F buffer pair for the I_T nexus if it has forwarded an earlier connection request from the SAS initiator port to that target and received AIP (SFT TAKEN) in response.

When an S&F expander device forwards a conneciton request received from a SAS initiator port to a SAS target port and the S&F expander device does not receive AIP (SFT TAKEN) before the SAS target port accepts the connection request, the S&F expander device shall create an S&F buffer pair for the I_T nexus including an S&F buffer context, and S&F in buffer, and an S&F out buffer. The S&F buffer shall accept the connection request fom the SAS initiator port on behalf of the SAS target port.

When an S&F expander device receives an OPEN address frame with the INITIATOR PORT bit set to zero (i.e., originated from a SAS target port), the S&F expander may or may not have an S&F context for the specified I_T nexus. If the value of the remote connection rate field in the OPEN address frame received from the SAS target port is set to 1h, the S&F expander device shall disable its S&F buffer service for that I_T nexus (i.e., an S&F expander closer to the SAS target port has already claimed responsibility for providing the S&F buffer service), and the S&F expander shall process the OPEN address frame as would an expander device that does not support the S&F capability.

If the S&F expander device receives an OPEN address frame originated from a SAS target port with the REMOTE CONNECTION RATE set to zero, the S&F expander device shall provide the S&F buffering service for the I_T nexus. If the S&F expander device already has an active S&F buffer context for the I_T nexus, it shall accept the connection request on behalf of the SAS initiator port and begin receiving frames from the SAS

target port in the S&F in buffer. If the S&F expander device does not have an active S&F buffer context for the I_T nexus, it shall create one and forward the connection request to the S&F initiator with the REMOTE CONNECTION RATE field set to 1h (DISABLE), and the CONNECTION RATE field set to the negotiated physical link rate of the phy to which it transmits the OPEN address frame. The S&F expander device shall not accept frames from the SAS target port (i.e. shall not send R_RDY credit) until the connection has been accepted by the SAS initiator port.

If the S&F expander device receives an OPEN address frame originating from a SAS target port (i.e., the INITIATOR PORT bit is set to zero) with the REMOTE CONNECTION RATE field set to a valid connection rate, and the S&F expander device does not have an active S&F buffer context for the I_T nexus, then the S&F expander device shall provide the S&F buffering service for the I_T nexus and forward the OPEN address frame along the partial pathway to the SAS initiator port, copying the value from the REMOTE CONNECTION RATE field to the CONNECTION RATE field and setting the REMOTE CONNECTION RATE field to a value of 1h (DISABLED).

SMP Support (still part of overview)

There are several changes proposed to SMP commands to support S&F expander devices. These changes allow discovery of S&F buffering capability, enable/disable of S&F buffer service, query of S&F buffer status, and commands to manage S&F buffer context for error handling. S&F buffering is not used for SMP connections.

SSP Error Handling (still part of overview)

There are a number of new potential error conditions to deal with, for example a stranded RESPONSE frame from a target that has closed the connection while the S&F expander device still waits for an opportunity to reconnect and transfer the RESPONSE frame to the initiator. This proposal defines the mechanisms to detect and recover from such conditions.

In terms of pecedent for this mode of operation it is worth noting a case in fibre channel that works similarly. In many ways, SAS protocol emulates the protocol evident in fibre channel arbitrated loop (FCAL). Once a connection is established between an initiator port and target port on the same loop, operation is as though the initiator port and target port are linked point-to-point. Fibre channel also supports a topology where an FC initiator and an FC target are on independent loops, with each loop connected through an FL-port to a common fabric. Initiator and target, though following FCAL protocol locally, are still able to process exchanges efficiently and effectively through an intermediate switched fabric. Loop tenancies are managed independently by the FL ports on opposite ends of the FC fabric. The intent is to model the use SAS S&F expander devices similarly, with loop-tenancies corresponding to connections, and with the SnF expander device performing the frame buffering function of the FC switched fabric, while making simplifications everywhere possible.

STP Error Handling (still part of overview)

Most STP error conditions are handled the same way they are handled without S&F buffering because SATA protocol interlock is maintained (i.e., buffering is not used) for all SATA frames except those containing a DATA FIS. Handling of errors where a SATA frame containing a DATA FIS are described below.

Forwarded DATA FIS returns SATA_R_ERR

Effective bandwidth aggregation requires that the S&F buffer hold at least one complete DATA FIS in-transit. and begin receipt of the next DATA FIS before transmitting the previous DATA FIS to the target. To accomplish this, the S&F expander must return SATA_R_OK to the sender, exchange SATA_SYNC, receive SATA_X_RDY and transmit SATA_R_RDY, before forwarding the held FIS to the destination. The destination may, then, experience an error in receiveing the forwarded FIS (e.g., a bit error causing a CRC mismatch), and would then return SATA_R_ERR to the S&F buffer in response. At that point, though, the STP port that sent the DATA FIS has received SATA_R_OK for the DATA FIS and, incorrectly, assumes that the destination received the FIS without error.

On a transfer from an STP target port to an STP initiator port (i.e., READ), the S&F expander continues to accept SATA frames from the STP target port and forwards them to the STP initiator port. The STP initiator port has detected the error, should abort outstanding commands, and retry the affected operations.

If an STP target port returns SATA_R_ERR after receiving a SATA frame containing a DATA FIS, the S&F buffer discards the SATA_R_ERR and continues forwarding SATA DATA frames received from the STP initiator port to the STP target port. The SATA device detects the error and, depending on whether the affected commands are queued or not, either returns a Set Device Bits FIS (i.e. in the queued case) or a Register Device to Host FIS (i.e., the non-queued case) to the STP initiator. If the affected command is a queued command, the the Set Device Bits FIS has the appropriate error bit set in the ERR field set to one, and the bit in the SActive field for any commands aborted by the error set to one (i.e., causing the corresponding bit in the STP initiator's SActive shadow register to be set to zero). The STP initiator port should discard any frames received for the affected command(s).

SATA frame received after forwarding a SATA frame containing a PIO SETUP FIS doesn't match

If the S&F buffer forwards a SATA frame containing a PIO SETUP FIS to the destination, then the next FIS received from there must match the PIO SETUP FIS. It might be a DATA FIS, so S&F buffer can't hold-off R_OK until it's forwarded to the originator of the PIO SETUP FIS, so the S&F buffer must perform this checking. Anytime a PIO SETUP FIS is forwarded, the S&F buffer stores the relevant fields in the S&F buffer context, and checks the next frame received against those values for consistency, returning R_ERROR if it doesn't match.

Receive a SATA frame frame containing an unsupported FIS type

SATA protocol requires a response of SATA_R_ERR for an unsupported FIS type. The problem is, different devices support different FIS types. An S&F buffer handles this case two ways:

a) An SMP command is defined to provide a list of valid FIS types. The S&F expander shall respond with SATA_R_ERR if it receieves a FIS with a type that does not match any of the types provided in the list.
b) The SATL shall hold off the SATA_R_OK or SATA_R_ERR reply for any received FIS type other than DATA until it's forwarded to the destination and the destination provides the response.

Sync escape

While an S&F buffer is transmitting a SATA frame (containing a DATA FIS?), if a SATA_SYNC is received, the S&F buffer shall stop transmitting the frame and send SATA_SYNC, thereby terminating transmission of the SATA frame.

STP Handling Affiliations (still part of overview)

The S&F buffer, acting as a proxy for the STP/SATA bridge port, maintains a state machine variable indicating whether the STP target port of the STP/SATA bridge should be affiliated or not at a given frame boundary. After transmitting the last frame received from an STP initiator prior to receiving a CLOSE (CLEAR AFFILIATION), the S&F expander transmits a CLOSE (CLEAR AFFILIATION) to the STP/SATA bridge. If there is no connection to the STP/SATA bridge and there are no frames in the buffer when an STP initiator sends a CLOSE (CLEAR AFFILIATION), the S&F expander device must open a connection to the STP/SATA bridge and close it with CLOSE (CLEAR AFFILIATION) to clear the affiliation. Any other time, the S&F buffer transmits CLOSE (NORMAL).

Suggested changes

Editor's Note 2: There are many changes yet forthcoming to address everywhere S&F buffers impact the existing text, but most of the key changes are identified here.

Add the following definitions to subclause 3.1 Definitions:

3.1.x partial connection: A temporary association between an S&F buffer in an S&F expander device and a SAS port, in which the S&F expander device performs the link-layer functions on behalf of a another SAS port (i.e., as though the connection were between two SAS ports).

3.1.x remote connection rate: The effective rate of dwords through the partial pathway between an S&F expander phy and a SAS phy that is in a SAS port having the SAS address specified in the destination SAS address of an OPEN address frame originating from a SAS phy (i.e., originating from a SAS phy that is in the SAS port having its SAS address specified in the source SAS address field of the OPEN address frame).

3.1.x S&F buffer: A buffer in an S&F expander device used to accumulate or source SSP or SATA frames received by or transmitted from an S&F expander port through partial connections.

3.1.x S&F expander device: An expander device that uses store-and-forward buffers (see 3.1.x) to move data from a receiving expander port to the transmitting expander port asynchronously through the use of partial connections.

3.1.x S&F buffer pair: An S&F in buffer and an S&F out buffer associated with the same I_T nexus. An S&F expander shall not use an S&F buffer for STP partial connecitons and SSP partial connections at the same time. If x and y are two SAS addresses, then the S&F buffer pair identified by an initiator port SAS address x and a target port SAS address y shall be a different S&F buffer pair than the S&F buffer pair identified by initiator port SAS address y and a target port SAS address x (i.e., when a SAS initiator and a SAS target reverse roles it represents a different I_T nexus and used a different S&F buffer pair).

3.1.x S&F expander port: An S&F expander device object that interfaces to the service delivery subsystem and to SAS ports in other devices. See 4.6.2.

3.1.x S&F in buffer: An S&F buffer that receives dwords from a SAS target port transmits dwords to a SAS initiator port.

3.1.x S&F out buffer: An S&F buffer that receives dwords from a SAS initiator port transmits dwords to a SAS target port.

3.1.x SATA frame: A unit of information exchanged between an ATA host adapter and an ATA device. A SATA frame consists of a SATA_SOF primitive, a Frame Information Structure (see...), a CRC calculated over the contents of the FIS, and a SATA_EOF primitive.

3.1.x store-and-forward (S&F): A technique in which frames are sent to an intermediate expander device where they are held in a buffer (i.e., an S&F buffer) and sent at a later time to the destination SAS port or to another expander port.

3.1.x store-and-forward buffer manager (SBM): An object within an expander function that manages allocation and deallocation of store-and-forward buffers and partial connections.

Modify the following definitions in subclause 3.1 Definitions:

3.1.31 connection rate: The effective rate of dwords through the pathway between a SAS initiator phy and a SAS target phy, established through the connection request or between a SAS phy and an S&F buffer in an S&F expander device.

Add the following to the symbols and abbreviations in subclause 3.2:

S&F	store-and-forward (see 3.1.x)
SBM	store-and-forward buffer manager (see 3.1.x)

Modify figure 10 as shown to add S&F capability:

Figure 10 defines the notation used for aggregation relationships between classes.

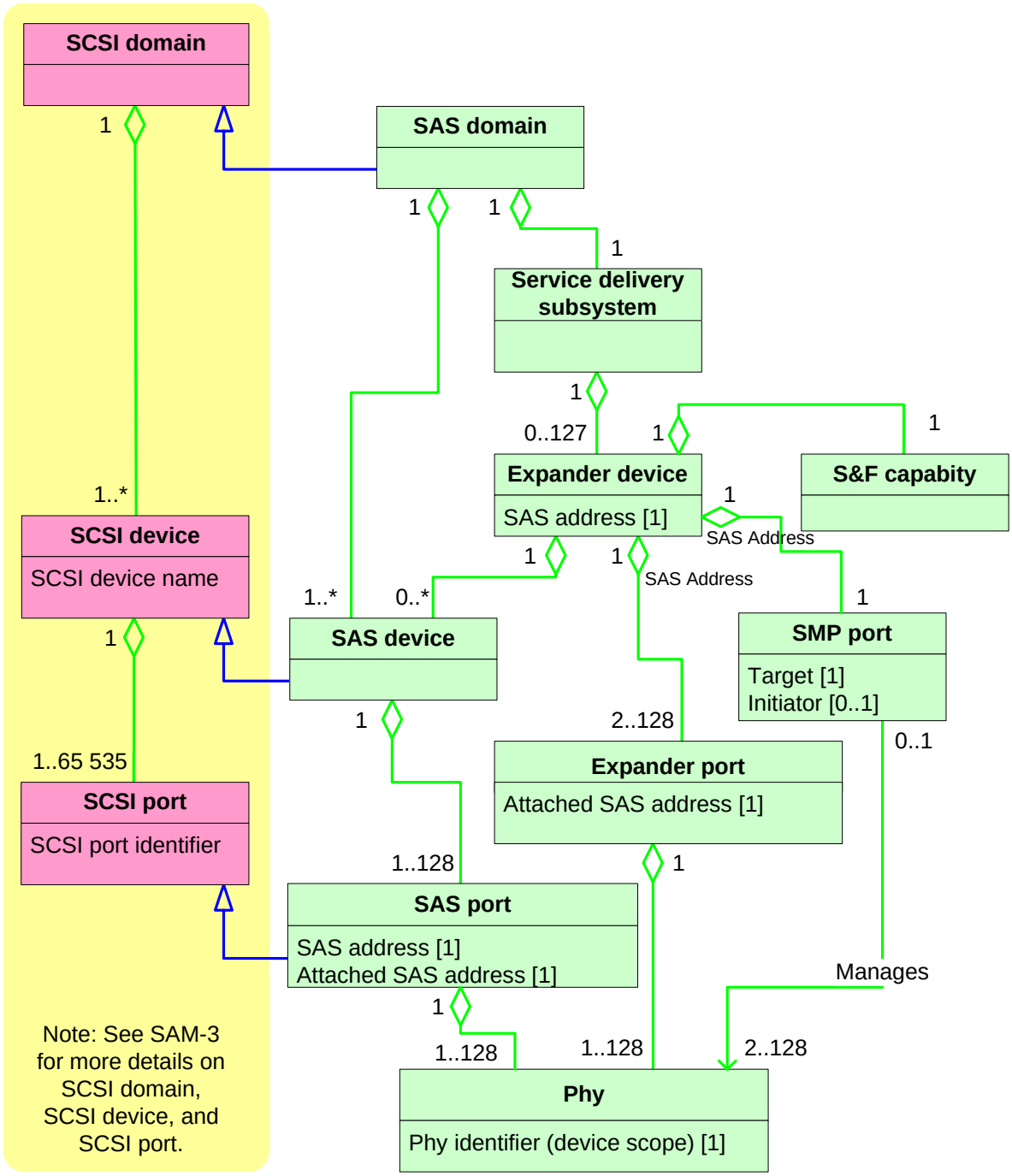

Figure 10 — Aggregation relationships in class diagrams

Editor's Note 3: Added S&F Capability class

Add the following subclause:

4.1.6 S&F buffer pairs

An S&F buffer pair consists of:

a) an S&F buffer context that maintains information about the state of the S&F buffer pair;
b) an S&F in buffer that receives SSP frames or SATA frames through a partial connection to a SAS target port and transmits SSP frames or SATA frames through a partial connection to a SAS initiator port; and
c) an S&F out buffer that receives SSP frames or SATA frames through a partial connection to a SAS initiator port and transmits SSP frames or SATA frames through a partial connection to a SAS target port.

Each S&F buffer is associated with an I_T nexus uniquely identified by an initiator SAS address and a target SAS address. An S&F buffer context shall be created when an S&F expander device processes a connection request from a SAS initiator port to a SAS target port. An S&F buffer pair acts as a proxy for the SAS target port through partial connections to the SAS initiator port, and the S&F buffer pair acts as a proxy for the SAS initiator port through partial connections to the SAS target port.

When both the S&F in buffer and S&F out buffer are empty and there are no partial connections active, an S&F expander device should discard the S&F in buffer and the S&F out buffer, but should maintain the context for the I_T nexus to service a subsequent connection request from SAS target port for a vendor-specific period of time.

In the S&F buffer context, the S&F expander device shall maintain:

a) the SAS addresses of the SAS initiator port and SAS target port;
b) the connection rate to use on the partial pathway from the S&F expander to the SAS target port;
c) the connection rate to use on the partial pathway from the S&F expander to the SAS initiator port;and
d) the partial connection status for each SAS port constituting the I_T nexus.

An S&F expander device may discard the context of an S&F buffer pair (e.g., using a least-recently used (LRU) algorithm to determine when to discard the context for an S&F buffer pair).

If an S&F expander device receives a connection request from a SAS target port for which it has no S&F buffer context, the S&F expander device may not be able to determine the connection rate to use to establish a connection with the SAS initiator port. In this case, the S&F expander should set the CONNECTION RATE field in the OPEN address frame to the negotiated physical link rate of the first phy along the pathway to the SAS initiator port. If the response to the connection request is OPEN_REJECT (COONNECTION RATE NOT SUPPORTED), then the S&F expander device should retry the connection request setting a different value in the CONNECTION RATE field.

Modify subclause 4.1.6 as follows:

4.1.~~6~~7 Expander devices (edge expander devices, ~~and~~ fanout expander devices, and S&F expander devices)

Expander devices are part of a service delivery subsystem and facilitate communication between multiple SAS devices. Expander devices contain two or more external expander ports. Each expander device contains one SMP target port and one management device server, contains one SMP initiator port and one management application client if it is self-configuring and may contain one SMP initiator port and one management application client if it is not self-configuring, and may contain SAS devices (e.g., an expander device may include an SSP target port for access to a logical unit with a peripheral device type set to 0Dh (i.e., enclosure services device) (see SPC-4 and SES-2)).

Figure 18 shows an expander device.

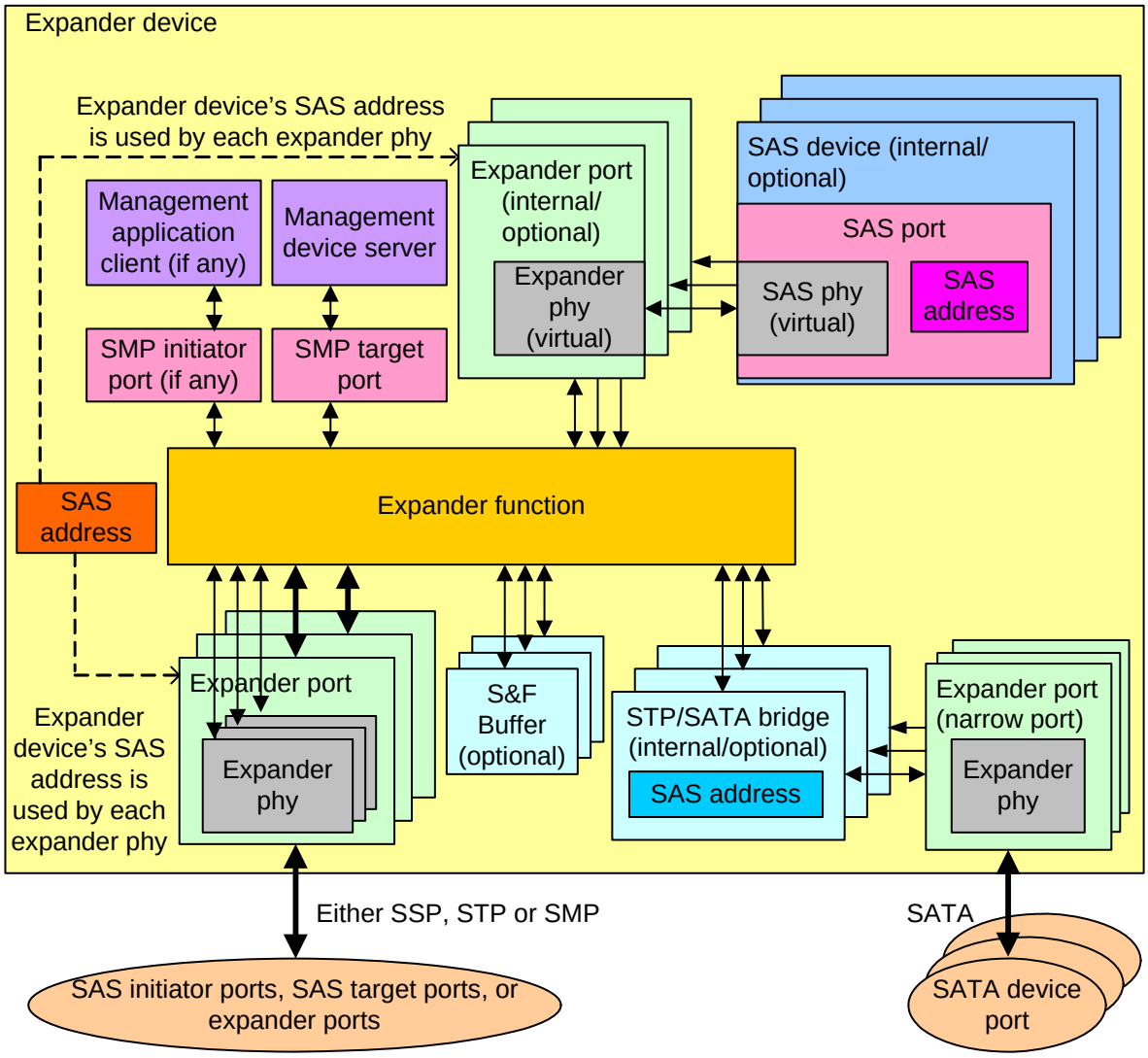

Figure 18 — Expander device

Editor's Note 4: Added S&F Buffer (optional) block

See 4.6 for a detailed model of an expander device.

Figure 19 defines the expander device class.

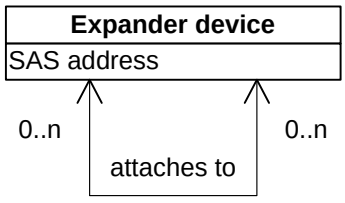

Figure 19 — Expander device class diagram

Each expander phy has one of the following routing attributes (see 4.6.7.1):

a) direct routing attribute;

b) table routing attribute; or
c) subtractive routing attribute.

Expander devices with expander phys with the table routing attribute contain an expander route table (see 4.6.7.3). An externally configurable expander device depends on a management application client within the SAS domain to use the discover process (see 4.7) and the configuration subprocess (see 4.8) to configure the expander route table. A self-configuring expander device contains a management application client and an SMP initiator port to perform the discover process (see 4.7) to configure its own expander route table.

An S&F expander device is an expander device with S&F capabilities: to establish partial connections and to use store-and-forward to transmit SSP frames and SATA frames received from source SAS ports to destination SAS ports.

Editor's Note 5: Subclause 4.1.10 Pathways does not address pathways to virtual phys. An S&F buffer could be viewed as having virtual phys, with a connection to an S&F buffer established along the pathway from a SAS phy to a virtual phy associated with the S&F buffer.

Modify subclause 4.1.11 as follows:

4.1.11 Connections

A connection is a temporary association between a SAS initiator port and a SAS target port. During a connection all dwords from the SAS initiator port are forwarded to the SAS target port, and all dwords from the SAS target port are forwarded to the SAS initiator port.

A connection is pending when an OPEN address frame has been delivered along a completed pathway to the destination phy but the destination phy has not yet responded to the connection request. A connection is established when an OPEN_ACCEPT is received by the source phy.

A connection enables communication for one protocol: SSP, STP, or SMP. For SSP and STP, connections may be opened and closed multiple times during the processing of a command (see 7.12).

The connection rate is the effective rate of dwords through the pathway between a SAS initiator phy and a SAS target phy, or between a SAS phy and an S&F expander phy, established through the connection request. Every phy shall support a 1,5 Gbps connection rate regardless of its physical link rate.

No more than one connection is active on a physical link at a time. If the connection is an SSP or SMP connection and there are no dwords to transmit associated with that connection, idle dwords are transmitted. If the connection is an STP connection and there are no dwords to transmit associated with that connection, SATA_SYNCs, SATA_CONTs, or vendor-specific scrambled data dwords (after a SATA_CONT) are transmitted. If there is no connection on a physical link then idle dwords are transmitted.

The number of connections established by a SAS port shall not exceed the number of SAS phys within the SAS port (i.e., only one connection per SAS phy is allowed). There shall be a separate connection on each physical link.

If multiple potential pathways exist between the SAS initiator port(s) and the SAS target port(s), multiple connections may be established by a SAS port between the following:

a) one SAS initiator port to multiple SAS target ports;
b) one SAS target port to multiple SAS initiator ports; or
c) one SAS initiator port to one SAS target port.

An S&F buffer shall have at most one connection to a SAS target port and at most one connection to a SAS initiator port at any time (i.e., a wide port shall not be used to establish multiple partial connections with an S&F buffer).

If an S&F buffer exists for a SAS initiator port and a SAS target port and the S&F expander receives a second connection request from a SAS port that already has a partial connection to the S&F buffer, the S&F buffer shall reject the connection request with OPEN REJECT (PATHWAY BLOCKED), even if the destination SAS port is a wide port with available phys.

Once a connection is established, the pathway used for that connection shall not be changed (i.e., all the physical links that make up the pathway remain dedicated to the connection until it is closed).

The remainder of subclause 4.10 is unchanged.

Add subclause 4.1.12 as follows:

4.1.12 Partial connections

A partial connection is a temporary association between a SAS initiator port or a SAS target port, and an S&F buffer pair in an S&F expander device.

During a partial connection between a SAS initiator port and an S&F buffer pair, SSP and SATA frames received from the SAS initiator port are forwarded to an S&F out buffer, and SSP and SATA frames from an S&F in buffer are transmitted to the SAS initiator port.

During a partial connection between a SAS target port and an S&F buffer pair, SSP and SATA frames received from the SAS target port are forwarded to an S&F in buffer, and SSP and SATA frames from an S&F out buffer are transmitted to the SAS target port.

Except as stated otherwise, all statements in this standard regarding connections shall also apply to partial connections.

4.1.12.1 Store-and-forward (S&F) buffers

An S&F expander device creates, maintains, and discards S&F buffer pairs. Each S&F buffer pair is identified by an initiator port SAS address in the SOURCE SAS ADDRESS field of an OPEN address frame, and a target port SAS address in the DESTINATION SAS ADDRESS FIELD of an OPEN address frame received from a SAS initiator port. The initiator port SAS address and target port SAS address define an I_T nexus are maintained in the S&F buffer context.

S&F buffer pairs exist within an S&F expander device. S&F buffer pairs are used to hold dwords transmitted to or from SAS ports in partial SSP connections and partial STP connections. S&F buffer pairs shall not be used to hold dwords transmitted in SMP connections (partial or otherwise). An S&F buffer used in an SSP partial connection should be large enough to hold at least one SSP frame with the maximum payload. An S&F buffer used in an STP partial connection should be large enough to hold at least one SATA FIS with the maximum payload.

> NOTE 7 - The S&F buffer protocol provides no way to subdivide a SATA frame, so If the S&F buffer cannot fully contain at least one DATA FIS, then the S&F buffer will achieve little bandwidth aggregation. During transfers from a fast domain to a slower domain, once the S&F buffer is full, dwords transferred from the faster domain will have to throttle to the slower domain's transfer rate (i.e., through SATA_HOLD / SATA HOLDA protocol) until the complete frame has been received by the S&F buffer. During transfers from a slow domain to a faster domain, the S&F buffer must establish a partial connection and begin transfer to the faster domain before receiving the entire SATA frame, the faster domain cannot reconnect to a different S&F buffer until it has completed receiving that same frame.

S&F buffers are further identified by the direction of flow of dwords transferred to or from the S&F buffer. An S&F buffer that receives dwords from a SAS initiator port or transmits dwords to a SAS target port is referred to as an S&F out buffer. An S&F buffer with dwords received from a SAS target port or transmitted to a SAS initiator port is referred to as an S&F in buffer.

4.1.12.2 Partial connection behavior

A SAS port originates and responds to partial connections in the same manner as it does for other connections. A partial connection is different from other connections in that an intermediate S&F expander device along the pathway between a SAS initiator port and a SAS target port performs the link-layer functions on behalf of one of the SAS ports for the duration of the partial connection. No more than one S&F expander along the pathway between a SAS initiator port and a SAS target port shall enable it's S&F function for any given I_T nexus, and the S&F expander which performs the S&F function shall be the S&F expander closest to the SAS target port. The way this is accomplished depends on whether the initial connection request originates from a target port or an initiator port.

If the connection request originates from a SAS initiator port, the SAS initiator port shall set the REMOTE CONNECTION RATE field in the OPEN address frame to a connection rate supported by a potential partial pathway from the S&F expander performing the S&F function to the SAS target port of the I_T nexus. If an S&F expander device along the potential partial pathway with it's S&F capability enabled receives the OPEN address frame with a value representing a valid connection rate in the REMOTE CONNECTION RATE field, then it shall send an AIP(SFT TAKEN) primitive back through the phy through which it received the OPEN address frame. When an S&F expander device nearer the initiator port receives the AIP(SFT TAKEN) it disables it's S&F capability for the I_T nexus. The S&F expander device closest to the target port never receives an AIP(SFT TAKEN), and so its SNF function remains enabled for the partial-connection.

If an S&F expander device receives a connection request with a target port SAS address in the SOURCE SAS ADDRESS field of the OPEN address frame and there is an active S&F buffer context for the I_T nexus identified by the SOURCE SAS ADDRESS field and the DESTINATION SAS ADDRESS FIELD in the OPEN address frame, and if there is space in the S&F buffer to receive a frame (i.e., the S&F buffer can grant credit for an SSP connection request, or the S&F buffer can respond to X_RDY with R_RDY for an STP connection request), then the S&F expander device shall accept the connection request (i.e., establish a partial connection with the SAS target port that originated the OPEN address frame).

If an S&F expander device does not have an active S&F buffer context for the I_T nexus identified by the SOURCE SAS ADDRESS field and the DESTINATION SAS ADDRESS field in the OPEN address frame when an S&F expander device receives a connection request with the INITAITOR PORT bit set to zero, then the S&F expander device shall check the REMOTE CONNECTION RATE field in the OPEN address frame. If the REMOTE CONNECTION RATE field is set to zero or a value representing a valid connection rate, the S&F expander device shall:

a) create an active S&F buffer context for the I_T nexus;
b) set the REMOTE CONNECTION RATE field in the OPEN address frame to a value of 1h;
c) set the CONNECTION RATE field in the OPEN address frame to a value representing the negotiated physical link rate of the next phy along the potential pathway to the SAS iniator port; and
d) forward the modified OPEN address frame to the next phy along the potential pathway to the SAS initiator port.

If an S&F expander device receives an OPEN address frame with the INITIATOR PORT bit set to zero and with the REMOTE CONNECTION RATE field set to 1h, the S&F expander device shall disable its S&F capability for the duration of that partial connection, shall process the OPEN address frame as would an expander device that does not support S&F capability.

An S&F expander device originates an OPEN address frame to establish a partial connection from an S&F buffer pair destined to a SAS port when it has frames to transmit to the destination SAS port. If an S&F expander device originates an OPEN address frame destined for a SAS initiator port, then the S&F expander device shall:

a) set the CONNECTION RATE field to a link rate supported by the potential partial pathway from the S&F expander device to the SAS initiator port;
b) set the REMOTE CONNECTION RATE field to 1h;
c) set the DESTINATION SAS ADDRESS field to the initiator SAS address from the S&F buffer context; and
d) set the SOURCE SAS ADDRESS field to the target SAS address from the S&F buffer context.

If an S&F expander device originates an OPEN address frame destined for a SAS target port, the S&F expander device shall:

a) set the CONNECTION RATE field to a link rate supported by the potential partial pathway from the S&F expander device to the SAS target port;
b) set the REMOTE CONNECTION RATE field to 1h;
c) set the DESTINATION SAS ADDRESS field to the target SAS address from the S&F buffer context; and
d) set the SOURCE SAS ADDRESS field to the initiator SAS address from the S&F buffer context.

The SMP PHY CONTROL command (see ...) may be used to explicitly disable S&F capabilities on a phy-by-phy basis in an S&F expander device, in which case the S&F expander port shall behave as an expander device that does not have S&F capability.

Editor's Note 6: Should the SMP CONFIGURE GENERAL and PHY CONTROL commands to disable S&F buffering return an error if the S&F buffer is in use?

Partial connections may be established for SSP and STP connections, but not for SMP connections.

Modify figure 28 as follows:

Figure 28 shows the state machines for SAS devices, their relationships to each other and to the SAS device, SAS port, and SAS phy classes.

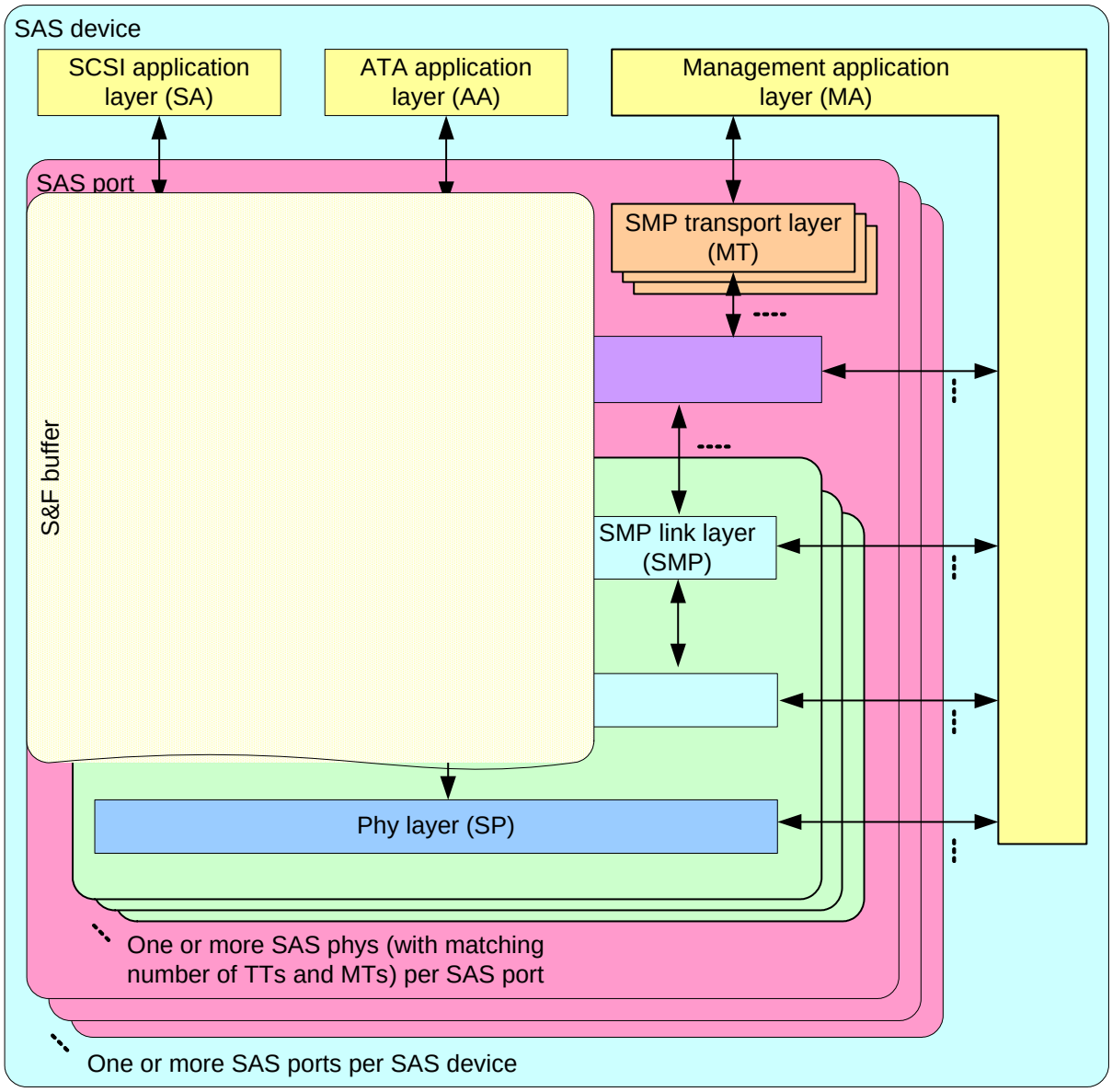

Figure 28 — State machines for SAS devices

Add subclause 4.6.5 as follows:

4.6.5 S&F Buffer Manager (SBM)

The SBM performs the following functions:

a) maps a destination SAS address in a connection request to an internal S&F buffer pair;
b) arbitrates for and grants or withholds S&F buffer resources for connection requests;
c) controls the flow of frames into and out of S&F buffers through partial connections with SAS initiator ports and SAS target ports; and
d) originates opening and closing of partial connections between S&F buffers and SAS initiator ports or SAS target ports.

The SBM may forward connection requests to the ECM if insufficient S&F buffer resources are available to establish a partial connection.

Modify table 15 as follows:

Table 15 — Expander phy to ECR to expander phy responses and confirmations

Message	Description
Arb Status (Normal)	Confirmation/response that AIP (NORMAL) has been received.
Arb Status (Waiting On Partial)	Confirmation/response that AIP (WAITING ON PARTIAL) has been received.
Arb Status (Waiting On Connection)	Confirmation/response that AIP (WAITING ON CONNECTION) has been received.
Arb Status (Waiting On Device)	Confirmation/response that: a) AIP (WAITING ON DEVICE) has been received; or b) the expander phy has completed the forwarding of an OPEN address frame and has entered the XL6:Open_Response_Wait state.
Arb Status (SFT Taken)	Confirmation/response that AIP (SFT TAKEN) has been received indicating the ECM shall disable its store-and-forward function for this connection.
Open Accept	Confirmation/response that OPEN_ACCEPT has been received.
Open Reject	Confirmation/response that OPEN_REJECT has been received.
Backoff Retry	Confirmation/response that: a) a higher priority OPEN address frame has been received (see 7.12.3); and b) the source SAS address and connection rate of the received OPEN address frame are not equal to the destination SAS address and connection rate of the transmitted OPEN address frame.
Backoff Reverse Path	Confirmation/response that: a) a higher priority OPEN address frame has been received (see 7.12.3); and b) the source SAS address and connection rate of the received OPEN address frame are equal to the destination SAS address and connection rate of the transmitted OPEN address frame.

Add subclause 4.6.6.6 as follows:

4.6.6.6 SBM interface

Editor's Note 7: Need to add in tables of requests and indications.

Requests are received from the ECM and/or ECR, and indications are transmitted to the ECM or ECR. The SBM also has an interface to S&F buffers, includes "create" and "delete", and messages to open or close partial connections with S&F buffer pairs. S&F buffers have "state" information in the S&F buffer context as follows:

a) initiator port SAS address;
b) target port SAS address;
c) connection rate for partial connections with the SAS initiator port;
d) connection rate for partial connections with the SAS target port;
e) out full/empty status;
f) in full/empty status; and
g) a list of NAKed tasks (i.e., tag sentries).

Add subclause 4.7.6.4 as follows:

4.6.7.4 S&F buffers

4.6.7.4.1 S&F buffers overview

An S&F expander device establishes an S&F buffer pair for a given I_T nexus to store and forward frames transferred between a SAS initiator port and a SAS target port through partial connections between the S&F buffer pair and the SAS ports. An S&F buffer pair consists of

a) an S&F out buffer which receives SSP or SATA frames from a SAS initiator port and transmits SSP or SATA frames to a SAS target port; and
b) an S&F in buffer which receives SSP or SATA frames from a SAS target port and transmits SSP or SATA frames to a SAS initiator port.

If an S&F expander device receives a connection request where the I_T nexus corresponding to the source SAS address and destination SAS address in the connection request has no S&F buffer context established, the expander device shall route that connection request as a non-S&F expander device would. When the connection request is accepted by the SAS port identified in the destination SAS address, the S&F expander shall establish an S&F buffer pair for the identified I_T nexus.

While an S&F buffer pair exists for a given I_T nexus, the S&F expander device shall act as a proxy for the SAS initiator port and for the SAS target port for the purpose of managing connections (i.e., partial connections) with either SAS port, and may independently receive and transmit frames to and from the S&F buffer pair through partial connections with either SAS port. With respect to the SAS initiator port associated with an S&F buffer pair, the S&F expander device may:

a) close a partial connection with the SAS initiator port while it maintains a partial connection with the SAS target port to receive frames from or transmit frames to the SAS target port;
b) open a partial connection to the SAS initiator port to transmit frames to the SAS initiator port from the S&F in buffer; and
c) accept a connection request received from the SAS initiator port on behalf of the SAS target port and store frames received from the SAS initiator port in the S&F out buffer, even when there is no partial connection between the SAS target port and the S&F buffer pair.

With respect to the SAS target port associated with an S&F buffer pair, the S&F expander device may:

a) close a partial connection with the SAS target port while it maintains a partial connection with the SAS initiator port to receive frames from or transmit frames to the SAS initiator port;
b) open a partial connection to the SAS target port to transmit frames to the SAS target port from the S&F out buffer; and
c) accept a connection request received from the SAS target port on behalf of the SAS initiator port and store frames received from the SAS target port in the S&F in buffer, even when there is no partial connection between the SAS initiator port and the S&F buffer pair.

4.6.7.4.2 S&F buffer SSP flow control

The S&F expander device shall implement flow control through partial connections as a SAS port does for normal connections. For partial SSP connections, an S&F expander device shall not transmit more R_RDY credits to a SAS port than the number of frames it has space to receive (i.e., assuming maximum payload) in the corresponding S&F buffer. If the S&F expander cannot grant R_RDY credit in a partial SSP connection because the S&F buffer is full, the S&F expander shall transmit CREDIT BLOCKED.

4.6.7.4.3 S&F buffer STP flow control

If an S&F buffer receives any frame other than a SATA DATA frame, the S&F buffer shall not return SATA_R_OK for that frame until it has received SATA_R_OK from the destination. If the S&F buffer receives any other response from the destination (e.g., SATA_R_ERR), the S&F buffer shall return the same response to the originator of the SATA frame.

If the S&F buffer receives a SATA DATA frame, the S&F buffer may return SATA_R_OK to the STP port that originated the SATA DATA frame, and exchange SATA_SYNC in both directions with the originator of the

SATA DATA frame. If the originating STP port sends a new SATA_X_RDY, the S&F buffer may respond with SATA_R_RDY and begin receiving the next SATA frame when there is sufficient space available to accept another SATA DATA frame of maximum payload. If the S&F buffer is not expected to have sufficient space avaliable to accept another SATA DATA frame of maximum payload within two microseconds, the S&F buffer shall transmit CLOSE (NORMAL) to the source of the last DATA frame. The S&F buffer shall reject subsequent connection requests with OPEN_REJECT (RETRY) until there is sufficient space in the S&F buffer to accept at least one SATA DATA frame of maximum payload.

Editor's Note 8: A different time value may be desired, but it should be short compared to the time it takes to transmit a SATA DATA frame of maximum payload across a link at the slower link rate.

NOTE 8 - This allows an STP initiator to establish a partial connection with a different S&F buffer while the last SATA DATA frame(s) drain from the one just closed.

When an S&F buffer detects EOF on a SATA frame received from an STP initiator port and there is not enough space left in the S&F out buffer to receive another FIS of maximum size, the S&F buffer shall originate closing the STP connection with the STP initiator port. An S&F out buffer shall transmit HOLD to an STP initiator port as an STP target port would to avoid overflowing the S&F out buffer. An S&F buffer pair shall not initiate closing a connection with an STP initiator port while HOLD is in effect. The S&F buffer shall not accept a new connection request from the STP initiator port until there is enough space in the S&F out buffer to accommodate a SATA frame frame of maximum payload.

Modify subclause 7.2.2 and table 71 as follows:

7.2.2 Primitive summary

Table 71 defines the primitives not specific to the type of connection.

Table 16 — Primitives not specific to type of connection (part 1 of 2)

Primitive	Use [a]	From [b]			To [b]			Primitive sequence type [c]
		I	E	T	I	E	T	
AIP (NORMAL)	NoConn		E		I	E	T	Extended
AIP (RESERVED 0)								
AIP (RESERVED 1)								
AIP (RESERVED 2)								
AIP (RESERVED WAITING ON PARTIAL)								
AIP (WAITING ON CONNECTION)			E					
AIP (WAITING ON DEVICE)			E					
~~AIP (WAITING ON PARTIAL)~~			~~E~~					
AIP (SFT TAKEN)			E		I	E		
ALIGN (0)	All	I	E	T	I	E	T	Single
ALIGN (1)								
ALIGN (2)								
ALIGN (3)								
BREAK	All	I	E	T	I	E	T	Redundant
BREAK_REPLY	All	I	E	T	I	E	T	Redundant
BROADCAST (CHANGE)	NoConn	I	E		I	E	T	Redundant
BROADCAST (SES)				T				
BROADCAST (EXPANDER)			E					
BROADCAST (ASYNCHRONOUS EVENT)				T				
BROADCAST (RESERVED 3)								
BROADCAST (RESERVED 4)								
BROADCAST (RESERVED CHANGE 0)								
BROADCAST (RESERVED CHANGE 1)								
CLOSE (CLEAR AFFILIATION)	STP	I					T	Triple
CLOSE (NORMAL)	Conn	I		T	I		T	
CLOSE (RESERVED 0)								
CLOSE (RESERVED 1)								
EOAF	NoConn	I	E	T	I	E	T	Single
ERROR	All		E		I	E	T	Single
HARD_RESET	NoConn	I	E		I	E	T	Redundant
NOTIFY (ENABLE SPINUP)	All	I	E				T	Single
NOTIFY (POWER LOSS EXPECTED)		I	E				T	
NOTIFY (RESERVED 1)					I	E	T	
NOTIFY (RESERVED 2)								
OPEN_ACCEPT	NoConn	I		T	I		T	Single

Table 16 — Primitives not specific to type of connection (part 2 of 2)

Primitive	Use [a]	From [b]			To [b]			Primitive sequence type [c]
		I	E	T	I	E	T	
OPEN_REJECT (BAD DESTINATION)	NoConn		E		I		T	Single
OPEN_REJECT (CONNECTION RATE NOT SUPPORTED)		I	E	T				
OPEN_REJECT (NO DESTINATION)			E					
OPEN_REJECT (PATHWAY BLOCKED)			E					
OPEN_REJECT (PROTOCOL NOT SUPPORTED)		I		T				
OPEN_REJECT (ZONE VIOLATION)			E					
OPEN_REJECT (RESERVED ABANDON 1)								
OPEN_REJECT (RESERVED ABANDON 2)								
OPEN_REJECT (RESERVED ABANDON 3)								
OPEN_REJECT (RESERVED CONTINUE 0)								
OPEN_REJECT (RESERVED CONTINUE 1)								
OPEN_REJECT (RESERVED INITIALIZE 0)								
OPEN_REJECT (RESERVED INITIALIZE 1)								
OPEN_REJECT (RESERVED STOP 0)								
OPEN_REJECT (RESERVED STOP 1)								
OPEN_REJECT (RETRY)		I		T				
OPEN_REJECT (STP RESOURCES BUSY)			E	T				
OPEN_REJECT (WRONG DESTINATION)		I		T			T	
SOAF	NoConn	I	E	T	I	E	T	Single

[a] The Use column indicates when the primitive is used:
a) NoConn: SAS physical links, outside connections;
b) Conn: SAS physical links, inside connections;
c) All: SAS physical links, both outside connections or inside any type of connection; or
d) STP: SAS physical links, inside STP connections.

[b] The From and To columns indicate the type of ports that originate each primitive or are the intended destinations of each primitive:
a) I for SAS initiator ports;
b) E for expander ports; and
c) T for SAS target ports.
Expander ports are not considered originators of primitives that are passing through from expander port to expander port.

[c] The Primitive sequence type columns indicate whether the primitive is sent as a single primitive sequence, a repeated primitive sequence, a continued primitive sequence, an extended primitive sequence, a triple primitive sequence, or a redundant primitive sequence (see 7.2.4).

Modify subclause 7.2.3and table 74 as follows:

7.2.3 Primitive encodings

Table 82 defines the primitive encoding for primitives not specific to type of connection.

Table 82 — Primitive encoding for primitives not specific to type of connection (part 1 of 2)

Primitive	Character			
	1st	2nd	3rd	4th (last)
AIP (NORMAL)	K28.5	D27.4	D27.4	D27.4
AIP (RESERVED 0)	K28.5	D27.4	D31.4	D16.7
AIP (RESERVED 1)	K28.5	D27.4	D16.7	D30.0
AIP (RESERVED 2)	K28.5	D27.4	D29.7	D01.4
~~AIP (RESERVED WAITING ON PARTIAL)~~	~~K28.5~~	~~D27.4~~	~~D01.4~~	~~D07.3~~
AIP (SFT TAKEN)	K28.5	D27.4	D01.4	D07.3
AIP (WAITING ON CONNECTION)	K28.5	D27.4	D07.3	D24.0
AIP (WAITING ON DEVICE)	K28.5	D27.4	D30.0	D29.7
AIP (WAITING ON PARTIAL)	K28.5	D27.4	D24.0	D04.7
ALIGN (0)	K28.5	D10.2	D10.2	D27.3
ALIGN (1)	K28.5	D07.0	D07.0	D07.0
ALIGN (2)	K28.5	D01.3	D01.3	D01.3
ALIGN (3)	K28.5	D27.3	D27.3	D27.3
BREAK	K28.5	D02.0	D24.0	D07.3
BREAK_REPLY	K28.5	D02.0	D29.7	D16.7
BROADCAST (CHANGE)	K28.5	D04.7	D02.0	D01.4
BROADCAST (SES)	K28.5	D04.7	D07.3	D29.7
BROADCAST (EXPANDER)	K28.5	D04.7	D01.4	D24.0
BROADCAST (ASYNCHRONOUS EVENT)	K28.5	D04.7	D04.7	D04.7
BROADCAST (RESERVED 3)	K28.5	D04.7	D16.7	D02.0
BROADCAST (RESERVED 4)	K28.5	D04.7	D29.7	D30.0
BROADCAST (RESERVED CHANGE 0)	K28.5	D04.7	D24.0	D31.4
BROADCAST (RESERVED CHANGE 1)	K28.5	D04.7	D27.4	D07.3
CLOSE (CLEAR AFFILIATION)	K28.5	D02.0	D07.3	D04.7
CLOSE (NORMAL)	K28.5	D02.0	D30.0	D27.4
CLOSE (RESERVED 0)	K28.5	D02.0	D31.4	D30.0
CLOSE (RESERVED 1)	K28.5	D02.0	D04.7	D01.4
EOAF	K28.5	D24.0	D07.3	D31.4
ERROR	K28.5	D02.0	D01.4	D29.7

Table 82 — Primitive encoding for primitives not specific to type of connection (part 2 of 2)

Primitive	Character			
	1st	2nd	3rd	4th (last)
HARD_RESET	K28.5	D02.0	D02.0	D02.0
NOTIFY (ENABLE SPINUP)	K28.5	D31.3	D31.3	D31.3
NOTIFY (POWER LOSS EXPECTED)	K28.5	D31.3	D07.0	D01.3
NOTIFY (RESERVED 1)	K28.5	D31.3	D01.3	D07.0
NOTIFY (RESERVED 2)	K28.5	D31.3	D10.2	D10.2
OPEN_ACCEPT	K28.5	D16.7	D16.7	D16.7
OPEN_REJECT (BAD DESTINATION)	K28.5	D31.4	D31.4	D31.4
OPEN_REJECT (CONNECTION RATE NOT SUPPORTED)	K28.5	D31.4	D04.7	D29.7
OPEN_REJECT (NO DESTINATION)	K28.5	D29.7	D29.7	D29.7
OPEN_REJECT (PATHWAY BLOCKED)	K28.5	D29.7	D16.7	D04.7
OPEN_REJECT (PROTOCOL NOT SUPPORTED)	K28.5	D31.4	D29.7	D07.3
OPEN_REJECT (ZONE VIOLATION)	K28.5	D31.4	D02.0	D27.4
OPEN_REJECT (RESERVED ABANDON 1)	K28.5	D31.4	D30.0	D16.7
OPEN_REJECT (RESERVED ABANDON 2)	K28.5	D31.4	D07.3	D02.0
OPEN_REJECT (RESERVED ABANDON 3)	K28.5	D31.4	D01.4	D30.0
OPEN_REJECT (RESERVED CONTINUE 0)	K28.5	D29.7	D02.0	D30.0
OPEN_REJECT (RESERVED CONTINUE 1)	K28.5	D29.7	D24.0	D01.4
OPEN_REJECT (RESERVED INITIALIZE 0)	K28.5	D29.7	D30.0	D31.4
OPEN_REJECT (RESERVED INITIALIZE 1)	K28.5	D29.7	D07.3	D16.7
OPEN_REJECT (RESERVED STOP 0)	K28.5	D29.7	D31.4	D07.3
OPEN_REJECT (RESERVED STOP 1)	K28.5	D29.7	D04.7	D27.4
OPEN_REJECT (RETRY)	K28.5	D29.7	D27.4	D24.0
OPEN_REJECT (STP RESOURCES BUSY)	K28.5	D31.4	D27.4	D01.4
OPEN_REJECT (WRONG DESTINATION)	K28.5	D31.4	D16.7	D24.0
SOAF	K28.5	D24.0	D30.0	D01.4

Modify table 86 as follows:

Table 86 — AIP primitives

Primitive	Description
AIP (NORMAL)	Expander device has accepted the connection request. This may be sent multiple times (see 7.12.4.2).
AIP (RESERVED 0)	Reserved. Processed the same as AIP (NORMAL).
AIP (RESERVED 1)	Reserved. Processed the same as AIP (NORMAL).
AIP (RESERVED 2)	Reserved. Processed the same as AIP (NORMAL).
AIP (WAITING ON CONNECTION)	Expander device has determined the routing for the connection request, but either the destination phys are all being used for connections or there are insufficient routing resources to complete the connection request. This may be sent multiple times (see 7.12.4.2).
AIP (WAITING ON DEVICE)	Expander device has determined the routing for the connection request and forwarded it to the output physical link. This is sent one time (see 7.12.4.2).
AIP (WAITING ON PARTIAL)	Expander device has determined the routing for the connection request, but the destination phys are all busy with other partial pathways. This may be sent multiple times (see 7.12.4.2).
~~AIP (RESERVED WAITING ON PARTIAL)~~	~~Reserved. Processed the same as AIP (WAITING ON PARTIAL).~~
AIP (SFT TAKEN)	An an expander device along the pathway to the destination has S&F capability, the S&F capability is enabled and will be performed by that expander device for the duration of the connection or partial connection. An S&F expander device shall: 1) transmit AIP (SFT TAKEN) to the phy over which it received an OPEN address frame when it activates it's S&F capability for that connection; and 2) deactivate its S&F capability for the duration of the connection or partial connection if it receives AIP (SFT TAKEN) from a phy to which it forwarded the OPEN address frame.

Modify subclause 7.2.5.8 as follows:

7.2.5.8 ERROR

ERROR should be sent by an expander device when it is forwarding dwords from a SAS physical link or SATA physical link to a SAS physical link and it receives an invalid dword or an ERROR.

NOTE 9 - Since an 8b10b coding error in one dword is sometimes not detected until the next dword (see table 58 in 6.3.5), expander devices should avoid deleting invalid dwords or ERRORs unless necessary (e.g., if the elasticity buffer is full) to avoid hiding evidence that an error has occurred.

ERROR primitives are forwarded through an S&F buffer just as they would be without an S&F buffer.

See 7.15 for details on error handling by expander devices.

Modify table 101 as follows:

Table 101 — IDENTIFY address frame format

Byte\Bit	7	6	5	4	3	2	1	0
0	Restricted (for OPEN address frame)	DEVICE TYPE			ADDRESS FRAME TYPE (0h)			
1	Restricted (for OPEN address frame)							
2	Reserved			S&F CAPABLE	SSP INITIATOR PORT	STP INITIATOR PORT	SMP INITIATOR PORT	Restricted (for OPEN address frame)
3	Reserved				SSP TARGET PORT	STP TARGET PORT	SMP TARGET PORT	Restricted (for OPEN address frame)
4	DEVICE NAME							
11								
12	SAS ADDRESS							
19								
20	PHY IDENTIFIER							
~~20~~21	Reserved							BREAK_REPLY CAPABLE
22	Reserved							
27								
28	(MSB) CRC							
31								(LSB)

Add a description of the S&F capable bit:

An S&F CAPABLE bit set to one specifies that an expander device has the capability to transfer SSP frames and SATA frames through an S&F buffer through partial connections to SAS initiator ports and SAS target ports at different connection rates. An S&F CAPABLE bit set to zero specifies that an expander device does not have store-and-forward capability.

Modify table 94 as follows:

Table 102 — OPEN address frame format

Byte\Bit	7	6	5	4	3	2	1	0
0	INITIATOR PORT	PROTOCOL			ADDRESS FRAME TYPE (1h)			
1	~~FEATURES~~ REMOTE CONNECTION RATE				CONNECTION RATE			
2	(MSB)	INITIATOR CONNECTION TAG						
3								(LSB)
4		DESTINATION SAS ADDRESS						
11								
12		SOURCE SAS ADDRESS						
19								
20	SOURCE ZONE GROUP							
21	PATHWAY BLOCKED COUNT							
22	(MSB)	ARBITRATION WAIT TIME						
23								(LSB)
24		MORE COMPATIBLE FEATURES						
27								
28	(MSB)	CRC						
31								(LSB)

Add a description of the REMOTE CONNECTION RATE field:

If a SAS initiator port specifies a value representing a valid connection rate (i.e., 8h for 1.5 Gbps, 9h for 3 Gbps, or Ah for 6 Gbps) in the REMOTE CONNECTION RATE field, the CONNECTION RATE field specifies the connection rate for the partial connection between the SAS initiator port and the S&F expander device providing an S&F buffer pair for the specified I_T nexus.

To request S&F service in an S&F expander device along a potential pathway to the SAS target port, a SAS initiator port shall set the REMOTE CONNECTION RATE field to a value representing a valid connection rate . The REMOTE CONNECTION RATE field shall be set to a value supported by at least one potential pathway from the S&F expander device providing the S&F service to the SAS target port.

If an S&F expander receives an OPEN address frame with the address of a SAS initiator port in the SOURCE SAS ADDRESS field and a value of zero or one in the REMOTE CONNECTION RATE field, it shall disable its S&F function for that I_T nexus and continue processing the OPEN address fame as it would if it did not support the S&F capability.

When an S&F expander device forwards an OPEN address frame with the INITIATOR PORT bit set to one and the value in the REMOTE CONNECTION RATE field is zero, the S&F expander shall process the OPEN address frame as it would if it did not support the S&F capability.

When an S&F expander device forwards an OPEN address frame with the INITIATOR PORT bit set to one and the REMOTE CONNECTION RATE field is not zero, if:

a) the value in the CONNECTION RATE field is greater than the negotiated physical link rate of the phy to which the OPEN address frame is to be routed; or
b) the expander device attached to the phy to wich the OPEN address frame is to be routed is not S&F capable (i.e., the S&F CAPABLE bit in the IDENTIFY address frame from that expander device is set to zero);

then the S&F expander device shall:

1) create an S&F buffer context for the I_T nexus, store the value from the CONNECTION RATE field in the S&F buffer context as the initiator SAS port connection rate, and store the value form the REMOTE CONNECTION RATE field in the S&F buffer context as the target SAS port connection rate;
2) set the REMOTE CONNECTION RATE field in the OPEN address frame to 1h to disable S&F buffering in S&F expander devices closer to the destination SAS port.;
3) set the CONNECTION RATE field in the OPEN address frame to the value of the negotiated physical link rate of the phy;
4) recalculate the CRC and forward the modified OPEN address frame to the selected phy; and
5) retrun AIP (SFT TAKEN) to the phy through which it received the OPEN address frame.

When originating an OPEN address frame destined for a SAS port, the S&F expander device shall set the CONNECTION RATE field to a value supported by at least one potential pathway to the SAS port identified by the DESTINATION SAS ADDRESS field, and shall set the REMOTE CONNECTION RATE field to 1h.

NOTE 10 - The S&F expander may use the values from a previous connection request received from the SAS initiator port for the same I_T nexus.

Editor's Note 9: Several changes need to be added to subclause 7.12 Connections. One key element is that when an S&F buffer creates a connection request destined to a SAS initiator port, it may set the ARBITRATION WAIT TIME field in the OPEN address frame to 8000h to gain priority over other connection requests. This is needed to make sure the S&F in buffer can drain frames efficiently and maintain buffer resources available to serivce S&F traffic through the interconnect.

Add the following text to the end of subclause 7.16.3 SSP frame transmission and reception:

If an S&F buffer detects a NAK on a frame transmitted to an SSP initiator port, the S&F buffer shall discard the NAK and establish a sentry for the affected I_T_L_Q nexus. The S&F buffer shall continue to forward frames received from the SSP target port with that tag to the SSP initiator port until an SMP PHY CONTROL command is received from the SSP initiator to abort that tag and the accompanying sentry.

An S&F buffer may delete the sentry for the I_T_L_Q nexus after either:

a) the S&F buffer receives and SMP PHY CONTROL command to abort the tag; or
b) the S&F buffer receives an SMP command to clear/reset the S&F buffer.

An S&F buffer that has an outstanding task tag but is otherwise empty may not be deleted except by an SMP command.

The SBM may delete an empty S&F buffer if:

a) there are no partial connections to the S&F buffer;
b) the S&F buffer is empty; and
c) there are no active sentrys associated with the S&F buffer.

Modify subclause 7.16.4 as follows:

7.16.4 SSP flow control

An SSP phy uses RRDY to grant credit for permission for the other SSP phy in the connection to transmit frames. Each RRDY increments credit by one frame. Frame transmission decrements credit by one frame. Credit of zero frames is established at the beginning of each connection.

SSP phys shall not increment credit past 255 frames.

To prevent deadlocks where an SSP initiator port and SSP target port are both waiting on each other to provide credit, an SSP initiator port shall never refuse to provide credit by withholding RRDY because it needs to transmit a frame itself. It may refuse to provide credit for other reasons (e.g., temporary buffer full conditions).

An SSP target port may refuse to provide credit for any reason, including because it needs to transmit a frame itself.

If credit is zero, SSP phys that are going to be unable to provide credit for 1 ms may send CREDIT_BLOCKED. The other phy may use this to avoid waiting 1 ms to transmit DONE (CREDIT TIMEOUT) (see 7.16.8).

If credit is zero, an SSP S&F buffer that is unable to provide credit shall send CREDIT_BLOCKED.

If credit is nonzero, SSP phys that are going to be unable to provide additional credit for 1 ms, even if they receive frames per the existing credit, may transmit CREDIT_BLOCKED.

If credit is nonzero, an SSP S&F buffer that is unable to provide additional credit, even if it receives frames per the existing credit, shall transmit CREDIT_BLOCKED.

After sending CREDIT_BLOCKED, an SSP phy or S&F buffer shall not transmit any additional RRDYs in the connection.

Editor's Note 10: Subclause 7.16.5 needs to be updated to explain that an S&F expander device must track task tags in SSP frames flowing through an S&F buffer to enforce interlocked frame transmission rules when transmitting frames to the destination SAS port. S&F expander devices also must isolate the affected task tag when a NAK or ACK/NAK timeout is detected and discard any frames received thereafter with the same task tag until the condition is cleared by an SMP BUFFER CONTROL request.

Add the following subclause 7.17.5 to explain how affiliations work with S&F expander devices

7.17.5 Affiliations and S&F expander devices

S&F expander devices introduce additional complexity in managing affilations. Establishing affiliations is no different. When an S&F expander establishes a partial connection with an STP target port with an STP/SATA bridge, an affiliation is established with the STP initiator port associated with the I_T nexus of the S&F buffer.

An STP initiator port specifies to close an affiliation by transmitting a CLOSE (CLEAR AFFILIATION) after having verified neither the SATA device nor the STP initiator has anymore SATA frames to transmit. However, it is possible the last frame to be transmitted between the STP initiator port and the STP target port originated from the STP initiator port.

When an S&F buffer receives a CLOSE (CLEAR AFFILIATION) from an STP inititiator port, the S&F buffer shall:

1) close the connection with the STP initiator port (i.e., reply with CLOSE);
2) complete transmission of any SATA frames in the S&F out buffer to the STP target port; and
3) close the connection to the STP target port by transmitting CLOSE (CLEAR AFFILIATION).

If an STP application client clears an affiliation by any other means, the application client should also issue an SMP S&F BUFFER CONTROL request with the buffer control field set to DELETE STP BUFFER and abort

any outstanding ATA commands to maintain synchronization between STP application client, SATA device, and S&F buffer context.

Editor's Note 11: Subclause 7.17.7 needs elaboration to explain S&F buffer handling of SATA_X_RDY crossing on the wire (or in the buffer). The STP initiator port is supposed to win arbitration, but the S&F buffer may have already accepted a partial connection from the STP target port and accepted a FIS when the S&F initiator port requests a connection to transmit SATA_X_RDY. In this case the S&F expander device accepts the connection request and a FIS from the STP initaitor port and retries the transmission to the STP target port after the STP target port is finished transmitting a FIS.

Add subclause 7.17.8 as follows:

7.17.8 STP partial connections with S&F buffer pairs

An S&F out buffer shall not close a partial connection to an STP initiator port except after receiving SATA_EOF from the STP initiator port for the previous SATA frame, and before transmitting SATA_R_RDY to receive the next SATA frame. If the S&F expander device receives a SATA_EOF from an STP initiator port and the number of dwords remaining in the S&F out buffer is less than the maximum size of a SATA frame, the S&F buffer shall transmit CLOSE (NORMAL) to the STP initiator port. If the STP initiator port sends a SATA_X_RDY, before the CLOSE arrives, the STP initiator port shall transmit CLOSE to complete closing the connection, and shall retry the SATA_X_RDY after establishing a new partial connection.

The S&F buffer shall reject subsequent connection requests from that STP initiator with OPEN_REJECT (RETRY) until there is at least enough space in the S&F out buffer to receive another maximum size FIS. The S&F buffer pair may continue rejecting connection requests from the STP initiator port until the S&F out buffer is empty.

If the management application client has knowledge that the maximum size of a SATA frame will not exceed a specific value, an SMP S&F BUFFER CONTROL command (see...) may be used to configure the maximum FIS size for an S&F expander device. The S&F buffer shall use this value to determine if there is sufficient space in the S&F buffer to accept another SATA frame of maximum size.

NOTE 11 - An STP initiator may try to limit the device-to-host FIS size by setting Sectors Per Block field in the SET MULTIPLE command, but this limit does not apply to FPDMA commands.

When forwarding a SATA frame containing a DATA FIS originating from an STP target port, the S&F buffer may close a partial connection to the STP initiator port after sending SATA_EOF, SATA_WTRM; receiving SATA_R_OK or SATA_R_ERR, and exchanging SYNC in both directions. If a SATA frame containing a DATA FIS originating from an STP target port is larger than the S&F in buffer and the S&F buffer becomes empty before the S&F buffer has received EOF from the STP target port, then the S&F buffer shall repeatedly transmit either SATA_HOLD or ALIGNs to the STP initiator port.

An S&F buffer shall be large enough to receive the maximum size SATA frame that does not contain a DATA FIS, and should be large enough to accommodate the maximum size SATA frame that contains a DATA FIS.

Editor's Note 12: Need to clarify in subclause 9.2.4 that an S&F expander device (as defined in SAS-2) does not support transport layer retries. S&F buffering still has to handle the non TLR case anyway, so that should be the initial priority.

Editor's Note 13: 9.2.4.3 Task frame - add handling of NAK, ACK/NAK timeout on SSP task frames

---

Add subclause 9.2.5 explaining transport-layer handling of link layer errors with S&F buffers

9.2.5 Transport layer handling of link layer errors with S&F buffering in use

9.2.5.1 Transport layer handling of link layer errors with S&F buffering in use overview

An S&F buffer may receive a NAK when forwarding an SSP frame to a destination SAS port after having returned an ACK for that frame to the source SAS port. Subclauses 9.2.5.2 through 9.2.5.6 describe the transport-layer handling of these errors.

To identify the affected I_T_L_Q nexus, and in order to observe the rules for interlocked frames, the S&F buffer shall recognize the value in the TAG field of the last SSP frame transmitted, and associate it with the appropriate I_T_L_Q nexus.

9.2.5.2 Handling NAK or ACK/NAK timeout on a COMMAND frame transmitted to an SSP target port

If the S&F buffer receives a NAK or an ACK/NAK timeout in response to a COMMAND frame transmitted to an SSP target port, the S&F buffer may retransmit the COMMAND frame one or more times in the same or in a new connection. If the S&F buffer continues to receive NAK or ACK/NAK timeout in response, the S&F buffer shall discard the COMMAND frame and establish a sentry for the affected I_T_L_Q nexus. The application client will experience an upper-layer protocol time-out. The application client should apply the procedure for handling application-level timeouts when S&F buffering is used (see 10.2.3).

9.2.5.3 Handling NAK or ACK/NAK timeout on an XFER_RDY frame transmitted to an SSP initiator port

If the S&F buffer receives a NAK in response to transmitting an XFER_RDY frame to an SSP initiator port, the NAK is discarded, and the application client experiences an upper-layer protocol timeout. The application client should apply the procedure for handling application-level timeouts when S&F buffering is used (see 10.2.3).

9.2.5.4 Handling NAK or ACK/NAK timeout on a DATA frame transmitted to an SSP target port

The SSP initiator port has received an ACK for the frame so the application client does not immediately abort the command, but rather continues transmitting the next DATA frame or waits for an XFER_RDY or RESPONSE frame from the SSP target port. If the SSP target port receives the next DATA frame, it will detect a DELIVERY FAILURE - DATA OFFFSET ERROR and will terminate the command with CHECK CONDITION STATUS with the sense code set to ABORTED COMMAND and the additional sense code set to DATA OFFSET ERROR. The application client should send an SMP S&F BUFFER CONTROL command to the S&F buffer with the BUFFER CONTROL field set to ABORT TASK and the TASK TAG field set to the tag of the I_T_L_Q nexus that returned the error. If the application client waits for an XFER_RDY or RESPONSE frame, then the application client will eventually time-out. The application client should apply the procedure for handling application-level timeouts when S&F buffering is used (see 10.2.3).

9.2.5.5 Handling NAK or ACK/NAK timeout on a DATA frame transmitted to an SSP initiator port

The S&F buffer discards the NAK. If there is more data for the command remaining to transfer, the SSP target port will transmit a subsequent DATA frame for that I_T_L_Q nexus, the SSP initiator port will detect a DATA OFFSET error and abort the commad. The application client should send an SMP S&F BUFFER CONTROL command to the S&F buffer with the BUFFER CONTROL field set to ABORT TASK and the TASK TAG field set to the tag of the I_T_L_Q nexus that returned the error. If there is no more data remaining to transfer for the command, the application client will experience an application-level timeout waiting for the RESPONSE frame for that I_T_L_Q nexus. The application client should apply the procedure for handling application-level timeouts when S&F buffering is used (see 10.2.3).

9.2.5.6 Handling a NAK or ACK/NAK timeout on a RESPONSE frame transmitted to an SSP initiator port

The S&F buffer discards the NAK. The S&F buffer may retransmit the RESPONSE frame in the same or a new partial connection a vendor-specific number of times as would an SSP target port (see 9.2.4.6). If retries fail, the application client will experience an application-level timeout. The S&F buffer shall not create a sentry

for the I_T_L_Q nexus, but shall discard the RESPONSE frame and the context for the I_T_L_Q nexus. The application client should apply the procedure for handling application-level timeouts when S&F buffering is used (see 10.2.3).

Add subclause 10.2.3 covering application client handling of SCSI timeouts for S&F buffers

10.2.3 Application client handling of SCSI timeouts when S&F buffering is used

When an S&F buffer is used between an SSP initiator port and an SSP target port, a frame that has received an ACK from the S&F buffer may receive a NAK when transmitted to the destination. In such cases S&F buffer discards the NAK, and in some cases (e.g., when an SSP initiator port returns a NAK for a RESPONSE frame), the application client will experience a timeout.

If the application client experiences a timeout and S&F buffering is used, the application client shall send a QUERY TASK task management function with Send Task Management Request () to determine whether the command was received successfully, and shall send an SMP S&F BUFFER CONTROL command to the S&F buffer with the BUFFER CONTROL field set to QUERY SSP BUFFER and the TASK TAG field set to the tag of the I_T_L_Q nexus that timed out to determine if the S&F buffer has a sentry active for that task. The application client shall process the timeout condition according to the Service Response returned for the QUERY TASK task management function and the TASK STATUS field in the S&F BUFFER CONTROL response as shown in table 103.

Table 103 — Application client recovery from application-level time-out

QUERY TASK response	TASK STATUS field	Application client recovery
FUNCTION SUCCEEDED	SENTRY ACTIVE	The S&F buffer has discarded a NAK but the device server is still processing the command. The application client shall abort the command (e.g., by sending an ABORT TASK task management function), and shall send an SMP S&F BUFFER CONTROL command to the S&F buffer with the BUFFER CONTROL field set to ABORT TASK and the TASK TAG field set to the tag of the I_T_L_Q nexus that timed out. The application client may then reuse the tag.
	FRAMES PRESENT	The S&F buffer did not discard a NAK and is waiting to transmit one or more frames to the SSP initiator port and/or the SSP target port.
	FRAMES NOT PRESENT	The application timeout was caused by the device server. The application client should invoke the appropriate application-level recovery (see SAM-3).
FUNCTION COMPLETE	SENTRY ACTIVE	The S&F buffer has discarded a NAK. bit the device server has completed processing of the command. The application client shall send an SMP S&F BUFFER CONTROL command to the S&F buffer with the BUFFER CONTROL field set to ABORT TASK and the TASK TAG field set to the tag of the I_T_L_Q nexus that timed out. The application client may then reuse the tag.
	FRAMES PRESENT	The S&F buffer did not discard a NAK and is waiting to transmit one or more frames to the SSP initiator port.
	FRAMES NOT PRESENT	The device server may have discarded the command. The application client may reuse the tag.

SMP Changes:

A new bit is added to byte 10 of the SMP REPORT GENERAL response, the S&F CAPABLE bit. This bit is set to one by S&F expander devices that have the S&F capability and is able to establish partial connections between S&F buffer pairs and SAS ports.

A bit is added to byte 12 of the SMP DISCOVER response, the S&F ENABLED bit. The S&F expander device sets this bit to one to indicate it has the S&F function enabled for the specified phy.

A new link rate encoding, Ah, is added to the nogitiated physical link rate field in the DISCOVER response as follows:

Ah: Phy is enabled; 6,0 Gbps physical link rate. This field shall be updated to this value after the speed negotiation sequence completes.

In the SMP CONFIGURE GENERAL request byte 8, this proposal adds a 3-bit wide ENABLE/DISABLE S&F BUFFERING field. The description of the ENABLE/DISABLE S&F BUFFERING field is:

The ENABLE/DISABLE S&F BUFFERING field is used to enable or disable S&F capability on all phys of an S&F expander device. This overrides previous per-phy settings set using the PHY CONTROL command. An S&F expander device shall enable or disable the S&F capability for all phys as shown in table 104.

Table 104 — ENABLE/DISABLE S&F BUFFERING field

Code	Function
000b	Do not change S&F buffer enable settings on any phys
001b	Enable S&F buffering on all phys
010b	Disable S&F buffering on all phys
011b - 111b	Reserved

An S&F expander device shall not create an S&F buffer context for a given I_T nexus unless S&F buffering is enabled both on an S&F expander phy beginning a partial pathway from the S&F expander device to the SAS target port and on an S&F expander phy beginning the partial pathway from the S&F expander device to the SAS initiator port. If the S&F function is disabled for a phy that would be used by a partial conneciton to an existing S&F buffer pair, that S&F buffer pair shall be deleted only when it no longer contains any frames for the I_T nexus (i.e., the S&F buffer function remains enabled as long as there are frames in transit through the S&F buffer pair between the SAS initiator port and SAS target port).

A similar field is added in byte-11 of the SMP PHY CONTROL request to provide the ability to enable or disable S&F service on a phy-by-phy basis. Both phys along the pathway through an S&F expander device between a SAS initiator port and a SAS target port have to have the S&F function enabled for the S&F expander to provide the S&F service for the I_T nexus.

The SMP S&F BUFFER CONTROL request shown in table 105 specifies functions processed by an S&F buffer for a specified I_T nexus.

Table 105 — S&F BUFFER CONTROL request

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (40h)							
1	FUNCTION (93h)							
2	Reserved							
3	REQUEST LENGTH (09h)							
4	Reserved							
6								
7	BUFFER OPERATION							
8	INITIATOR PORT SAS ADDRESS (see Editor's Note 14:)							
15								
16	TARGET PORT SAS ADDRESS							
23								
24	PARAMETER SELECT							
25	Reserved							
26	TASK TAG							
27								
28	(MSB)	PARAMETER VALUE						
29								(LSB)
30	Reserved							
39								
40	(MSB)	CRC						
43								(LSB)

Editor's Note 14: Alternatively, the S&F expander device could use the SOURCE SAS ADDRESS field from the SMP OPEN address frame as the initiator port SAS address.

The S&F BUFFER CONTROL request is used to query an STP or SSP buffer, to delete an SSP or STP buffer, to abort a specific task with frames in an S&F buffer, or to set parameters affecting STP flow control for an STP S&F buffer.

Buffer operations are specified in the BUFFER OPERATION field as shown in table 106.

Table 106 — BUFFER CONTROL field

Code	Mneumonic	Description
00h:	QUERY SSP BUFFER	Returns NOT PRESENT or PRESENT in the BUFFER PAIR STATUS field of the BUFFER CONTROL response for the buffer pair associated with the I_T nexus identified in the INITIATOR PORT SAS ADDRESS field and TARGET PORT SAS ADDRESS field.
01h:	QUERY STP BUFFER	
02h:	DELETE SSP BUFFER	Discards the SSP or STP buffer context for the I_T nexus identified in the INITIATOR PORT SAS ADDRESS field and TARGET PORT SAS ADDRESS field.
03h:	DELETE STP BUFFER	
04h:	ABORT TASK	Discards any frames in the SSP buffer pair for the I_T nexus identified in the INITIATOR PORT SAS ADDRESS field and TARGET PORT SAS ADDRESS field that have a tag matchnig the value in the TASK TAG field. It also terminates the tag sentry [a] if one is active for the specified task.
05h:	SET PARAMETER	Set the S&F buffer configuration parameter specified in the PARAMETER SELECT field to the value specified in the PARAMETER VALUE field.
06h:	GET PARAMETER	Return the configuration parameter specified in the PARAMETER SELECT field in the RETURNED PARAMETER field in the S&F BUFFER CONTROL response.
07h - FFh:	Reserved	

[a] A tag sentry is established when an S&F buffer detects an error condition, such as ACK/NAK timeout, which requires the affected task to be aborted. The tag sentry discards any frames received with the affected tag until it receives an SMP S&F BUFFER CONTROL command specifying a buffer operation of ABORT TASK with a task tag matching the tag of an existing tag sentry.

Values of the PARAMETER SELECT field are shown in table 107.

Table 107 — PARAMETER SELECT field

Code	Mneumonic	Description
00h:	Reserved	
01h:	STP MAX FRAME SIZE	Sets (i.e., from the parameter value field) or returns (i.e., in the RETURNED PARAMETER field of the S&F BUFFER CONTROL response) the maximum STP frame size in bytes. The S&F expander shall set this field to the default value of 2008h at power-on.
02h - FFh:	Reserved	

The S&F buffer control response is shown in table 108.

Table 108 — S&F BUFFER CONTROL response

Byte\Bit	7	6	5	4	3	2	1	0
0	SMP FRAME TYPE (41h)							
1	FUNCTION (80h)							
2	FUNCTION RESULT							
3	RESPONSE LENGTH (03h)							
4	Reserved							
5	BUFFER PAIR STATUS				TASK STATUS			
6	IN BUFFER STATUS				OUT BUFFER STATUS			
7	INITIATOR PARTIAL CONNECTION STATUS				TARGET PARTIAL CONNECTION STATUS			
8	RETURNED PARAMETER							
9								
10	Reserved							
11								
12	(MSB) CRC							
15								(LSB)

The BUFFER PAIR STATUS field may return the values shown in table 109.

Table 109 — BUFFER PAIR STATUS field

Code	Name	Description
0h	Reserved	
1h	Buffer is not present	The S&F expander device does not have an S&F buffer pair for the specified I_T nexus.
2h	Buffer is present	The S&F expander device has an S&F buffer pair or S&F buffer context for the specified I_T nexus.
3h - Fh	Reserved	

The TASK STATUS field is valid when the TASK TAG field in the S&F BUFFER CONTROL request is non-zero. The task status may return the values shown in table 110.

Table 110 — TASK STATUS field

Code	Name	Description
0h	Reserved	
1h	FRAMES NOT PRESENT	There are no SSP frames in the S&F buffer pair with a task tag matching the TASK TAG field in the S&F BUFFER CONTROL request,
2h	FRAMES PRESENT	There are SSP frames in the S&F buffer pair with a task tag matching the TASK TAG field in the S&F BUFFER CONTROL request,
3h	SENTRY ACTIVE	The S&F buffer has an active sentry for frames with a task tag matching the TASK TAG field in the S&F BUFFER CONTROL request, Any frames received for the specified task are discarded.
4h - Fh	Reserved	

[a] An SMP initiator port may deactivate the sentry for the specified task by sending an SMP S&F BUFFER CONTROL command specifying a buffer operation of ABORT TASK and the same TASK TAG field used for this query.

The IN BUFFER STATUS field may return the values shown in table 111.

Table 111 — IN BUFFER STATUS field

Code	Name	Description
0h	Reserved	
1h	FULL	The specified S&F buffer has no space available for more SSP frames or SATA frames.
2h	EMPTY	The specified S&F buffer has no SSP frames or SATA frames.
3h	NOT EMPTY	There are SSP frames or SATA frames in the specified S&F buffer, and space for more.
4h - Fh	Reserved	

The OUT BUFFER STATUS field may return the values shown in table 112.

Table 112 — OUT BUFFER STATUS field

Code	Name	Description
0h	Reserved	
1h	FULL	The specified S&F buffer has no space available for more SSP frames or SATA frames.
2h	EMPTY	The specified S&F buffer has no SSP frames or SATA frames.
3h	NOT EMPTY	There are SSP frames or SATA frames in the specified S&F buffer, and space for more.
4h - Fh	Reserved	

The INITIATOR PARTIAL CONNECTION STATUS field may return the values shown in table 9.

Table 113 — INITIATOR PARTIAL CONNECTION STATUS field

Code	Name	Description
0h	Reserved	
1h	NOT CONNECTED	The specified S&F buffer pair does not have a partial connection or pending partial connection to a SAS initiator port.
2h	CONNECTED	The specified S&F buffer pair has a partial connection or pending partial connection to a SAS initiator port.
3h	CONNECTION PENDING	The specified S&F buffer pair has a pending partial connection a SAS initiator port.
4h	CREDIT BLOCKED	The specified S&F buffer pair is unable to accept SSP frames or SATA frames from the SAS initiator port, has reported CREDIT BLOCKED, and is waiting for the SAS initator port to initiate closing the connection.
5h - Fh	Reserved	

The TARGET PARTIAL CONNECTION STATUS field may return the values shown in table 114.

Table 114 — INITIATOR PARTIAL CONNECTION STATUS field

Code	Name	Description
0h	Reserved	
1h	NOT CONNECTED	The specified S&F buffer pair does not have a partial connection or pending partial connection to a SAS initiator port.
2h	CONNECTED	The specified S&F buffer pair has a partial connection or pending partial connection to a SAS initiator port.
3h	CONNECTION PENDING	The specified S&F buffer pair has a pending partial connection a SAS initiator port.
4h	CREDIT BLOCKED	The specified S&F buffer pair is unable to accept SSP frames or SATA frames from the SAS initiator port, has reported CREDIT BLOCKED, and is waiting for the SAS initator port to initiate closing the connection.
5h - Fh	Reserved	

Editor's Note 15: Other information that may be useful to include in the S&F BUFFER CONTROL response for a buffer operation of QUERY SSP BUFFER or QUERY STP BUFFER includes: Information about # frames in buffer, buffer capacity in full-frames, buffer pairs available in free buffer pool, total number of buffer pairs in expander (being used plus in free pool), frame credits (received and transmitted, w.r.t. init and target).