BREA Technologies, Inc.
14902 Mesita Drive
Houston, TX 77083-3209
P: (281) 530-3063
F: (281) 988-0358
BillG@breatech.com

October 29, 2002

To: T10 Technical Committee
From: Bill Galloway
Subj: SAS Simple Relative Offset

It does not make sense to add a sequence count to SAS to detect lost data frames. An I_T_L_Q based sequence count will be difficult to handle across wide links. Relative Offset will be needed in future versions of the SAS specification and a very simple form of Relative Offset can be used to detect lost data frames. This proposal makes Relative Offset mandatory for all SAS devices but requires all data to be transmitted in order with no gaps. By requiring the data to be transmitted in order with no gaps, the Relative Offset field becomes a "byte" sequence count instead of a frame sequence count.

The following changes are required to implement this proposal:

9.2.1 SSP frame format
Table 72 defines the SSP frame format.

Table 72 — SSP frame format

Byte	7	6	5	4	3	2	1	0
0	FRAME TYPE							
1	(MSB)	HASHED DESTINATION SAS ADDRESS						
3								(LSB)
4	Reserved							
5	(MSB)	HASHED SOURCE SAS ADDRESS						
7								(LSB)
8	Reserved							
9	Reserved							
10	Reserved						TIMEOUT	Reserved
11	Reserved						NUMBER OF FILL BYTES	
12	Reserved							
13	Reserved							
15								
16	(MSB)	TAG						
17								(LSB)
18	(MSB)	TARGET PORT TRANSFER TAG						
19								(LSB)
20	(MSB)	RELATIVE OFFSET						
23								(LSB)
24		INFORMATION UNIT						
m								
	Fill bytes, if needed							
n - 3	(MSB)	CRC						
n								(LSB)

Add paragraph:

For DATA frames, the RELATIVE OFFSET field indicates the application client buffer offset as described by SAM-3. The relative offset shall be a multiple of four (i.e., each DATA frame shall begin on a word boundary). For all other frame types, this field shall be ignored.
This field shall be zero for the first read DATA frame and first write DATA frame of a command.

9.2.2.3 XFER_RDY information unit
Table 78 defines the transfer ready IU. The XFER_RDY frame is sent by a target port to request write data from the initiator port.

Table 78 — XFER_RDY information unit

Byte	7	6	5	4	3	2	1	0
0	(MSB)	RELATIVE OFFSET						
3								(LSB)
4	(MSB)	WRITE DATA LENGTH						
7								(LSB)
8		Reserved						
11								

Add paragraph:

The RELATIVE OFFSET field indicates the initial application client buffer offset of the write data the initiator port may transmit to the logical unit (using DATA frames). The relative offset shall be a multiple of four (i.e., each DATA frame shall begin on a word boundary). This field shall be zero for the first XFER_RDY frame of a command unless the FIRST BURST SIZE field in the Disconnect-Reconnect mode page is not zero (see 10.1.1.1.5).

The WRITE DATA LENGTH field indicates how many bytes of write data the initiator port may transmit to the logical unit (using DATA frames) starting at the relative offset. If the value in the MAXIMUM BURST SIZE field in the Disconnect-Reconnect mode page is not zero, the value in the WRITE DATA LENGTH field is constrained by the value in the MAXIMUM BURST SIZE field (see 10.1.1.1.4).

9.2.3.3 XFER_RDY frame rules
The target port shall only transmit XFER_RDY frames in response to a COMMAND frame for a write or bidirectional command.

The target port shall not transmit an XFER_RDY frame for a given I_T_L_Q until it has received all write DATA frames for the previous XFER_RDY frame, if any, and has provided link layer acknowledgement for all of the previous write DATA frames for that I_T_L_Q. The target port shall not transmit an XFER_RDY frame for a given I_T_L_Q after it has transmitted a RESPONSE frame that terminates the task for that I_T_L_Q (e.g., a RESPONSE frame with STATUS or a RESPONSE frame for an ABORT task management request).

The initial XFER_RDY frame for a given command shall set the relative offset to the value of the FIRST BURST SIZE field in the Disconnect-Reconnect mode page (see 10.1.1.1.5). If any more XFER_RDY frames are required, the relative offset field shall be set to the value of the previous XFER_RDY frame relative offset plus the previous XFER_RDY frame write data length.

9.2.3.4 DATA frame rules
The initiator port shall only transmit DATA frames in response to a XFER_RDY frame or an implied XFER_RDY frame as a result of the FIRST BURST SIZE field.

The target port shall only transmit DATA frames in response to a COMMAND frame for a read or bidirectional command.

A wide port shall only transmit DATA frames for a given I_T_L_Q on one link at a time. The wide port may switch links for DATA frames once all of the previous DATA frames have been acknowledged by the link layer. Read DATA frames and write DATA frames for the same I_T_L_Q may be transmitted simultaneously and may be on the same or different physical links.

The initiator port shall not transmit a DATA frame for a given I_T_L_Q after it has sent a TASK frame that terminates the task for that I_T_L_Q.

The target port shall not transmit a DATA frame for a given I_T_L_Q after it has transmitted a RESPONSE frame that terminates the task for that I_T_L_Q (e.g., a RESPONSE frame with STATUS or a RESPONSE frame for an ABORT task management request).

The initial read DATA frame for a given command shall set the relative offset to zero. If any more read DATA frames are required, the relative offset field shall be set to the value of the previous read DATA frame relative offset plus the previous read DATA frame data length.

The initial write DATA frame for a given command shall set the relative offset to zero. If any more write DATA frames are required, the relative offset field shall be set to the value of the previous write DATA frame relative offset plus the previous write DATA frame data length.

9.2.5.2 Target port error handling

…

If a target port receives a DATA frame with more write data than expected, it shall discard the frame and terminate the command with a CHECK CONDITION status with a sense key of ABORTED COMMAND and an additional sense code of TOO MUCH WRITE DATA.

If a target port receives a DATA frame with a relative offset that was not expected, it shall discard the frame and terminate the command with a CHECK CONDITION status with a sense key of ABORTED COMMAND and an additional sense code of RELATIVE OFFSET ERROR (new 4B / xx).

9.2.5.3 Initiator port error handling

…

If an initiator port receives an XFER_RDY frame requesting more write data than expected, it shall transmit an ABORT TASK to abort the command.

If an initiator port receives an XFER_RDY frame with a relative offset that was not expected, it shall transmit an ABORT TASK to abort the command.

…

If an initiator port receives a DATA frame with more read data than expected, it shall discard the frame and transmit an ABORT TASK to abort the command. It may receive a RESPONSE for the command before being able to abort it.

If an initiator port receives a DATA frame with a relative offset that was not expected, it shall discard the frame and transmit an ABORT TASK to abort the command. It may receive a RESPONSE for the command before being able to abort it.

9.2.6.2.4 ST_ISF3:Prepare_Send_Data_Out state

9.2.6.2.4.1 State description

This state shall construct a DATA frame. This state shall include the following in the frame (these were received either from the SCSI initiator device's application layer or included in an XFER_RDY Arrived parameter):

a) tag; ~~and,~~
b) target port transfer tag; and,
c) relative offset.

This state shall generate and include the following in the frame:
a) information unit type;
b) hashed destination SAS address;
c) hashed source SAS address;
d) timeout bit set to zero;
e) number of fill bytes; and,
f) ~~f)~~ fill bytes.

9.2.6.2.4.2 Transition ST_ISF3:Prepare_Send_Data_Out to ST_ISF1:Send_Frame
This transition shall occur after the ST_ISF3:Prepare_Send_Data_Out state has constructed a DATA frame.

9.2.6.2.5 ST_IRD1:Receive_Data_In state
9.2.6.2.5.1 State description
The ST_IRD state machine shall be initiated when a Data-In Arrived parameter is received from the ST_IFR (frame router) state machine.

This state shall check the length and relative offset of the DATA information unit. If the length of the information unit exceeds the amount of data remaining to be transfered for the data-in command, or the relative offset is not the expected offset, then this state shall send a Delivery Failure parameter to the ST_IPR:Process_Received_Response state. This state machine shall terminate after sending the parameter.

9.2.6.3.2 ST_TFR1:Target_Frame_Router state
…
If the information unit type is DATA, then this state shall check that the target port transfer tag corresponds to a target port transfer tag sent in a previous XFER_RDY frame. If the target port transfer tag in the frame does not correspond, then this state shall send a SCSI Command Received protocol service indication to the SCSI target device's application layer indicating a Service Delivery Subsystem Failure. This indication shall include the tag.

If the information unit type is DATA, then this state shall check that the relative offset corresponds to the expected relative offset. If the relative offset in the frame does not correspond, then this state shall send a SCSI Command Received protocol service indication to the SCSI target device's application layer indicating a Service Delivery Subsystem Failure. This indication shall include the tag.

9.2.6.3.5 ST_TTS3:Prepare_Send_Data_In state

9.2.6.3.5.1 State description
This state fetches the data from the Device Server Buffer and constructs a DATA frame. This state shall use the tag received from the ST_TTS2:Send_Frame state to construct the frame.

This state shall generate the following to be used in the frame:
a) information unit type;
b) hashed destination SAS address;
c) hashed source SAS address;
d) timeout bit set to zero;
e) number of fill bytes; ~~and,~~
f) relative offset; and,

g) fill bytes.

9.2.6.3.6 ST_TTS4:Receive_Data_Out state

9.2.6.3.6.1 State description

…

If this state was entered as the result of receiving a DATA frame from the ST_TFR state machine, then this state shall check the length of the data. If the length of the data exceeds that specified by the XFER_RDY frame that requested the data, then this state shall send a Data-Out Received data-out delivery service confirmation to the SCSI target device's application layer with a delivery result argument of DELIVERY SUCCESSFUL (TOO MUCH WRITE DATA). This confirmation shall include the tag. The ST_TTS state machine shall terminate after sending the confirmation.

If this state was entered as the result of receiving a DATA frame from the ST_TFR state machine, then this state shall check the relative offset. If the relative offset was not expected, then this state shall send a Data-Out Received data-out delivery service confirmation to the SCSI target device's application layer with a delivery result argument of DELIVERY SUCCESSFUL (RELATIVE OFFSET ERROR). This confirmation shall include the tag. The ST_TTS state machine shall terminate after sending the confirmation.

9.2.6.3.7 ST_TTS5:Prepare_XFER_RDY state

9.2.6.3.7.1 State description

This state shall construct an XFER_RDY frame. This state shall use the following received from the Receive Data-Out data-out delivery service request to construct the frame:

a) tag;
b) target port transfer tag; ~~and,~~
f) relative offset; and,
c) write data length.

This state shall generate the following to be used in the frame:

a) information unit type;
b) hashed destination SAS address;
c) hashed source SAS address;
d) timeout bit set to zero; and,
e) number of fill bytes;