**T10/07-391 revision 0**

Date: September 13, 2007

To: T10 Committee (SCSI)

From: George Penokie (IBM)

Subject: SAS-2: Limiting SAS Target response to OPEN_REJECT (RETRY)

# 1 Overview

In a SAS topology with a large imbalance between the number of initiators and targets (e.g., one or two initiators with 50 or more targets) with a heavy workload it is not unusual for the minority SAS device to become bandwidth limited. When this happens the bandwidth limited SAS device will start issuing OPEN_REJECT (RETRY) to most open attempts. In response to the OPEN_REJECT (RETRY) many SAS devices will immediately issue another OPEN. This causes a flood of OPEN/OPEN_REJECT (RETRY)s which accomplishes nothing except to add congestion to an already congested topology.

This proposal brings the DISCONNECT TIME LIMIT field of the Disconnect-Reconnect mode page into SAS-2.

# 2 Proposed SAS-2 changes

### 10.2.7.2 Disconnect-Reconnect mode page

### 10.2.7.2.1 Disconnect-Reconnect mode page overview

The Disconnect-Reconnect mode page (see SPC-4) provides the application client the means to tune the performance of a service delivery subsystem. Table 1 defines the parameters which are applicable to SSP. If any field in the Disconnect-Reconnect mode page is not implemented, the value assumed for the functionality of the field shall be zero (i.e., as if the field in the mode page is implemented and the field is set to zero).

The application client sends the values in the fields to be used by the device server to control the SSP connections by means of a MODE SELECT command. The device server shall then communicate the field values to the SSP target port. The field values are communicated from the device server to the SSP target port in a vendor-specific manner.

SAS devices shall only use the parameter fields defined below in this subclause. If any other fields within the Disconnect-Reconnect mode page of the MODE SELECT command contain a non-zero value, the device server shall terminate the MODE SELECT command with CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and the additional sense code set to INVALID FIELD IN PARAMETER LIST.

**Table 1 — Disconnect-Reconnect mode page for SSP**

| Byte\Bit | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
|---|---|---|---|---|---|---|---|---|
| 0 | PS | SPF (0b) | PAGE CODE (02h) | | | | | |
| 1 | PAGE LENGTH (0Eh) | | | | | | | |
| 2 | Reserved | | | | | | | |
| 3 | Reserved | | | | | | | |
| 4 | (MSB) | BUS INACTIVITY TIME LIMIT | | | | | | |
| 5 | | | | | | | | (LSB) |
| 6 | | DISCONNECT TIME LIMIT | | | | | | |
| 7 | | | | | | | | |
| 8 | (MSB) | MAXIMUM CONNECT TIME LIMIT | | | | | | |
| 9 | | | | | | | | (LSB) |
| 10 | (MSB) | MAXIMUM BURST SIZE | | | | | | |
| 11 | | | | | | | | (LSB) |
| 12 | Reserved | | | | | | | |
| 13 | Reserved | | | | | | | |
| 14 | (MSB) | FIRST BURST SIZE | | | | | | |
| 15 | | | | | | | | (LSB) |

The PARAMETERS SAVEABLE (PS) bit is defined in SPC-4.

The SUBPAGE FORMAT (SPF) bit shall be set to zero for access to this mode page.

The PAGE CODE (PS) field shall be set to 02h.

The PAGE LENGTH field shall be set to 0Eh.

The BUS INACTIVITY TIME LIMIT field is defined in 10.2.7.2.2.

The DISCONNECT TIME LIMIT field is defined in 10.2.7.2.3.

The MAXIMUM CONNECT TIME LIMIT field is defined in 10.2.7.2.4.

The MAXIMUM BURST SIZE field is defined in 10.2.7.2.5.

The FIRST BURST SIZE field is defined in 10.2.7.2.6.

**10.2.7.2.2** BUS INACTIVITY TIME LIMIT **field**

The value in the BUS INACTIVITY TIME LIMIT field contains the maximum time in 100 μs increments that an SSP target port is permitted to maintain a connection (see 4.1.12) without transferring a frame to the SSP initiator port. When this time is exceeded, the SSP target port shall prepare to close the connection (i.e., by requesting

to have the link layer transmit DONE). This value may be rounded as defined in SPC-4. A value of 0000h in this field specifies that there is no bus inactivity time limit. The bus inactivity time limit is enforced by the port layer (see 8.2.3).

### 10.2.7.2.3 DISCONNECT TIME LIMIT field

The DISCONNECT TIME LIMIT field specifies the minimum time in 100 ms increments that the target port shall wait between a receiving an OPEN_REJECT and issuing an OPEN address frame. This value may be rounded as defined in SPC-4. A value of 0000h specifies that there is no disconnect time limit. The disconnect time limit is enforced by the port layer (see 8.2.3).

### 10.2.7.2.4 MAXIMUM CONNECT TIME LIMIT field

The value in the MAXIMUM CONNECT TIME LIMIT field contains the maximum duration of a connection (see 4.1.12) in 100 $\mu$s increments (e.g., a value of 0001h in this field means that the time is less than or equal to 100 $\mu$s and a value of 0002h in this field means that the time is less than or equal to 200 $\mu$s). When this time is exceeded, the SSP target port shall prepare to close the connection. If an SSP target port is transferring a frame when the maximum connection time limit is exceeded, the SSP target port shall complete transfer of the frame before preparing to close the connection. This value may be rounded as defined in SPC-4. A value of 0000h in this field specifies that there is no maximum connection time limit. The maximum connection time limit is enforced by the port layer (see 8.2.3).

### 10.2.7.2.5 MAXIMUM BURST SIZE field

For read data, the value in the MAXIMUM BURST SIZE field contains the maximum amount of data that is transferred during a connection by an SSP target port per I_T_L_Q nexus without transferring at least one frame for a different I_T_L_Q nexus. If the SSP target port:

   a)  has read data to transfer for only one I_T_L_Q nexus, and
   b)  has no requests to transfer write data for any I_T_L_Q nexus;

then the SSP target port shall prepare to close the connection after the amount of data specified by the MAXIMUM BURST SIZE field is transferred to the SSP initiator port.

For write data, the value shall specify the maximum amount of data that an SSP target port requests via a single XFER_RDY frame (see 9.2.2.3).

This value shall be specified in 512-byte increments (e.g., a value of one in this field means that the number of bytes transferred to the SSP initiator port for the nexus is less than or equal to 512 and a value of two in this field means that the number of bytes transferred to the SSP initiator port for the nexus is less than or equal to 1 024). A value of zero in this field shall specify that there is no maximum burst size.

In terms of the SCSI transport protocol services (see 10.2.1), the device server shall limit the Request Byte Count argument to the Receive Data-Out () protocol service and the Send Data-In () protocol service to the amount specified in this field.

### 10.2.7.2.6 FIRST BURST SIZE field

If the ENABLE FIRST BURST field in the COMMAND frame is set to zero, the FIRST BURST SIZE field is ignored.

If the ENABLE FIRST BURST field in the COMMAND frame is set to one, the value in the FIRST BURST SIZE field contains the maximum amount of write data in 512-byte increments that may be sent by the SSP initiator port to the SSP target port without having to receive an XFER_RDY frame (see 9.2.2.3) from the SSP target port (e.g., a value of one in this field means that the number of bytes transferred by the SSP initiator port is less than or equal to 512 and a value of two in this field means that the number of bytes transferred by the SSP initiator port is less than or equal to 1 024).

Specifying a non-zero value in the FIRST BURST SIZE field is equivalent to an implicit XFER_RDY frame for each command requiring write data where the WRITE DATA LENGTH field of the XFER_RDY frame is set to 512 times the value of the FIRST BURST SIZE field.

The rules for data transferred using the value in the FIRST BURST SIZE field are the same as those used for data transferred for an XFER_RDY frame (i.e., the number of bytes transferred using the value in the FIRST BURST SIZE field is as if that number of bytes was requested by an XFER_RDY frame).

If the amount of data to be transferred for the command is less than the amount of data specified by the FIRST BURST SIZE field, the SSP target port shall not transmit an XFER_RDY frame for the command. If the amount of data to be transferred for the command is greater than the amount of data specified by the FIRST BURST SIZE field, the SSP target port shall transmit an XFER_RDY frame after it has received all of the data specified by the FIRST BURST SIZE field from the SSP initiator port. All data for the command is not required to be transferred during the same connection in which the command is transferred.

A value of zero in this field shall specify that there is no first burst size (i.e., an SSP initiator port shall transmit no write DATA frames to the SSP target port before receiving an XFER_RDY frame).

The first burst size is handled by the SCSI transport protocol services (see 10.2.1) and the SSP transport layer (see 9.2.6).