First Burst Simplification

Bob Nixon

bob.nixon@emulex.com

Revision 1 Incorporate requests from SAS meeting 7-8 July 2003.

1 Overview

It is difficult for initiators in multi-initiator SCSI domains to properly support first burst transfers because the transfer size may change unpredictably; however, a SCSI initiator will fail if it uses an incorrect first burst size. This proposal presents the issue and an approach to simplify support for first burst.

2 Issue

SAS 1.0 specifies support for the FIRST BURST SIZE field in the Disconnect - Reconnect mode page. Parameters in the Disconnect - Reconnect mode page pertain to a specific target port regardless of the initiator port or target logical unit. As these parameters are not specifically read-only in either SPC-3 or SAS, an initiator may alter them by a MODE SELECT command. Such a change immediately affects all other initiators using any logical unit via the same port In a multi-initiator SCSI domain such as SAS, there is no reason to expect various initiators to coordinate setting of first burst. In fact, as this value is provided explicitly for transport tuning, each initiator may see some advantage (a hair better performance, an increase in stability) to forcing its own optimal value (or zero) for first burst.

The result of an initiator not handling a change of the first burst size is that the initiator will experience write errors, so a viable initiator has to design to accomodate unpredictable changes of the first burst size.

Designing an initiator for unpredictable changes of the first burst size requires active assurance that the setting of first burst is at least known by the initiator (if not actually set to a preferred value) before every comand is sent. Since this involves awareness of unit attentions and sending of MODE SENSE/SELECT commands, the initiator transport layer becomes functionally entangled with the initiator application(s). This is bad, as evidenced by the acceptance of 02-403 (Elliot), which attempted to untangle exactly this issue.

In SAS, setting first burst to zero disables the feature, but this does not resolve the problem: Another initiator may at any time set it nonzero. In fact, it was pointed out to me that a transport has to support any possible setting of first burst size in order to deliver a MODE SELECT to disable it.

SPC-3 (thanks to 02-403) hints that using a setting of FIRST BURST SIZE (e.g., zero) to disable first burst is an incomplete implementation...in the specification of the FIRST BURST SIZE field it says "SCSI transport protocols supporting this field shall provide an additional mechanism to enable and disable the first burst function." SAS does not provide such an additional mechanism. By comparison, FCP-2 provides an additional mechanism by means of a flag during login that enables or disables first burst on a per-initiator basis. (A later interoperability profile for FCP-2 required unconditionally disabling first burst.)

This proposal defines an additional mechanism for SAS initiators to enable and disable the first burst function. An initiator may then choose to disable first burst to avoid issues arising from its full generality (e.g., for ports that allow initiators to change its value), while retaining first burst when its value exceeds its cost of implementation (e.g., for ports that do not allow initiators to change its value).

3 Proposed resolution

3.1 Technique

Provide an "additional mechanism" to disable first burst on a per initiator basis by reassigning a currently reserved bit in the SSP COMMAND IU as the FB bit. The FB bit shall indicate whether the target is to honor (bit=0) or ignore (bit=1) first burst size for the command.

3.2 Legacy device compatibility

The choice of bit values causes legacy initiators, for which the FB bit is reserved, to operate with targets in accord with SAS 1.0 (i.e., operation in accord with the FIRST BURST SIZE field is required).

Initiators that support setting the FB bit to one need to assure they do not do so in an SSP COMMAND IU sent to a legacy SSP target device. They may do so by examining the response to an INQUIRY command before sending commands that include data out. An SSP target that returns standard INQUIRY data that does not include a version descriptor for SAS 1.1 or later should be presumed not to support the FB bit setting to one.

4 Instructions to editor

In subclause 9.2.2.1, table 24, add a flag field in byte 9 bit 7 labeled “FB”.

In subclause 9.2.2.1, add a paragraph “The FB bit directs the target SSP transport state machine to attend or ignore the FIRST BURST SIZE field in the Disconnect - Reconnect mode page for the target port (see 10.2.6.1.5). If the value of the FIRST BURST SIZE field in the Disconnect - Reconnect mode page for the target port is zero, a target device may ignore the FB field. If the value in the FB bit is zero, the initiator and target SSP transport state machines shall transfer any data in accord with the value of the FIRST BURST SIZE field in the Disconnect - Reconnect mode page for the target port. If the value in the FB bit is one, the initiator and target SSP transport state machines shall transfer any data as though the value of the FIRST BURST SIZE field in the Disconnect - Reconnect mode page for the target port were zero (i.e., first burst data transfer shall be disabled for the command).

In subclause 9.2.2.3 second paragraph, eliminate the last sentence.

In subclause 9.2.2.3 third paragraph, change the first sentence to “In the initial XFER_RDY frame for a given command, the SSP target port shall set the requested offset to the number of bytes of first burst data enabled for the command (see 9.2.2.1)“. Join the third paragraph to the second paragraph.

In subclause 9.2.2.4 list item b, change “if the FIRST BURST SIZE field in the Disconnect-Reconnect mode page is not zero (see 10.2.6.1.5).“ to “if first burst data is enabled for the command (see 9.2.2.1)“.

In subclause 9.2.6.2.2.2.1, change the paragraph “If the request is for a data-out command, then the request also includes the number of bytes for the first burst size for the logical unit“ to “If the request is for a data-out command, then the request also includes the number of bytes of first burst data enabled for the command (see 9.2.2.1)“.

In subclause 10.2.6.1.5 first paragraph, prefix the first sentence with “For commands sent with first burst data disabled (see 9.2.2.1), the First Burst Data Size shall be ignored. For commands sent with first burst data enabled,”.