

Date: September 11, 1996

TO: X3T10 Committee (SCSI)

SUBJECT: SAM/SPC Link event timeout and the PLDA profile

The SCSI working group, on September 11, discussed the original r0 proposal at some length. The discussion concluded that the concept of a ULP timer for the target was a good idea. When specified in any form, it will probably be in SPC-2 and SAM-2, unless it is accepted as a public review comment for SPC-1 (unlikely from the sentiment expressed in the meeting).

The working group asked for some changes and too distinguish the interconnect event from one concept in the PLDA document which has to do with defining next sequence timeouts during a command.

It appears that the PLDA considers this timer in effect from the time a target sends an FCP sequence that transfers sequence initiative to the initiator (e.g., FCP_XFR_RDY) until the next sequence from that initiator for that exchange. This was not accepted as valid SCSI behavior, so some change to PLDA may be required as pointed out by Bob Snively. The initiator command ULP timer was considered the correct timer for the command rather than the PLDA timer.

The working group felt that the target should not time an initiator's individual response times during a command just as the initiator does not time individual target response times. The initiator normally has an initiator-based ULP timer running for each command that is specific to the command, set by the peripheral driver, and varies widely by command.

The working group felt that holding off aborting tasks for a minimum time was a good idea, but they wanted to extend the concept to specify a maximum time after which an initiator knows that its commands have been aborted. That is, this proposal is changed to include a minimum wait time and a maximum delay time. This will cause a complete rewrite of the SPC text since additional bytes will be needed in the page. There is no room left for adding the second field. This change would also permit less encoding of the timer value since more space can be allocated to the U/M and time fields.

The discussion also included CA and ACA active states in the target. If the timer goes off, all commands are aborted with no status. CA or ACA states would be removed at the same time. Persistent reservations would remain in effect if they would be kept over power cycles, otherwise any persistent reservations would be reset also.

The working group also discussed whether this timer was a link level function, a protocol specific function, or a ULP function. It was decided that it is appropriate at the ULP level and it applies to all protocols when reworded to encompass a broader set of events. For example, a reselection timeout event in SIP was considered a candidate event to start the timer.

After review of these proposals, the committee will reconsider them in November, 1996. The proposal may also require some additional explanation in SAM-2 since we are dealing with an architectural concept and not a protocol specific item. When specific changes are identified, an r2 version will be produced.

As with all standards, at implementation time, some items are discovered that were not considered when developing the base standard. This timeout mechanism need to be added to SPC and the concept possibly added to a future version of SAM so that the PLDA can continue to be a profile document and not a specification document. No other known SCSI-specific functions need to be defined in SCSI standards to support the PLDA profile requirements.

Gary Stephens

1 Attachment

From 240r0

The PLDA profile, rev 1.4, calls out a timer that permits a target to abort all tasks from an initiator if it does not make some specific contact with the target after a significant transport level event. For FC-AL, this is a link initialization event.

In this case, either the initiator or target AL_PA or its native address identifier may change. If the target just holds on to uncompleted tasks, it will eventually run out of resources. The PLDA defines a resource recovery timer for SCSI, independent of other FC-4s. Since this is not a Fibre Channel or FC-AL timer, it is a SCSI architecture and specification problem.

Historically, we have other timers in the Disconnect-reconnect mode page that specify target timeouts for other events. In particular, the Bus inactivity limit, has been used to escape from a hung bus problem for a missing REQ or ACK signal. In this case, only one task is aborted as a result of an unexpected disconnect event.

However, the PLDA timeout permits all tasks to be aborted and resources recovered for failure of an initiator to contact the target. With Fibre Channel and SSA, the number of initiators can increase dramatically and link events in a fabric, private loop, or web are outside the control of the two SCSI devices.

Some definition of a timeout needs to be added to SPC, probably in the Disconnect-reconnect page, byte 13 to specify a timeout. PLDA compliant devices will set this time to the required time, and may or may not permit it to be changed as with other mode page fields. The PLDA will then specify the value required for this field and not impose an non-architected process on these devices.

In Hawaii, at the Fibre Channel PLDA meeting, it was felt that this was a SCSI problem not specific to the PLDA, but a general problem that affects all transport levels. Therefore, the following is the proposed new field for the mode page in SPC.

Beginning of proposed SPC text

Bit	7	6	5	4	3	2	1	0
Byte	RESScale		RESValue					
13	RESScale		RESValue					

The RESScale field specifies the unit of measure of the value in the RESValue field. The units of measure codes for RESScale are defined in table x. A value of 00b is specified to have no meaning for backward compatibility with existing target implementations. If byte 13 is zero (i.e., it appears as reserved value) the target is not specifying a timeout value and the action of the target shall be vendor unique.

Table x - RESScale field values

Value	Unit of measure
00b	No time; RESValue not meaningful
01b	microseconds * 16
10b	milliseconds * 16
11b	seconds * 16

The RESValue field specifies a coarse grained amount of time that a target shall wait after notification of a possible resetting interconnect event within which each initiator shall connect with the target and reestablish that it is the same initiator and intends to continue its tasks. The range of valid values for the RESValue field is 0-63, where zero means zero times the time units specified in the RESScale field when RESScale is not 00b. The means of contact is protocol specific.

The RES value is calculated as follows when RESScale is not 00b:

$$\text{RESValue} * 16 * \text{unit of measure}$$

This structure permits timeout values from 0 microseconds to about 16.8 minutes (1008 seconds) where the maximum value in microseconds is slightly more than one millisecond and the maximum value in milliseconds is slightly more than one second.

If an initiator does not make the proper connection with a target the target shall abort all tasks for that initiator. If other initiators make the proper connection within the specified time period, all uncompleted tasks shall attempt to continue to completion.

End of proposed SPC text

End 240r0 proposal text