

To: INCITS T10 Committee
 From: Paul Entzel, Quantum
 Date: 8 May 2006
 Document: T10/06-236r0
 Subject: FCP-4 Fix race condition between REC ACC and FCP_XFER_RDY

1 Revision History

Revision 0:
 Posted to the T10 web site on 8 May 2006.

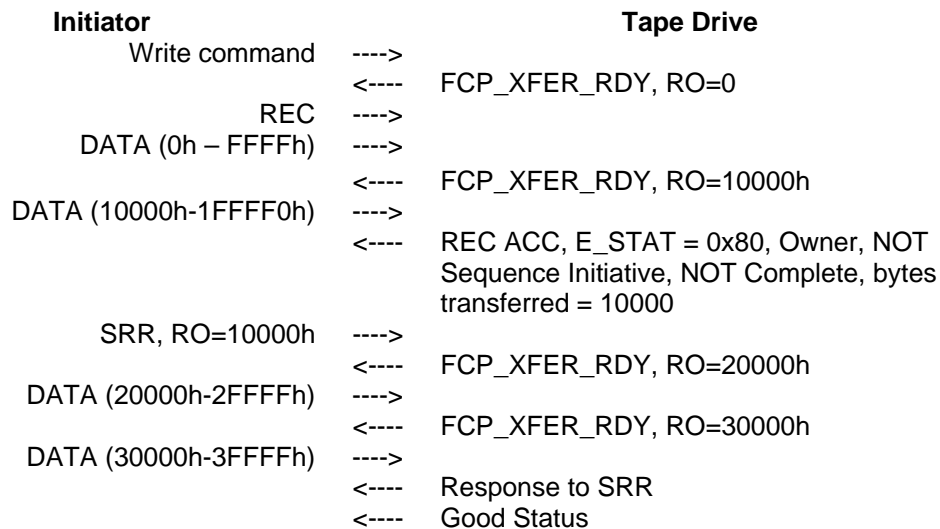
2 General

In Figure C.5 there is a paragraph that appears to have been included to deal with a race condition between the ACC for a REC and a FCP_XFER_RDY. The paragraph states:

"Wait REC_TOV*. If FCP_XFER_RDY is returned, continue with the Exchange. (ACC to REC arrived before FCP_XFER_RDY, out of order). Otherwise continue recovery."

The diagram shows this wait period beginning at the receipt of the ACC for an REC.

Recently, we have experienced a problem that is similar to this in a Fibre Channel tape drive that was performing extended recovery operations on the medium during a write operation. The buffer filled and a new write command was sent when there was no room in the buffer for the data associated with the command. The HBA polled with an REC every few seconds until at last there was room in the buffer. The REC and the FCP_XFER_RDY were timed so closely that they may have actually passed each other in the Fibre. Here is a ladder diagram of the condition:



This isn't exactly covered by the text in table C.5, since the FCP_XFER_RDY was received by the initiator port before the ACC for the REC, not after. The HBA has not violated the standard since there is no requirement in the normative part of the standard governing this case. However, the problem could have been avoided had the HBA ignored the REC ACC payload and not performed a retry based on the fact that it had received traffic for the exchange in question since it had sent the REC. This proposal will

attempt to correct this condition by adding requirements within the normative clauses in FCP-4 to cover the case already described in table C.5 and expanding it to cover the case described by the ladder diagram above.

3 Reference

T10/FCP-4 revision 0.

4 Changes to FCP-4

12.4.1.4 FCP_XFER_RDY IU recovery

This procedure shall be used only by FCP devices that have agreed to Sequence level recovery.

If the ACC for an REC indicates that an FCP_XFER_RDY IU was sent by the target FCP_Port (i.e., by indicating that the initiator FCP_Port holds Sequence Initiative, that all bytes were not transferred, and that the Exchange is not complete), but not received by the initiator FCP_Port, the initiator FCP_Port shall issue an SRR in a new Exchange to request retransmission of the FCP_XFER_RDY IU. [To avoid race conditions between the ACC for an REC and an FCP_XFER_RDY IU, the initiator FCP_Port shall wait REC_TOV after receiving the ACC for an REC before sending an SRR to recover for a lost FCP_XFER_RDY IU. If the initiator FCP_Port receives an FCP_XFER_RDY IU for an exchange after sending an REC or within REC_TOV time after receiving an ACC for an REC, it shall continue normal processing of the FCP I/O operation for that exchange and ignore the contents of the ACC.](#)

The target FCP_Port shall first transmit the ACC for the SRR and then shall retransmit the FCP_XFER_RDY IU in a new Sequence containing the same Relative Offset as the originally transmitted FCP_XFER_RDY IU. After the FCP_XFER_RDY IU is successfully received, the FCP I/O operation continues normally.

For examples of this type of recovery, see figure C.5 and figure C.6.