Response to HP INCITS No Vote on FCP-2
David Peterson, Cisco Systems
(02-440r0)

Resolution for the comments in *italics* below are provided inline starting with "Resolution".

HP INCITS No Vote on FCP-2
19 July 2002 - INCITS No vote
9 September 2002 - added issue #7 from Santosh

Technical content from Santosh Rao, Mallikarjun Chandalapaka, and Rob Elliott

IBM's initial NO vote in INCITS letter ballot 4450 on FCP-2 provides the opportunity to correct a number of issues have come to our attention since T10 letter ballot and INCITS public review. Since these seem to reflect errors in the proposed standard, it would be prudent to fix them now rather than introduce an FCP-2 amendment. Since they are errors, not new features, it does not seem appropriate to defer them until FCP-3.

HP Issue #1
1. Veritas' FCP-2 public review comment about backup application/tape drive problems unearthed problems with the clearing effects of protocol events that affect more SCSI standards than FCP-2. Since that time, T10 has worked to address the issues for all SCSI protocols, not just FCP-2. As IBM's INCITS no vote indicates, the suggested solution for FCP-2 and SSC-2 developed by T10 was incomplete. Rather than the mode page correction proposed by IBM, we suggest that all SCSI clearing effects be removed from FCP-2 and the issue be left to other T10 standards:

a) In section 4.9 "Clearing effects of task management, FCP, FC-FS, and FC-AL-2 actions", delete these rows and associated notes from Table 4 "Clearing effects of link related functions" and Table 5 "Clearing effects of initiator actions":

- *Open Tasks (FCP Exchanges) Aborted*
- *Target mode page parameters restored from saved pages*
- *Pre-existing ACA, Unit Attention, and Deferred error conditions cleared*
- *Device reservations*
- *Persistent device reservations*
- *Prevent Allow Medium Removal state cleared to allow removal*
- *Buffered data for XOR, EXTENDED COPY, COPY*
- *Access controls data*
- *AccessID enrollment state to pending enrolled*

All the other rows are FCP-2 specific and should remain.

b) Delete Table 6 "Management of mode pages during PRLI and PRLO."

c) Define "I_T nexus loss" to provide linkage to the SCSI architecture and command set changes being made.

Resolution: implemented changes in FCP-2 Rev 8.
Item a) – deleted the rows. Additionally the columns "Failed discovery after LIP" and "Failed discovery after OLS" because they are related to authentication failure following an initialization event. Authentication failures cause a LOGO (already defined in the Clearing effects table).
Item b) – deleted table 6.
Item c) – defined I_T nexus loss and added text specifying events that shall cause an "I_T nexus loss" notification and "Transport Reset" notification.

HP Issue #2

*The Time to wait for a response to Read Exchange Concise (REC) is specified as one Resource Allocation Timeout(Extended Link Service) (R_A_TOV(ELS)). All other extended link services (ELSes) use 2*R_A_TOV.*

*T10 needs to affirm that R_A_TOV is intended or change it to 2*R_A_TOV.*

Resolution:

a. Specified in clause 12.5.2 that the time to wait for a response to an REC is 2xR_A_TOV.
b. Changed figure C.26 and figure C.29 to indicate the time to wait is 2xR_A_TOV.

HP Issue #3

There is the possibility of a data corruption when using FCP-2 sequence error recovery under the following conditions:

- *The Read Exchange Concise Timeout Value (REC_TOV) timer expires*
- *REC is issued and it times out after R_A_TOV (or 2*R_A_TOV, depending on how issue #2 is resolved).*
- *REC is aborted and retried. By this time, the Resource Recovery Timeout Value (RR_TOV) has expired at target and it has discarded state.*
- *The REC retry receives an FCP FC-4 Link Service Reject (FCP_RJT) response due to the target having discarded state.*

The initiator cannot distinguish between a REC FCP_RJT due to FCP_CMD loss and target discard of exchange state. Under these conditions, the initiator will either always have to abort the exchange (rendering FCP-2 no better than FCP in error resilience in FCP_CMD loss scenarios), or attempt command retry and run the risk of data corruption, due to the command already having completed previously.

The root cause of the problem is that FCP-2 allows the targets to discard exchange state information RR_TOV after sending the response (in the case FCP_CONF is not in use), whereas initiators are permitted to continue to attempt FCP-2 SLER on a REC timeout by retrying the REC. The REC timeout value is R_A_TOV which can be greater than RR_TOV.

Resolution:

a) RR_TOV:
   a. If RETRY bit is set to 0: 2 sec.
   b. If RETRY bit is set to 1: ≥ REC_TOV + 2xR_A_TOV$_{ELS}$ + 1 sec.
b) ULP_TOV:
   a. If RETRY bit is set to 0: ≥ Operation-specific timer + E_D_TOV + 1 sec.
   b. If RETRY bit is set to 1: ≥ Operation-specific timer + 2xRR_TOV

HP issue #4

FCP-2 task retry identification and FCP_CONF features are optional to implement per the standard. When FCP_CONF is not in use and task retry identification is not enabled, there is a potential for data corruption under the following conditions:

- *An exchange completes and its originator ID (OX_ID) is re-used for the next exchange, issued within RR_TOV after the previous use of that OX_ID.*
- *In the new exchange, the FCP_CMD is lost and the initiator issues REC.*

Since the target never saw the new exchange, the REC response is sent with information about the previous exchange. The initiator commences sequence recovery based on incorrect REC response data. There is an exposure to the risk of data corruption when this condition occurs.

Resolution: Defer to FCP-3. This issue is discussed in clause 4.6 and is well understood.

HP issue #5
Per Section 9.4.1, the target is allowed to terminate an exchange for which it has detected an error and send a FCP_RSP with an appropriate CHECK CONDITION status and sense data that describes the error.

Since the initiator is not aware of the possible sense key and additional sense code that the target can return in the above scenarios, it cannot commence FCP-2 SLER when the target resorts to behavior as described above.

This causes FCP-2 SLER usage to be non-dependable in scenarios such as "Lost Write Data, not last frame of sequence."

Resolution: the current text addresses the behavior when device errors (i.e., not Sequence errors) are detected at the target. It is appropriate for a target to return CHECK CONDITION for a detected device error. Text in clause 12.3.5 is applicable to target detected Sequence errors. Added text specifying a CHECK CONDITION shall not be sent if data retransmission is enabled and the target detects a Sequence error. Also added a reference to clause 12.3.5.

HP issue #6
Per Section 12.5.2, the initiator shall abort (send ABTS + RRQ) a REC that does not complete within R_A_TOV(ELS) and retry the REC. This error recovery scenarios are rendered useless since the target is allowed to discard exchange state within RR_TOV after sending a FCP_RSP and RR_TOV can be < R_A_TOV(ELS).

The RR_TOV timer value needs to be re-defined for the purpose of FCP-2 SLER, or a new timer value needs to be used in its place which allows targets to discard exchange state.

Resolution: See issue #3 resolution. The dual use of this timer has become a problem (e.g., certainly don't need to wait an extended period of time for authentication to occur following loop initialization). Need to resolve this issue in FCP-3.

HP issue #7 (post letter ballot)
We have found an inconsistency between the SLER error recovery described in Section 12.4.1.7 and the Annex figures C.17 & C.19.

Fig C17/C19 state that on a lost read data, unacknowledged classes, the ACC for the REC indicates that the target does not hold sequence initiative and the exchange is "open".

*This is incorrect, since the REC ACC will indicate that the target does not hold sequence initiative and the exchange is *complete*, when FCP_CONF is not in use. (The target would have sent the FCP_RSP, following the completion of data phase.)*

Fig C17/C19 should be modified to indicate that :
- No FCP_CONF is in use
- the sequence initiative is not held by the target.
- exchange is complete.
- Initiator's "Data Transfer Count" is less than target's "Data Transfer Count" which is reported in the REC ACC.

Resolution: Reworded text in figures C.17 and C.19.