August 4, 2000 T10/00-266 revision 2

To: John Lohmeyer, chairperson, T10
From: Bob Snively
Date: July 20, 2000
Subject: Draft of error recovery clauses, FCP-2 revision 5

This document contains a draft of Annex D to be included in FCP-2, revision 5, based on T10/00-137r5. Revision 1 adds the requested corrections discovered during the FCP-2 working group meeting of July 10, 2000.

Revision 2 contains the complete draft of clause 12 (retries) and Annex D (examples of retries) as of August 4, 2000.

All information is included except the improvements proposed by David Baldwin to improve certain ambiguous recovery cases.

Brocade Communications Systems, Inc.
1901 Guadalupe Parkway n San Jose, CA 95131
T 408.487.8000 n F 408.487.8101
www.brocade.com

12 Class independent link error recovery procedure

12.1 Overview

12.1.1 Exchange level error recovery

FCP-2 provides several mechanisms for devices to identify protocol errors caused by frames and responses that have been corrupted and discarded in accordance with the requirements of FC-FS. See 12.2 for a list of these mechanisms.

To recover from these errors, all FCP-2 compliant initiators shall be capable of invoking the recovery abort function to terminate a failing exchange and to recover the associated resources as described in 12.3. All FCP-2 compliant targets shall be capable of executing the requested recovery abort to finish clearing the exchange and to recover the associated resources. The failed command may then be reissued by higher level programs according to protocols beyond the scope of this standard.

FCP-2 has expanded the error detection capabilities defined by the FCP standard by allowing the optional use of the REC ELS to monitor the progress of active Exchanges. An FCP-2 device may optionally accept or reject error detection inquiries.

12.1.2 Sequence level error recovery

To recover from errors, FCP-2 compliant devices may optionally perform retransmission procedures that allow the commands to be completed without requiring higher level programs to perform command retries. Such recovery is desirable for SCSI logical units that depend critically on command ordering and maintaining records of internal device state. The SCSI initiator and the SCSI target shall agree to perform retransmission using the SRR ELS by setting the retry bit to 1 in PRLI before performing the retransmission of individual IUs. (See 6.2.6.7). An FCP-2 device that has agreed to perform retransmission shall use and accept the REC and SRR ELSs as defined by this standard to perform the retransmission.

Even after agreeing to perform retransmission, the initiator may choose to request a recovery abort and the target shall be able to accept and perform the recovery abort.

While the error detection and error recovery procedures are class independent, acknowledged classes of services may use the acknowledgement mechanism as an additional error detection feature.

This clause defines the error detection and recovery mechanisms for fabrics that guarantee in-order frame delivery. However, if continuously increasing sequence count is used and if support for recovery qualifiers is fully implemented as defined in FC-FS, the same recovery mechanisms can used for fabrics that do not guarantee in-order frame delivery, as shown in the examples in Annex D.

Examples of error detection and error recovery are provided in Annex D and Annex E.

12.2 FCP Error Detection

12.2.1 Overview of FCP Error Detection

The subclauses of 12.2 describe the initial events that indicate an error may have occurred. The error may be recovered at the Exchange level or at the Sequence level.

12.2.2 FCP Error Detection using protocol errors for all classes of service.

The Exchange originator (SCSI Initiator) shall detect the following errors. It may optionally further identify and recover additional errors as described in 12.4.

a) After REC_TOV times out, no reply Sequence (i.e. FCP_XFER_RDY IU, FCP_DATA IU, or FCP_RSP IU) has been received for the FCP_CMND IU;

b) after REC_TOV times out following the sending of FCP_DATA IU(s), no FCP_RSP or FCP_XFER_RDY IU has been received;

c) a Sequence error is detected in a Sequence transmitted from a target to an initiator;

d) a read command completed with the data count smaller than FCP_DL and FCP_RESID_UNDER is set to 0;

e) a read-type command completed with the data count smaller than FCP_DL, FCP_RESID_UNDER is set to 1, and the data count plus FCP_RESID is not equal to FCP_DL; and

f) an ABTS is received.

The Exchange Responder (SCSI target) shall detect that REC_TOV times out and an expected FCP_CONF IU has not been received.

The Exchange Responder shall also initiate error detection and recovery after a Sequence error is detected in a Sequence transmitted from an initiator to a target. (See 12.3.5.)

An Exchange Responder may optionally further identify and recover additional errors as described in 12.4.

12.2.3 Error Detection mechanisms for acknowledged classes of Service

Acknowledged classes of service provide the additional FCP error detection mechanisms described below.

The Exchange originator (SCSI Initiator) shall detect the following errors. It may optionally further identify and recover additional errors as described in 12.4.

a) After E_D_TOV times out, no ACK has been received for the FCP_CMND IU;

b) after E_D_TOV times out, no ACK has been received for FCP_DATA IU(s) (see example in figure D.22);

c) after E_D_TOV times out and no ACK has been received for the FCP_CONF IU; and

d) an ACK with the F_CTL Abort Sequence Condition bits set to Abort Sequence, Perform ABTS is received. (See FC-FS.)

The Exchange Responder (SCSI target) shall detect the following errors. It may optionally further identify and recover additional errors as described in 12.4.

a) after E_D_TOV times out and no ACK has been received for the FCP_XFER_RDY IU (see example in figure D.6);

b) after E_D_TOV times out and no ACK has been received for FCP_DATA IU(s) (see example in figure D.21); and

c) after E_D_TOV times out and no ACK has been received for the FCP_RSP IU.

If an ABTS is transmitted by a Sequence Initiator because it had detected a missing ACK and the BA_ACC response to the ABTS indicates the Sequence was correctly received by the Sequence Recipient, no error detection or recovery is required.

12.3 Exchange level recovery using recovery abort

12.3.1 Recovery abort requirements

The recovery abort is an FC-FS protocol that recovers FCP_Port resources associated with an exchange that is being terminated, either because of a task management request or because of an error.

Recovery abort may be used whether or not the SCSI devices have agreed to Sequence level error recovery.

All FCP-2 initiators shall be capable of invoking the recovery abort protocol to terminate failing commands for later retry. (See 9.1.1.4.) All FCP-2 targets shall be capable of accepting and completing the recovery abort protocol.

12.3.2 SCSI initiator invocation of recovery abort

The SCSI initiator terminating the exchange sends an ABTS sequence the Parameter field set to ABORT EXCHANGE. The ABTS sequence is generated using the OX_ID and RX_ID of the exchange to be aborted. FC-FS allows ABTS to be generated by an FCP_Port regardless of whether or not it has sequence initiative.

Recovery abort may be invoked even if Sequence Initiative is not held. Following the transmission of ABTS, any Device_Data Frames received for this Exchange shall be discarded until the BA_ACC with “Last Sequence of Exchange” bit set to one is received from the SCSI target.

Recovery abort may not take effect immediately. For example, if ABTS is sent following transmission of a Read command, the SCSI initiator may receive some or all of the requested read data before receiving the BA_ACC to the ABTS. The SCSI initiator shall be capable of receiving this data and providing BB_Credit in order for the SCSI target to send the BA_ACC.

After the execution of a task management function that clears tasks, recovery abort shall be invoked for all ambiguous Exchanges not successfully terminated with an FCP_RSP IU status set to COMMAND CLEARED. (See 9.1.1.4).

Following receipt of the BA_ACC in response to an ABTS, and after $R\_A\_TOV_{SEQ\_QUAL}$ has elapsed, the SCSI initiator shall transmit RRQ.

If a BA_ACC, BA_RJT, LOGO, or PRLO is not received from the SCSI target within 2 times $R\_A\_TOV_{ELS}$, second level error recovery as described in 12.5 shall be performed.

12.3.3 SCSI target response to recovery abort

When an ABTS (Abort Exchange) is received at the SCSI target, it shall abort the designated Exchange and return one of the following responses:

a) the SCSI target shall discard the ABTS and return LOGO if the N_ or NL_Port issuing the ABTS is not currently logged in (i.e. no PLOGI);

b) the SCSI target shall return BA_RJT with Last Sequence of Exchange bit set to one if the received ABTS contains an assigned RX_ID and a FQXID that is unknown to the SCSI target; or

c) the SCSI target shall return BA_ACC with Last Sequence of Exchange bit set to 1.

Upon transmission of any of the above responses, the SCSI target may reclaim any resources associated with the designated Exchange after $R\_A\_TOV_{SEQ\_QUAL}$ has elapsed or a Reinstate Recovery Qualifier (RRQ) extended link service request has been received.

If the RX_ID is FFFFh, SCSI targets shall qualify the FQXID of the ABTS based only upon the combined values of D_ID, S_ID, and OX_ID, not RX_ID.

If the Exchange resources were not reclaimed upon responding to the ABTS, they shall be reclaimed at the time the response to the RRQ is sent.

When an RRQ is received at the SCSI target, it shall return one of the following responses:

a) the SCSI target shall discard the RRQ and return LOGO if the N_ or NL_Port issuing the RRQ is not currently logged in (i.e. no PLOGI);

b) the SCSI target shall return LS_RJT with Last Sequence of Exchange bit set to one if the received RRQ contains an RX_ID, other than FFFFh, that is unknown to the SCSI target. The reason code shall be logical error with a reason code explanation set to Invalid OX_ID-RX_ID combination; or

c) the SCSI target shall return ACC with Last Sequence of Exchange bit set to 1.

12.3.4 Additional error recovery by SCSI initiator

This procedure may be used whether or not the SCSI devices have agreed to Sequence level recovery.

If ULP_TOV times out and the Exchange is not complete, the application client shall clear the exchange resources using the ABORT TASK task management request or the initiator shall clear the exchange resources using the recovery abort protocol. (See 9.1.2.)

12.3.5 Additional error recovery by SCSI target

This procedure may be used whether or not the SCSI devices have agreed to Sequence level recovery.

If a SCSI target detects a Sequence error, it shall discard the Sequence(s) based on the Exchange error policy specified by the F_CTL Abort Sequence Condition bits in the first frame of the Exchange. (See FC-FS.)

For acknowledged classes of service, if a SCSI target detects a Sequence error, it may abort the sequence by sending an ABTS with the Parameter field to ABORT SEQUENCE. If a Recovery Qualifier range is returned in the BA_ACC for the ABTS the target shall send a RRQ ELS after $R\_A\_TOV_{SEQ\_QUAL}$ times out after receipt of the BA_ACC.

For unacknowledged classes of service, the target shall not attempt recovery for Sequence errors. The target shall depend on initiator time outs for recovery.

SCSI targets shall implement RR_TOV as described in 11.3 to facilitate recovery of resources allocated to a SCSI initiator that is no longer responding. The SCSI target may send a LOGO to the SCSI initiator and terminate all open Exchanges for that SCSI initiator upon detection of the following:

a) The SCSI initiator has failed to perform SCSI target Exchange authentication within RR_TOV (see FC-PLDA); or

b) RR_TOV times out without the SCSI initiator transmitting any expected Sequence for any open Exchange at this SCSI target (e.g., FCP write data in response to an FCP_XFER_RDY IU).

12.4 Sequence level error detection and recovery

12.4.1 Using information from REC to perform Sequence level recovery

12.4.1.1 Polling Exchange state with REC

REC is periodically transmitted by the initiator to poll each outstanding Exchange to determine if a SCSI task is progressing properly and if any Sequences have been received incorrectly. Timing of polling with the REC ELS is controlled by REC_TOV. REC_TOV is normally selected to be long enough that processing the transfers of initiative in the exchange and completing the exchange will occur before REC_TOV times out. If REC_TOV times out, then an REC ELS is performed. The information returned in the REC ACC payload is compared with

the expected state information known by the initiator and target. If the information is inconsistent, indicating that a link error occurred, optional error recovery actions may be performed to complete the Exchange. Subclauses 12.4.1.2 through 12.4.1.8 define optional error detection and recovery procedures for acknowledged and unacknowledged classes of service.

12.4.1.2 Detection of errors while polling with REC

If an Exchange Originator receiving an acknowledged service Sequence detects a Sequence error, it shall send an ACK Frame with the F_CTL Abort Sequence Condition bits set to Abort Sequence, Perform ABTS before issuing the REC. The REC for the Exchange containing the FCP_CMND IU shall be issued in a new Exchange.

If the response to the new Exchange issuing the REC is an LS_RJT with a reason code of command not supported, the initiator shall assume the target is a device not supporting error detection using REC. The device shall perform recovery using recovery abort as documented in 12.3.

If an ACC, LS_RJT, LOGO, or PRLO is not received from the SCSI target within 2 times $R\_A\_TOV_{ELS}$, second level error recovery as described in 12.5 shall be performed.

12.4.1.3 FCP_CMND IU Recovery using information from REC

This procedure may be used whether or not the SCSI devices have agreed to Sequence level recovery.

If the FCP_CMND IU was not received by the target (i.e., the initiator receives an LS_RJT for the REC with the reason code of logical error and reason code explanation set to Invalid OX_ID-RX_ID combination), retransmit the FCP_CMND IU using a new Exchange. If the precise delivery function is enabled, the CRN value shall remain the same in the retransmitted FCP_CMND IU.

If the ACC for the REC indicates that the FCP_CMND IU was received by the target and that no reply Sequence has been sent (i.e. by indicating that the initiator does not hold Sequence Initiative, and that the Exchange is not complete), the command is in process and no recovery is needed at this time. At a minimum interval of REC_TOV, the REC shall be retransmitted to more quickly determine if a reply Sequence has been lost.

For examples of such recoveries, see figure D.1 and figure D.2

12.4.1.4 FCP_XFER_RDY IU Recovery

This procedure shall be used only by SCSI 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 (i.e. by indicating that the initiator holds Sequence Initiative, that all bytes were not transferred, and that the Exchange is not complete), but not received by the initiator, issue an SRR in a new Exchange to request retransmission of the FCP_XFER_RDY IU. The target 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 D.5 and figure D.6.

12.4.1.5 FCP_RSP IU Recovery

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

An error in transmitting an FCP_RSP IU is detected if:

a) the ACC for the REC ELS indicates that an FCP_RSP IU was sent by the target and no FCP_CONF IU was requested (i.e. E_STAT indicates that the Exchange is complete), but the initiator has not yet received the FCP_RSP IU; or

b) the ACC for the REC ELS indicates that an FCP_RSP IU Sequence was sent by the target and an FCP_CONF IU was requested (i.e. E_STAT indicates that the Exchange is not complete, that the initiator has initiative, and that, if the data transfer was from the initiator to the target, the data transfer indicates that all of the bytes expected to be transferred by the command have been transferred.)

When an error in transmitting an FCP_RSP IU is detected, the initiator shall issue an SRR FC-4 Link Service frame in a new Exchange to request retransmission of the FCP_RSP IU. The target shall first transmit the ACC for the SRR, then shall retransmit the FCP_RSP IU in a new Sequence.

A command that was terminated prior to transferring data by a CHECK CONDITION requesting the FCP_CONF IU may have the same REC values as a command for which an FCP_XFER_RDY IU was not received by the initiator. For a write command with a non-zero FCP_DL, the parameters for the SRR shall indicate that an FCP_XFER_RDY IU is expected from the target. The target is aware of the actual present state of the transfer and response and shall either retry the FCP_XFER_RDY IU or, if the actual data transfer length for the command was zero, retry the FCP_RSP.

For non-tagged command queuing operations, the target shall retain the Exchange information until

a) the next FCP_CMND IU has been received for that LUN from the same SCSI initiator;

b) a FCP_CONF IU is received for the Exchange; or

c) after RR_TOV times out.

For tagged command queuing operations, the target shall retain Exchange information until

a) a FCP_CONF IU is received for the Exchange;

b) no REC has been received for the Exchange after a time of 3 times REC_TOV; or

c) status retention resources are exhausted.

The Exchange information retained shall include data transfer information, data descriptors, and FCP_RSP IU information.

Examples of FCP_RSP IU recoveries are provided in figure D.8 through figure D.12.

12.4.1.6 FCP_DATA IU Recovery - Write

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

If the ACC for an REC indicates that an FCP_DATA IU was sent by the initiator, but not received by the target (i.e. the data received count in the REC response is smaller than what the initiator sent, and the target indicates it does not hold Sequence Initiative) then the initiator shall send a SRR FC-4 Link Service frame in a new Exchange to request retransmission of an FCP_XFER_RDY IU to request the missing data. The target discards the Sequence in error, but does not initiate any recovery action for Class 3. (See 12.3.5.) After first transmitting the ACC for the SRR, the target transmits an FCP_XFER_RDY IU, in a new Sequence, with the Relative Offset parameter specified by the SRR and the initiator responds with the requested data.

The Sequence count for a retransmitted FCP_DATA IU shall start at zero, even if continuously increasing Sequence count is being used.

FCP_DATA shall be retransmitted in a new Sequence. For acknowledged classes, the SEQ_CNT shall be one greater than that used to transmit the last Sequence, usually the ABTS. For unacknowledged classes, the SEQ_CNT may start at zero, even if continuously increasing sequence count is being used.

Examples of data recovery during write operations are provided in figure D.13 through figure D.16.

12.4.1.7 FCP_DATA IU Recovery - Read

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

If the ACC for the REC indicates that data was sent by the target but not successfully received by the initiator (i.e. by indicating a data sent count greater than the initiator has successfully received), then the initiator shall send a SRR FC-4 Link Service frame in a new Exchange to request retransmission of the FCP_DATA IU that was not successfully received. The initiator shall set the Relative Offset field in the SRR to that of the next data requested. If the initiator is unable to determine the Relative Offset of the next data requested, the initiator shall set the Relative Offset to zero. The target shall first transmit the ACC for the SRR, then shall retransmit the requested data specified by the SRR in a new Sequence, and then complete the Exchange in the normal manner, including transmitting or retransmitting the FCP_RSP IU. If the target responds to the SRR with an LS_RJT and a reason code of command not supported, and an FCP_RSP IU has not already been sent, the target shall send an FCP_RSP IU with CHECK CONDITION status and sense information containing a sense key of HARDWARE ERROR and an ASC/ASCQ of INITIATOR DETECTED ERROR MESSAGE RECEIVED.

FCP_DATA shall be retransmitted in a new Sequence. For acknowledged classes, the SEQ_CNT shall be one greater than that used to transmit the last Sequence, usually the ABTS. For unacknowledged classes, the SEQ_CNT may start at zero, even if continuously increasing sequence count is being used.

It is the responsibility of the initiator to determine the appropriate action required (e.g. retry, allow ULP time out, or return status to ULP) based on the information determined by REC and other internal states. The SCSI target does not initiate error recovery for Class 3. (See 12.3.5.)

Examples of data recovery during read operations are provided in figure D.17 through figure D.20.

12.4.1.8 FCP_CONF IU Recovery

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

This recovery procedure is used by target devices using all service classes.

Target devices that implement confirmed completion shall set the RX_ID to a unique value other than FFFFh for each Exchange to enable unambiguous recovery.

If the SCSI initiator has sent the FCP_CONF IU, the reply to the REC from the target shall be a LS_RJT with the reason code of logical error and reason code explanation set to Invalid OX_ID-RX_ID combination. The SCSI target shall assume that the FCP_CONF IU was sent and release the Exchange.

If the SCSI initiator has received the FCP_RSP IU with the FCP_CONF_REQ bit set to 1 and has not sent the FCP_CONF IU before the REC is received, the REC reply shall be an ACC indicating the Exchange is still open. In this case the target shall wait REC_TOV and, if the FCP_CONF IU has not been received, send another REC. The target shall repeat this process until the FCP_CONF IU is received, a new FCP_CMND IU is received with the same OX_ID as the Exchange waiting for the FCP_CONF IU, or until the Exchange is aborted.

If another FCP_CMND IU is received by the target with the same OX_ID as an Exchange waiting for a FCP_CONF IU and with the RX_ID unassigned, the SCSI target shall assume that the FCP_CONF IU was sent and release the Exchange.

12.4.2 Additional error recovery requirements

12.4.2.1 Error indicated in ACK

If an ACK is received with the F_CTL Abort Sequence Condition bits set to Abort Sequence, Perform ABTS, the Sequence Initiator shall send an ABTS for the Sequence. After R_A_TOV times out, an RRQ shall be sent by the Sequence Initiator.

12.4.2.2 Missing ACK

FC-FS requires that an ABTS(Sequence) be transmitted by a Sequence Initiator detecting a missing ACK. If no ACK has been received within E_D_TOV, the target shall abort the sequence by sending an ABTS request with the Parameter field set to ABORT SEQUENCE. If a Recovery Qualifier range is returned in the BA_ACC for the ABTS the target shall send a RRQ ELS at least $R\_A\_TOV_{SEQ\_QUAL}$ after receipt of the BA_ACC. Adjustment of subsequent sequence counts may be required as specified by FC-FS.

12.4.2.3 Distinguishing exchange to be aborted

When OX_ID values are reused within R_A_TOV and RX_ID values are not used, it is possible for a missing ACK to a FCP_RSP IU to allow the target to attempt to abort a more recent Exchange using the same OX_ID. To prevent that, a target using acknowledged service behavior and performing error recovery shall:

a) set RX_ID to a value other than FFFFh to distinguish outstanding Exchanges as described in FC-FS and FC-PH; or

b) always request FCP_CONF IU.

If a Sequence error is detected for an FCP_DATA IU performing a Data Out action (IU T6 or T7), the target shall send an ACK Frame with the Abort Sequence Condition bits set to “Abort Perform ABTS”.

Examples of data recovery for acknowledged services are shown in Annex D.

Recovery abort shall be invoked for Exchanges that were not successfully recovered by the specified error recovery procedures.

12.5 Second-level error recovery

12.5.1 ABTS

If a response to an ABTS is not received within 2 times $R\_A\_TOV_{ELS}$, the SCSI initiator may send the ABTS again, attempt other retry operations allowed by FC-FS, or explicitly logout the SCSI target. If those retry operations attempted are unsuccessful, the SCSI initiator shall explicitly logout (i.e. use FC-FS Logout, LOGO) the SCSI target. All outstanding Exchanges with that SCSI target are terminated at the SCSI initiator.

12.5.2 REC

If a response to an REC is not received within 2 times $R\_A\_TOV_{ELS}$, the SCSI initiator shall:

1)send an ABTS(Exchange) for the REC followed by an RRQ if a BA_ACC is received for the ABTS; and

2)send another REC in a new Exchange.

If the response to the second REC is not received within 2 times $R\_A\_TOV_{ELS}$, the SCSI initiator may:

1)send an ABTS(Exchange) for the REC followed by an RRQ if a BA_ACC is received for the ABTS;

2)perform a recovery abort for all outstanding exchanges for that target; and

3)perform an implicit logout for that target.

Other retry mechanisms after the second REC fails shall comply with FC-FS, but are otherwise vendor specific.

See figure D.26, figure D.27, figure D.28, and figure D.29.

12.5.3 SRR

If a response to an SRR is not received within 2 times R_A_TOVELS, the SCSI initiator shall:

1)send an ABTS(Exchange) for the SRR followed by an RRQ if a BA_ACC is received for the ABTS;

2)may perform a recovery abort for the original exchange;

3)may perform a recovery abort for all other exchanges to the same target; and

4)may perform an implicit logout for that target

Other retry mechanisms after the SRR fails shall comply with FC-FS, but are otherwise vendor specific.

See figure D.30 through figure D.33.

12.6 Responses to FCP-level Frames before PLOGI or PRLI

If a SCSI target receives an FCP_CMND IU from an FCP_Port that is not successfully logged on to the target using either an implicit or explicit Login (i.e. PLOGI), it shall discard the FCP_CMND IU and, in a new Exchange, send LOGO to that FCP_Port. No Exchange is created in the SCSI target for the discarded request, and the Originator of the discarded request terminates the Exchange associated with the discarded request and any other open Exchanges for the SCSI target sending the LOGO.

If a SCSI target receives an FCP_CMND IU from an FCP_Port that has not successfully completed either implicit or explicit Process Login (i.e. PRLI) with the target, it shall discard the FCP_CMND IU and send PRLO to the SCSI initiator. No Exchange is created in the recipient FCP_Port for the discarded request, and the Originator of the discarded request terminates the Exchange associated with the discarded request.

If a SCSI device receives a frame of category 0001b or 0011b (solicited data or solicited control) and the SCSI device is has not performed a successful implicit or explicit PLOGI and PRLI with the source of the frame, the SCSI device shall discard and ignore the content of the frame. If the PLOGI is not completed, the SCSI device may transmit a LOGO extended link service request to the source of the unexpected frame. If the PLOGI is completed, but the PRLI is not completed, the SCSI device may transmit a PRLO extended link service request to the source of the unexpected frame.

Annex D Error detection and recovery action examples

(Informative) [Draft, based on T10/00-137r5]

D.1 Introduction

This annex diagrams various error detection and recovery procedures for SCSI devices conforming to this profile.

Table D.1 - Diagram Drawing Conventions

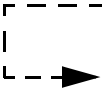
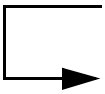
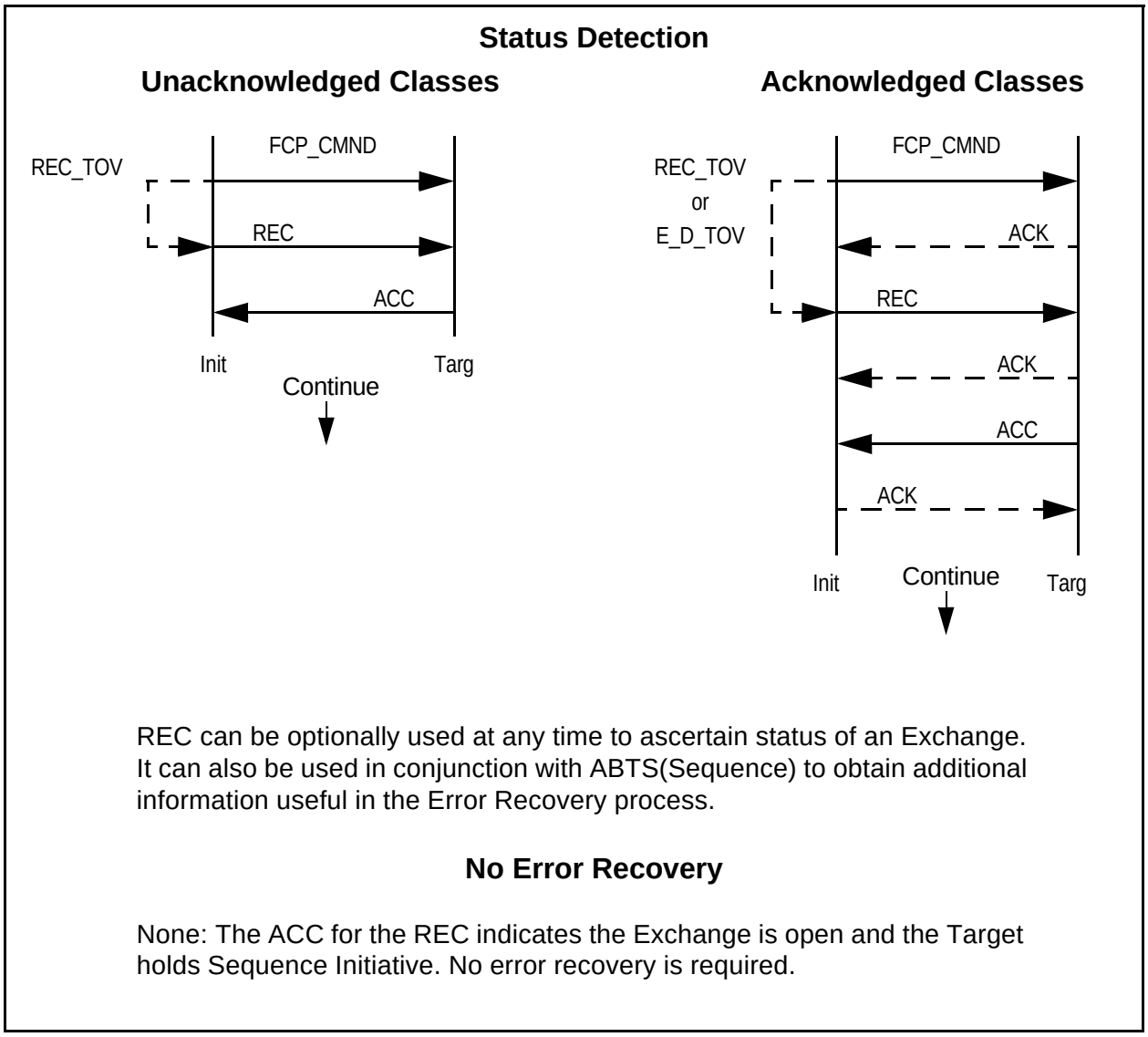
Drawing Convention	Meaning
	Acknowledged or Unacknowledged Frame
	Acknowledgement Frame
	Time-out value exceeded, caused transmission of IU or ELS
	IU or ELS received is processed to transmit IU or ELS
X	Frame lost or dropped
CI Continue 	Error detection complete. Operation continues with specified Error Recovery if continuously increasing sequence count prerequisites are met.
Continue 	Error detection complete. Operation continues with specified Error Recovery if continuously increasing sequence count prerequisites are not met.

Figure D.1 - Lengthy FCP_CMND or Lost ACK



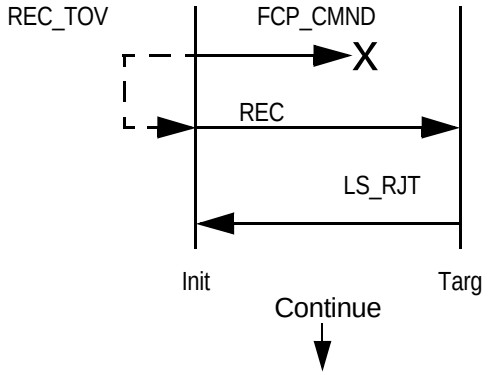
REC can be optionally used at any time to ascertain status of an Exchange. It can also be used in conjunction with ABTS(Sequence) to obtain additional information useful in the Error Recovery process.

No Error Recovery

None: The ACC for the REC indicates the Exchange is open and the Target holds Sequence Initiative. No error recovery is required.

Figure D.2 - FCP_CMND Lost, Unacknowledged Classes

Error Detection



Error Recovery

The LS_RJT (Logical Error, Invalid OX_ID/RX_ID combination) for the REC indicates the Exchange is unknown. The FCP_CMND is retransmitted in a new Exchange using the same CRN and in the same class of service.

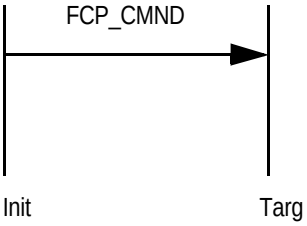
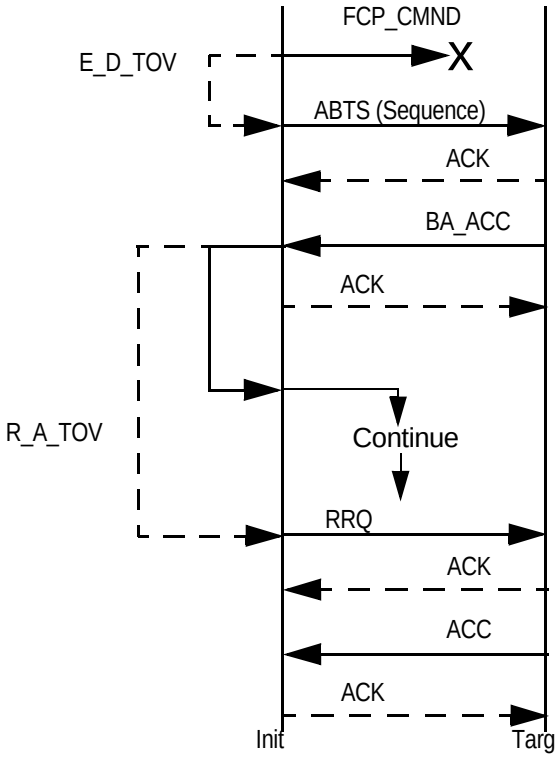


Figure D.3 - FCP_CMND Lost, Acknowledged Classes

Error Detection



The BA_ACC indicates that the FCP_CMND was never received. (BA_ACC payload: SEQ_ID Validity = invalid, Low SEQ_CNT = 0, High SEQ_CNT = SEQ_CNT of ABTS frame). Both the initiator and target establish Recovery Qualifiers. The value of R_A_TOV for in-order delivery is zero. The use of REC to determine status for error recovery is optional.
The FCP_CMND is retransmitted in a new Exchange using the same CRN and in the same class of service.

Error Recovery

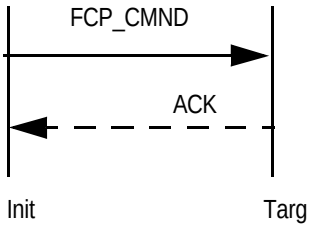
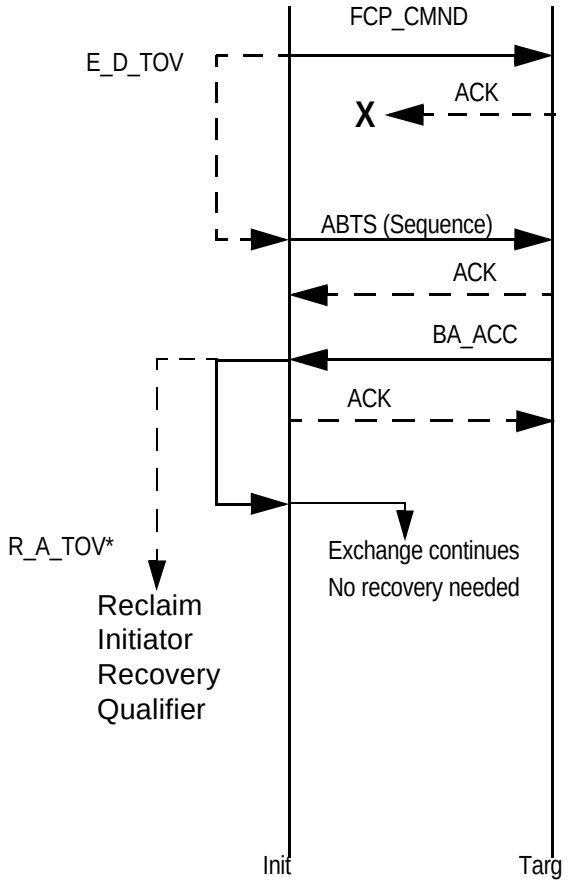


Figure D.4 - FCP_CMND Acknowledgement Lost, Acknowledged Classes

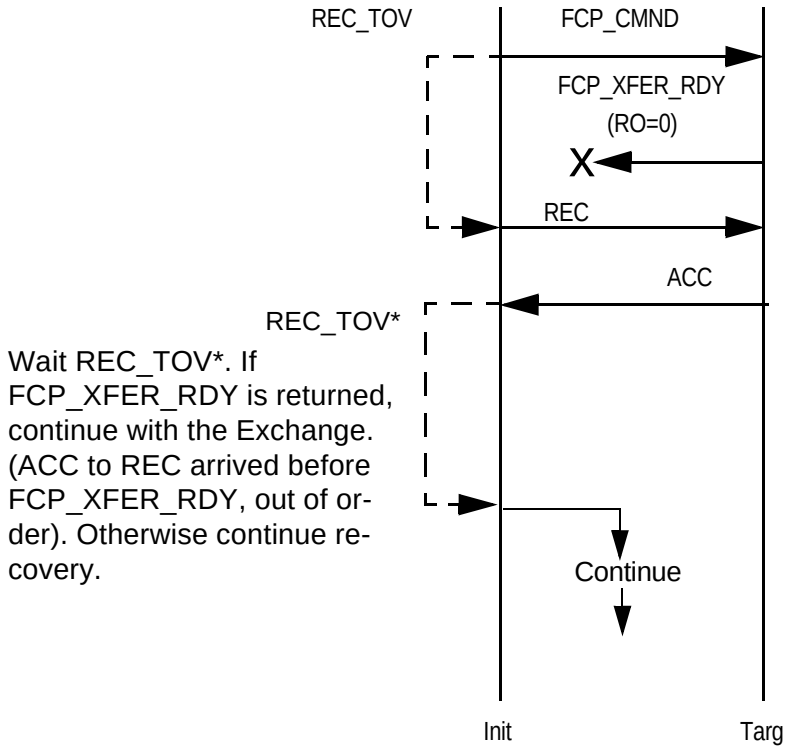
Error Detection



No error recovery required. The BA_ACC Payload indicates that FCP_CMND was received. The BA_ACC payload is SEQ_ID Valid, SEQ_ID value of FCP_CMND, low SEQ_CNT = high SEQ_CNT = SEQ_CNT of FCP_CMND. Note that the issuance of RRQ is not necessary in this case, since the target will not have established a Recovery Qualifier. However, the initiator cannot reclaim the resources associated with its Recovery Qualifier until R_A_TOV time-out expires. The value for R_A_TOV is 0 for in-order delivery. The use of REC to determine status for error recovery is optional.

Figure D.5 - FCP_XFER_RDY Lost, Unacknowledged Classes

Error Detection



Error Recovery

The ACC for the REC indicates the Initiator holds Sequence Initiative and the Exchange is open. The Initiator sends an SRR requesting the FCP_XFER_RDY be resent. The Target resends the FCP_XFER_RDY using the same Relative Offset. A new SEQ_ID must be used for retransmission of the FCP_XFER_RDY. For in-order delivery, the value of REC_TOV* is 0.

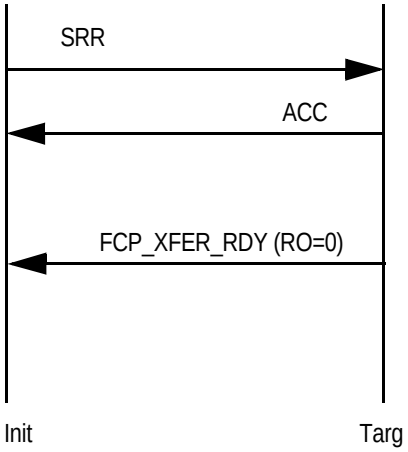
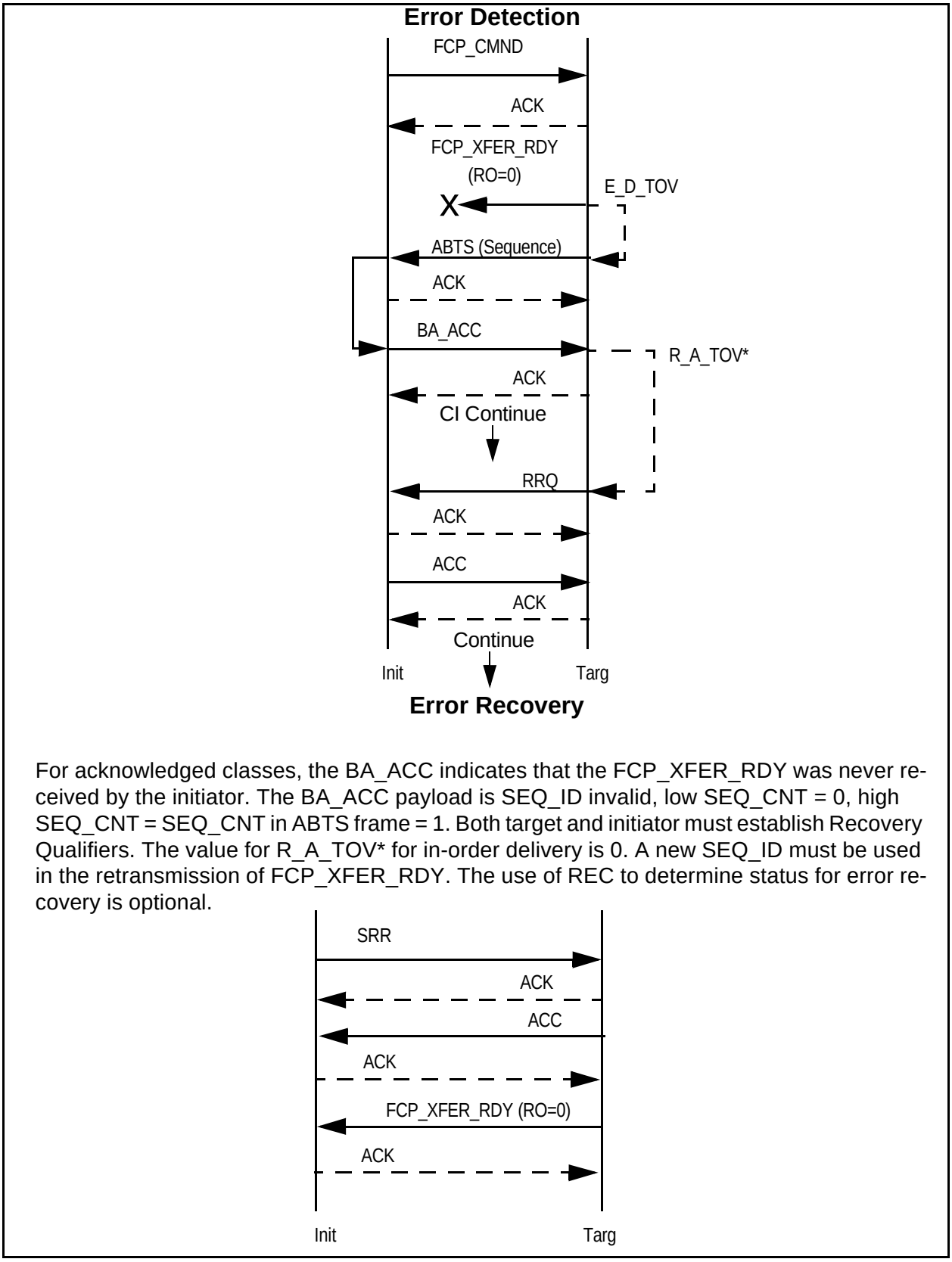
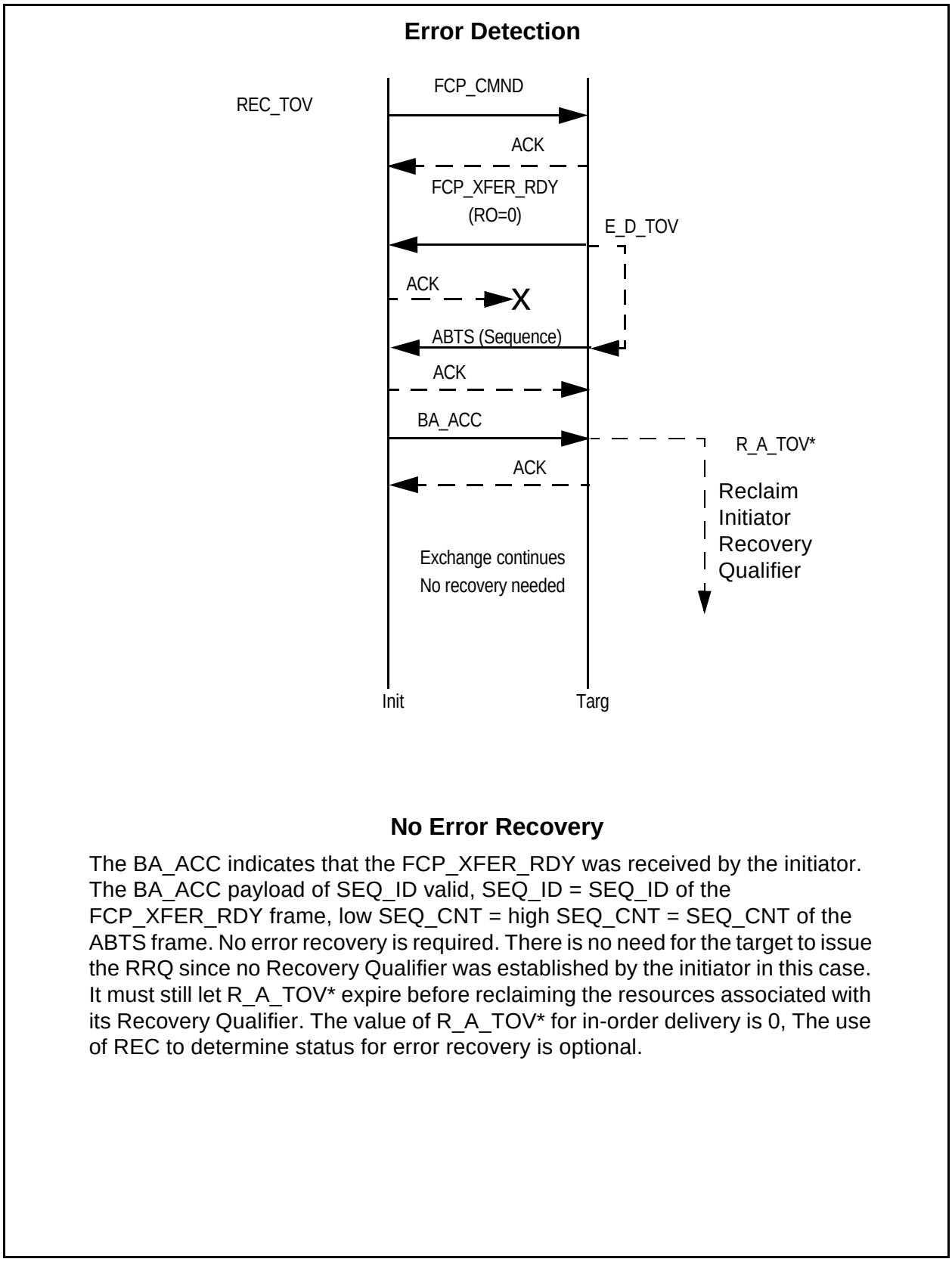


Figure D.6 - FCP_XFER_RDY Lost, Acknowledged Classes



For acknowledged classes, the BA_ACC indicates that the FCP_XFER_RDY was never received by the initiator. The BA_ACC payload is SEQ_ID invalid, low SEQ_CNT = 0, high SEQ_CNT = SEQ_CNT in ABTS frame = 1. Both target and initiator must establish Recovery Qualifiers. The value for R_A_TOV* for in-order delivery is 0. A new SEQ_ID must be used in the retransmission of FCP_XFER_RDY. The use of REC to determine status for error recovery is optional.

Figure D.7 - FCP_XFER_RDY Received, ACK Lost, Acknowledged Classes

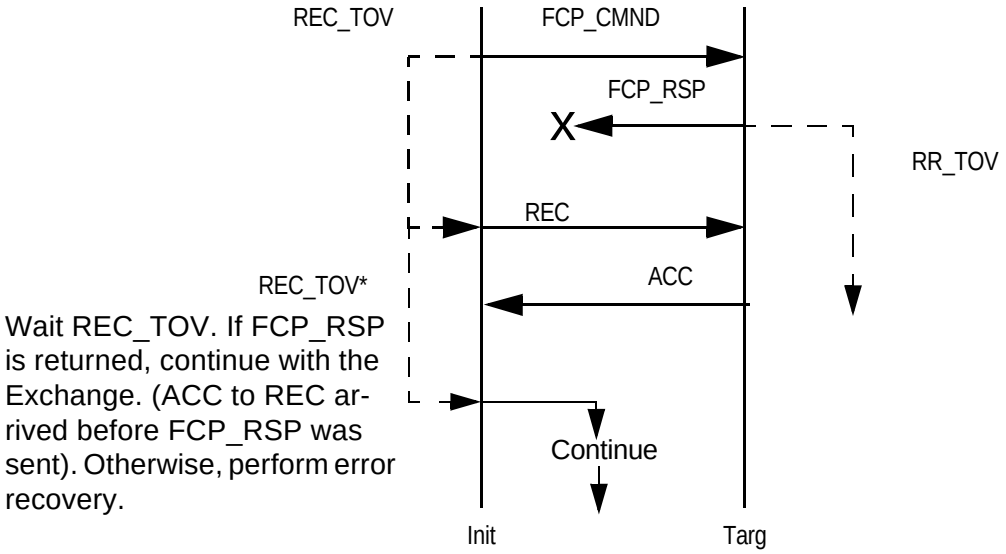


No Error Recovery

The BA_ACC indicates that the FCP_XFER_RDY was received by the initiator. The BA_ACC payload of SEQ_ID valid, SEQ_ID = SEQ_ID of the FCP_XFER_RDY frame, low SEQ_CNT = high SEQ_CNT = SEQ_CNT of the ABTS frame. No error recovery is required. There is no need for the target to issue the RRQ since no Recovery Qualifier was established by the initiator in this case. It must still let R_A_TOV* expire before reclaiming the resources associated with its Recovery Qualifier. The value of R_A_TOV* for in-order delivery is 0, The use of REC to determine status for error recovery is optional.

Figure D.8 - FCP_RSP Lost, FCP_CONF not requested, Unacknowledged Classes

Error Detection



Error Recovery

The ACC for the REC indicates the Initiator holds Sequence Initiative and the Exchange is complete. The Initiator sends an SRR requesting the FCP_RSP be resent. The Target retransmits the FCP_RSP. For in-order delivery, the value of REC_TOV* is 0.

For sequence level recovery, the target must keep the context of this Exchange as defined in 12.XXX.

For exchange level recovery, the context of the Exchange can be purged by the target after FCP_RSP is transmitted.

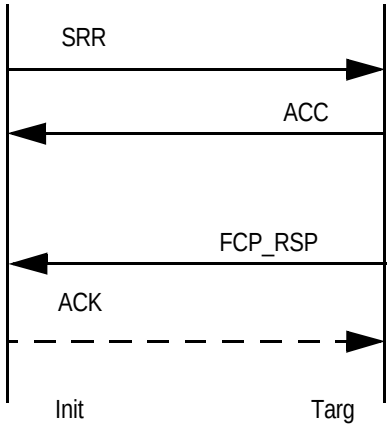
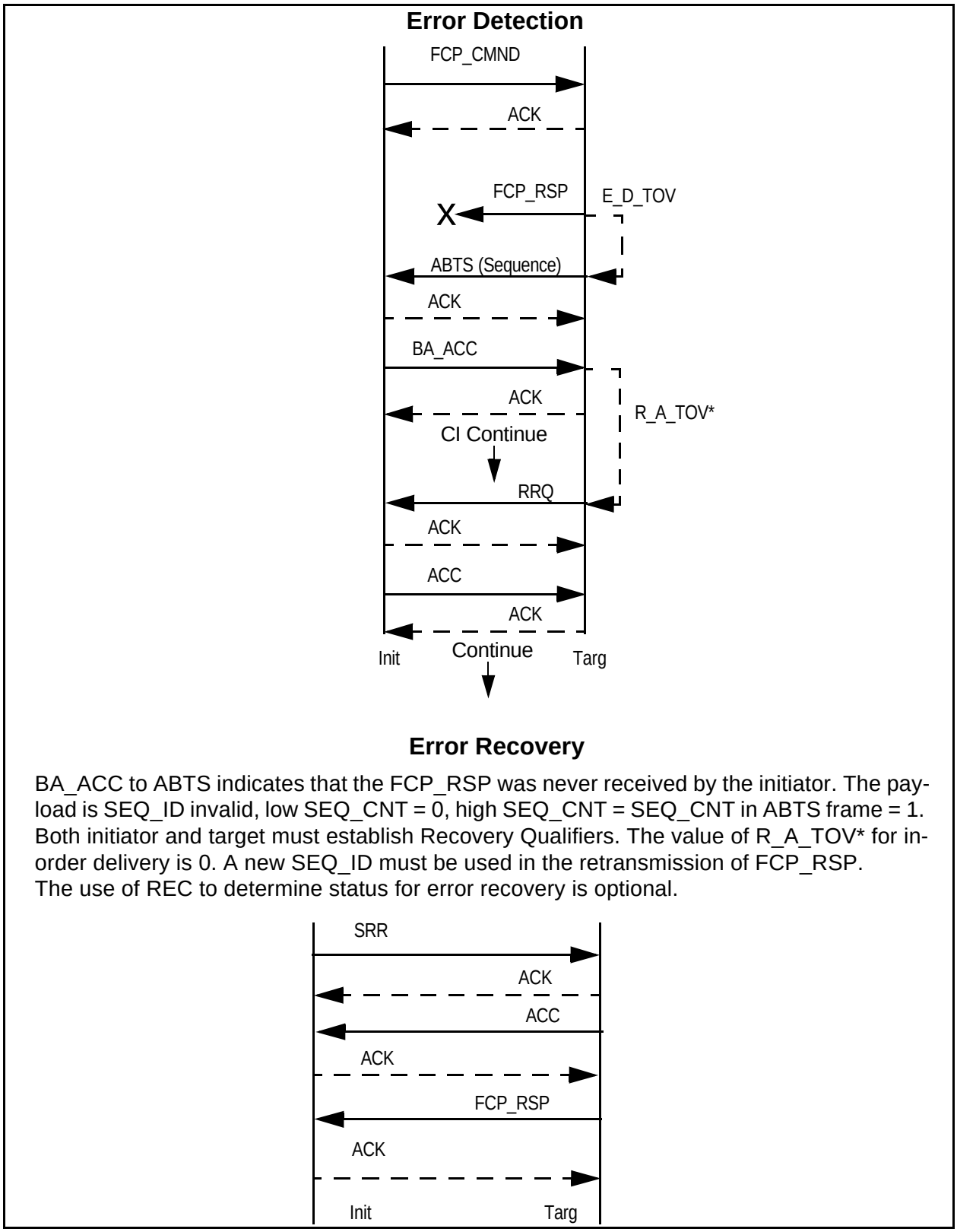
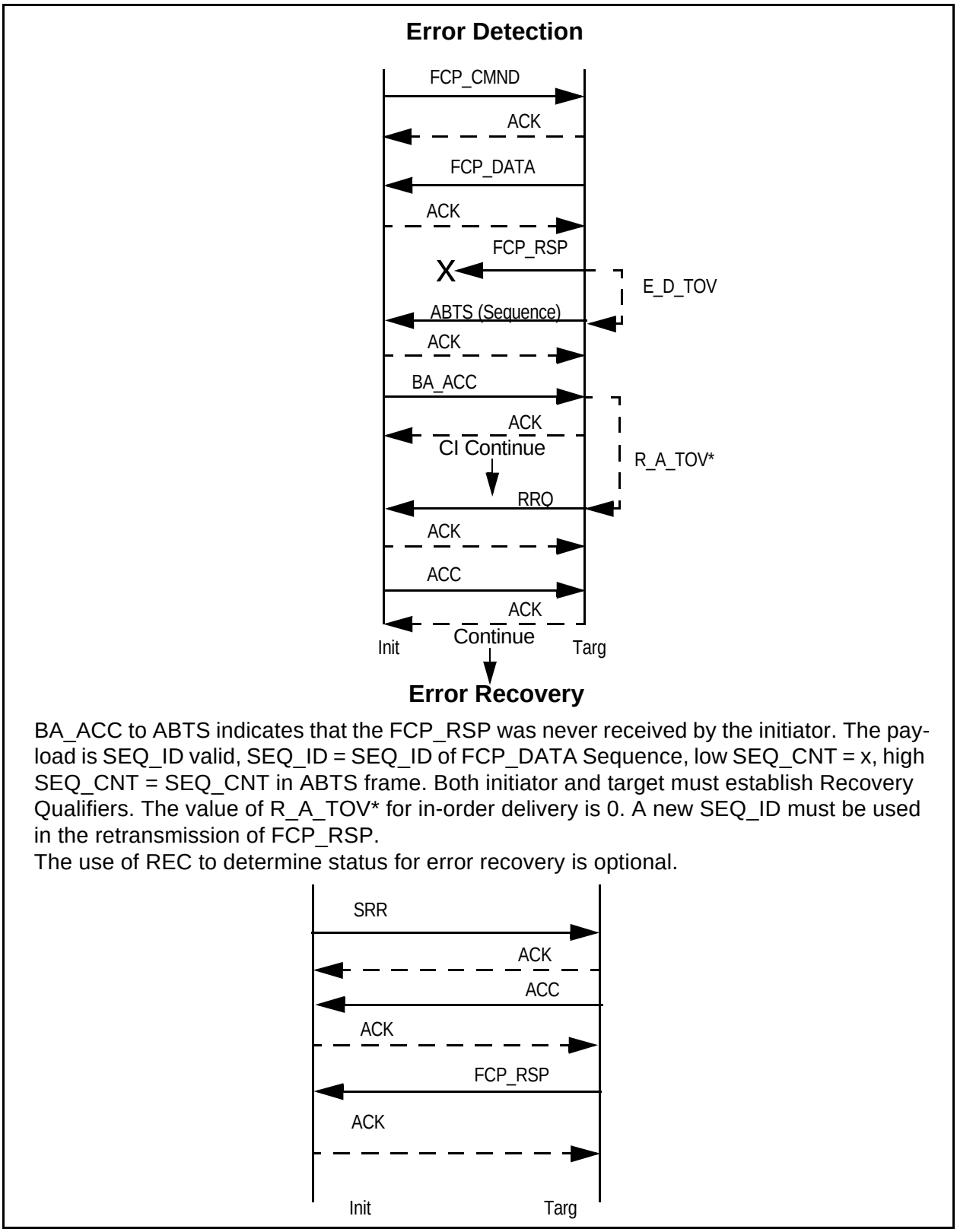


Figure D.9 - FCP_RSP Lost, FCP_CONF not requested, Acknowledged Classes



BA_ACC to ABTS indicates that the FCP_RSP was never received by the initiator. The payload is SEQ_ID invalid, low SEQ_CNT = 0, high SEQ_CNT = SEQ_CNT in ABTS frame = 1. Both initiator and target must establish Recovery Qualifiers. The value of R_A_TOV* for in-order delivery is 0. A new SEQ_ID must be used in the retransmission of FCP_RSP.
The use of REC to determine status for error recovery is optional.

Figure D.10 - FCP_RSP Lost after Read Command, FCP_CONF not requested, Acknowledged Classes



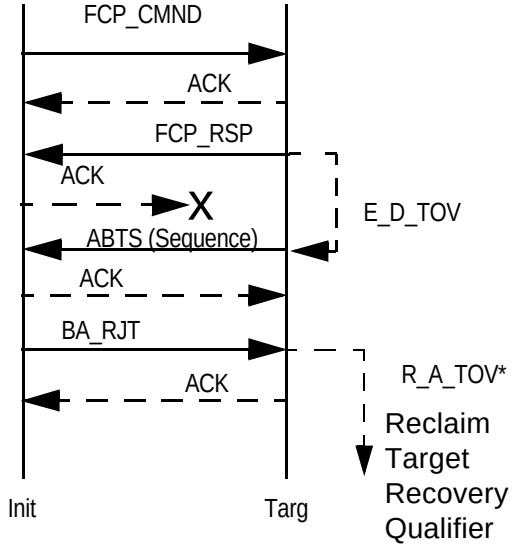
BA_ACC to ABTS indicates that the FCP_RSP was never received by the initiator. The payload is SEQ_ID valid, SEQ_ID = SEQ_ID of FCP_DATA Sequence, low SEQ_CNT = x, high SEQ_CNT = SEQ_CNT in ABTS frame. Both initiator and target must establish Recovery Qualifiers. The value of R_A_TOV* for in-order delivery is 0. A new SEQ_ID must be used in the retransmission of FCP_RSP.

The use of REC to determine status for error recovery is optional.

Figure D.11 - FCP_RSP Received, ACK Lost, Acknowledged Classes, Example 1

Error Detection

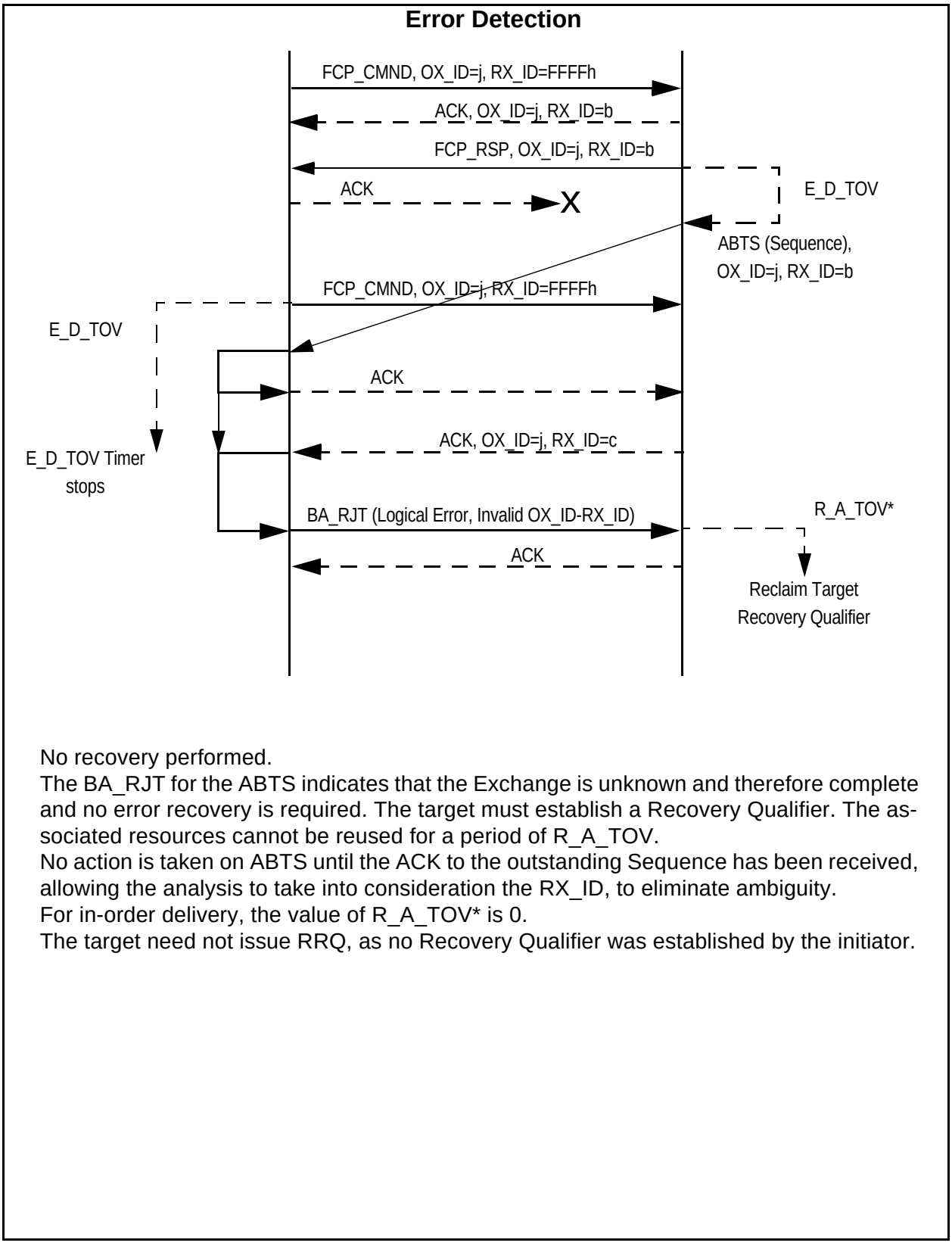
Acknowledged Classes



No Error Recovery

No error recovery required.
The BA_RJT for the ABTS indicates with reason code of Logical Error and the reason explanation code of Invalid OX_ID-RX_ID Combination that the Exchange is unknown and therefore complete. No error recovery is required. The target must establish a Recovery Qualifier. The associated resources cannot be reused for a period of R_A_TOV. For in-order delivery, the value of R_A_TOV* is 0. The target does not need to issue RRQ as no Recovery Qualifier was established by the initiator.

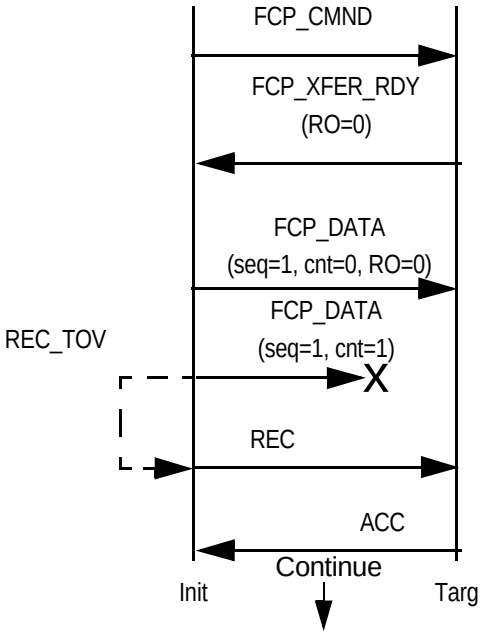
Figure D.12 - FCP_RSP Received, ACK Lost, Acknowledged Classes, Example 2



No recovery performed.
The BA_RJT for the ABTS indicates that the Exchange is unknown and therefore complete and no error recovery is required. The target must establish a Recovery Qualifier. The associated resources cannot be reused for a period of R_A_TOV.
No action is taken on ABTS until the ACK to the outstanding Sequence has been received, allowing the analysis to take into consideration the RX_ID, to eliminate ambiguity.
For in-order delivery, the value of R_A_TOV* is 0.
The target need not issue RRQ, as no Recovery Qualifier was established by the initiator.

Figure D.13 - Lost Write Data, Last Frame of Sequence, Unacknowledged Classes

Error Detection



Error Recovery

The ACC for the REC indicates the Target does not hold Sequence Initiative and the Exchange is open. The Initiator sends an SRR requesting a FCP_XFER_RDY be resent using the specified Relative Offset. The Target retransmits the FCP_XFER_RDY using the specified Relative Offset. New SEQ_IDs are used for FCP_XFER_RDY and FCP_DATA retransmission.

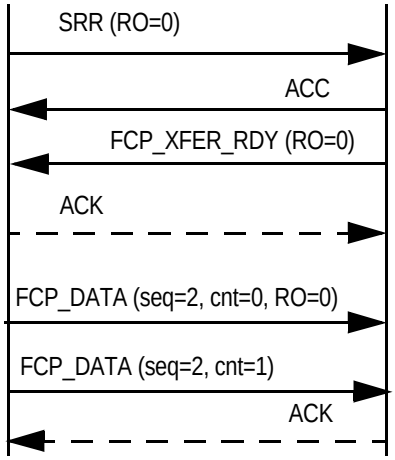
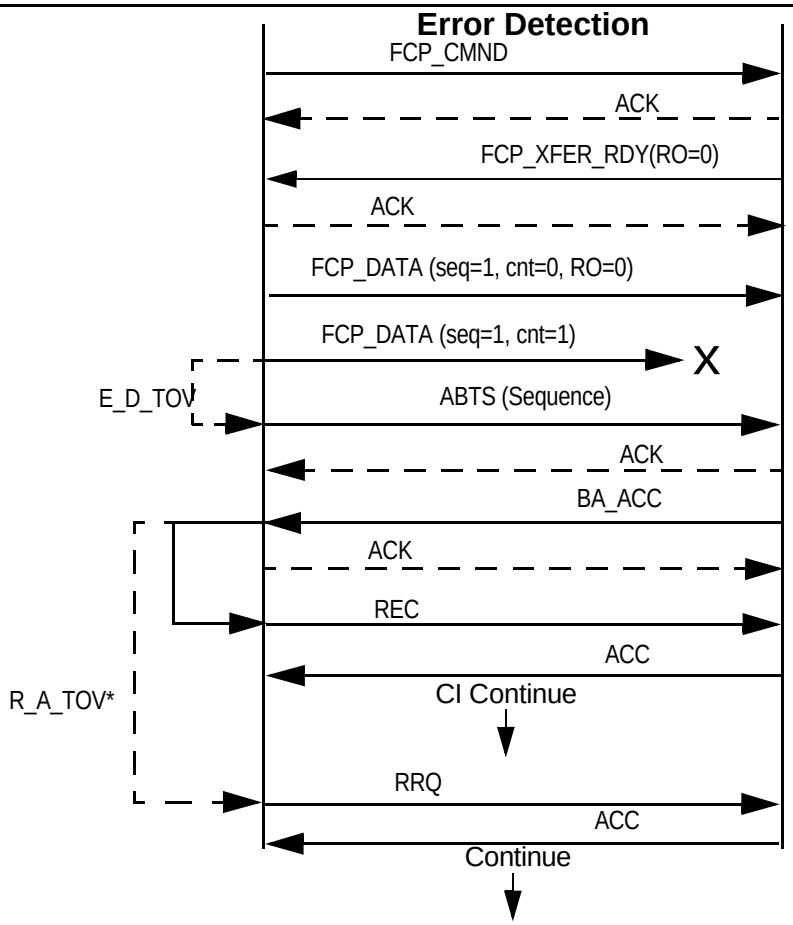


Figure D.14 - Lost Write Data, Last Frame of Sequence, Acknowledged Classes



Error Recovery

Using the information contained in the REC, the recovery may be performed as in figure D.13. The BA_ACC may also be used. The payload is SEQ_ID = valid, SEQ_ID = SEQ_ID of FCP_CMND, low SEQ_CNT of 0, high SEQ_CNT of 2.

New SEQ_IDs shall be used for retransmitting FCP_XFER_RDY and FCP_DATA. ACKs for REC/ACC and RRQ/ACC are not shown. The value of R_A_TOV* for in-order delivery is 0.

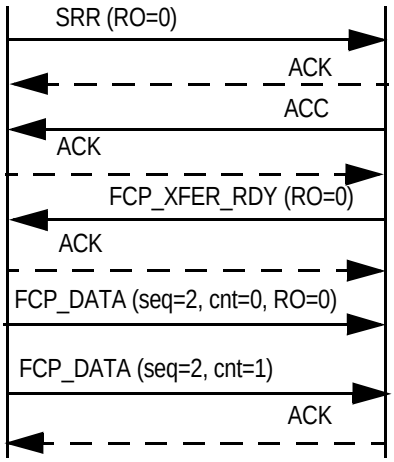
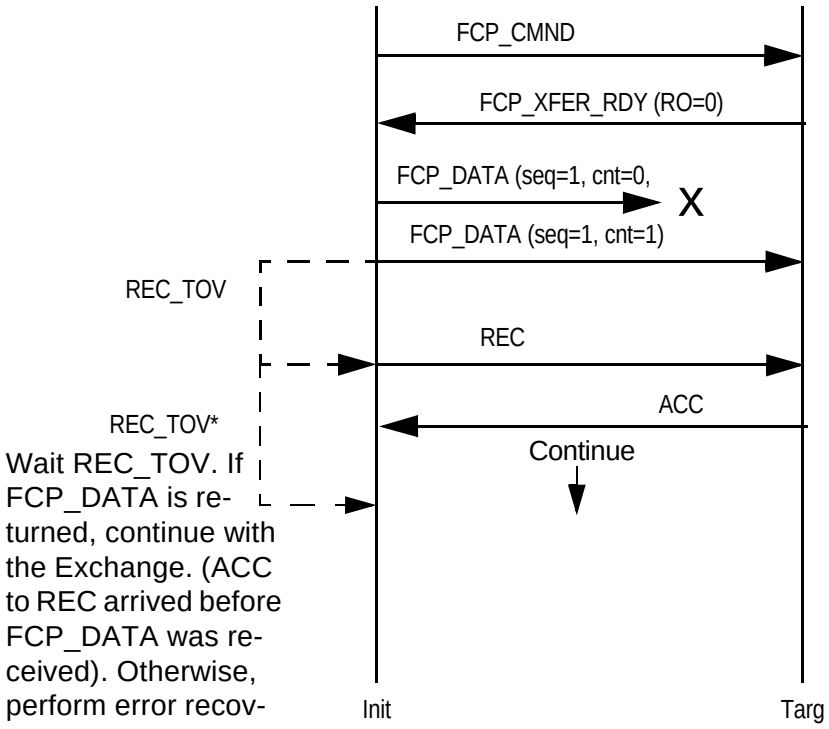


Figure D.15 - Lost Write Data, Not Last Frame of Sequence, Unacknowledged Classes

Error Detection



Error Recovery

The ACC for the REC indicates the Target does not hold Sequence Initiative and the Exchange is open. The Initiator sends an SRR requesting a FCP_XFER_RDY be resent using the specified Relative Offset. The Target retransmits the FCP_XFER_RDY using the specified Relative Offset. New SEQ_IDs are used for FCP_XFER_RDY and FCP_DATA retransmission.

If the lost frame arrives at the target prior to the expiration of REC_TOV, the exchange proceeds and no recovery is required.

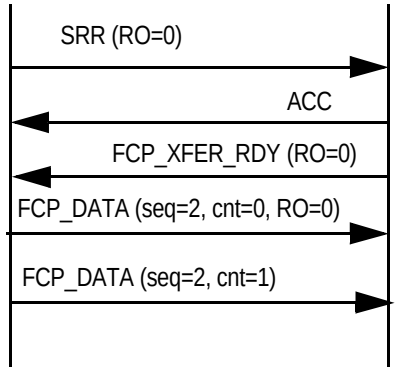
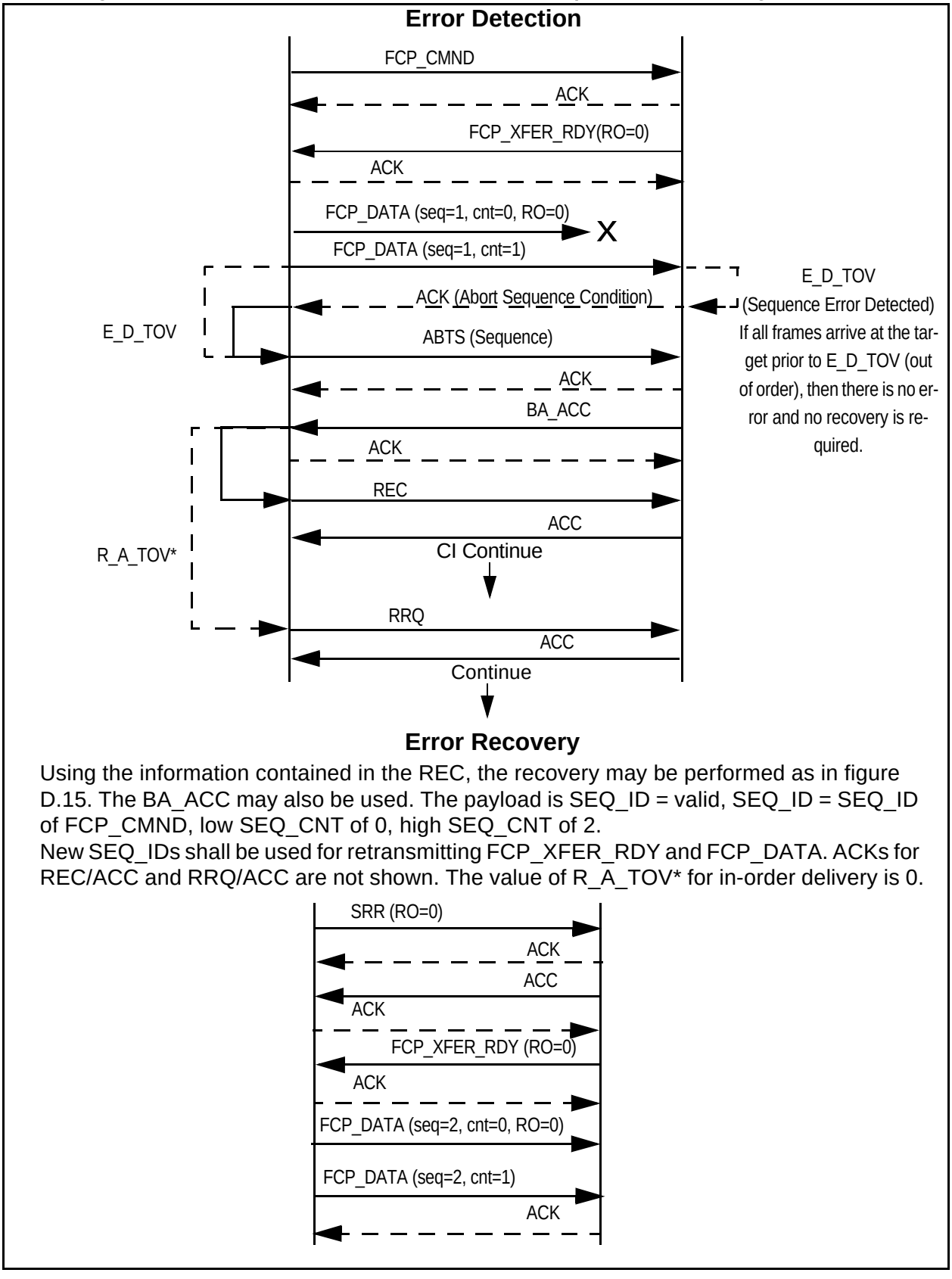


Figure D.16 - Lost Write Data, Not Last Frame of Sequence, Acknowledged Classes

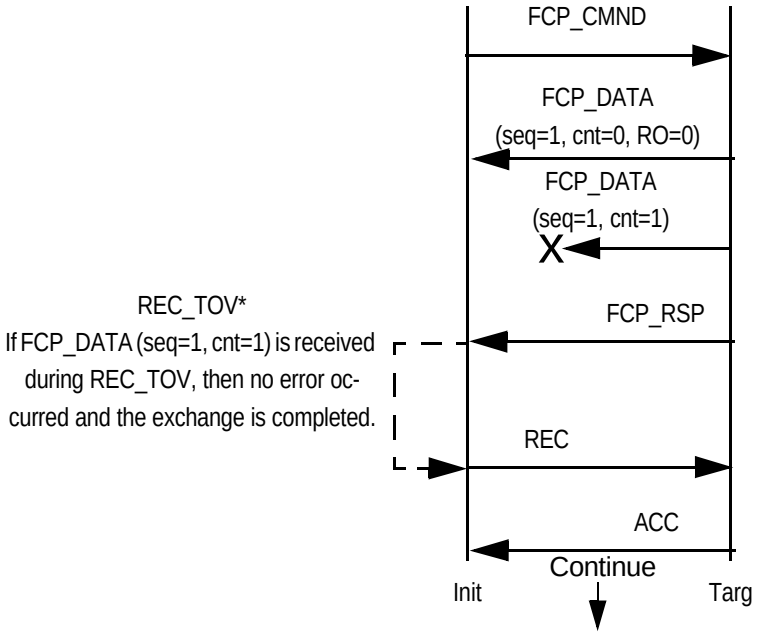


Using the information contained in the REC, the recovery may be performed as in figure D.15. The BA_ACC may also be used. The payload is SEQ_ID = valid, SEQ_ID = SEQ_ID of FCP_CMND, low SEQ_CNT of 0, high SEQ_CNT of 2.

New SEQ_IDs shall be used for retransmitting FCP_XFER_RDY and FCP_DATA. ACKs for REC/ACC and RRQ/ACC are not shown. The value of R_A_TOV* for in-order delivery is 0.

Figure D.17 - Lost Read Data, Last Frame of Sequence, Unacknowledged Classes

Error Detection



Error Recovery

The ACC for the REC indicates the Target does not hold Sequence Initiative and the Exchange is open. The Initiator sends an SRR requesting a FCP_DATA be resent using the specified Relative Offset. FCP_DATA retransmission uses a new SEQ_ID. For in-order delivery, the value of REC_TOV* is 0.

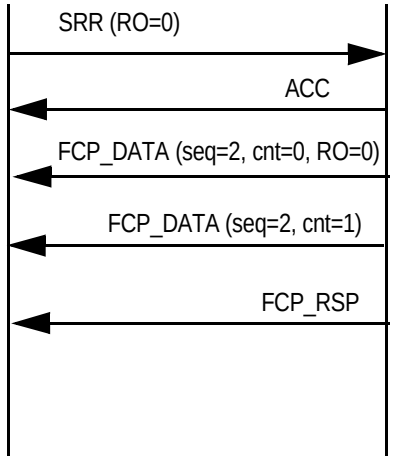
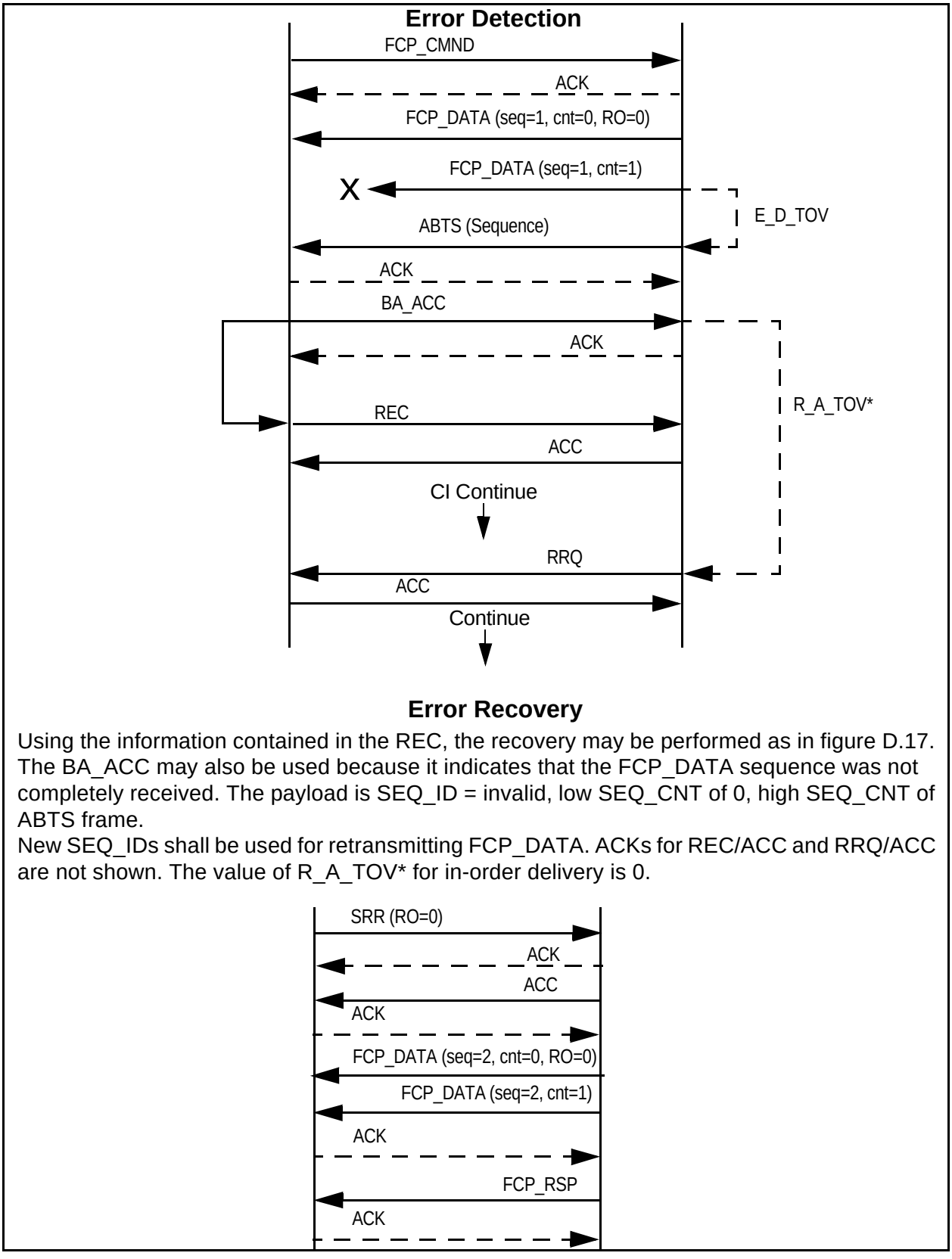


Figure D.18 - Lost Read Data, Last Frame of Sequence, Acknowledged Classes

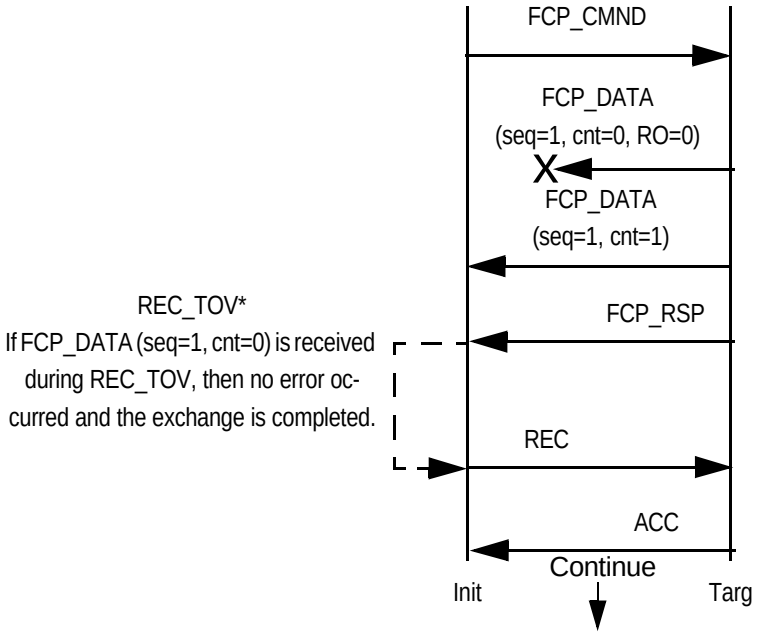


Using the information contained in the REC, the recovery may be performed as in figure D.17. The BA_ACC may also be used because it indicates that the FCP_DATA sequence was not completely received. The payload is SEQ_ID = invalid, low SEQ_CNT of 0, high SEQ_CNT of ABTS frame.

New SEQ_IDs shall be used for retransmitting FCP_DATA. ACKs for REC/ACC and RRQ/ACC are not shown. The value of R_A_TOV* for in-order delivery is 0.

Figure D.19 - Lost Read Data, Not Last Frame of Sequence, Unacknowledged Classes

Error Detection



Error Recovery

The ACC for the REC indicates the Target does not hold Sequence Initiative and the Exchange is open. The Initiator sends an SRR requesting a FCP_DATA be resent using the specified Relative Offset. FCP_DATA retransmission uses a new SEQ_ID. For in-order delivery, the value of REC_TOV* is 0.

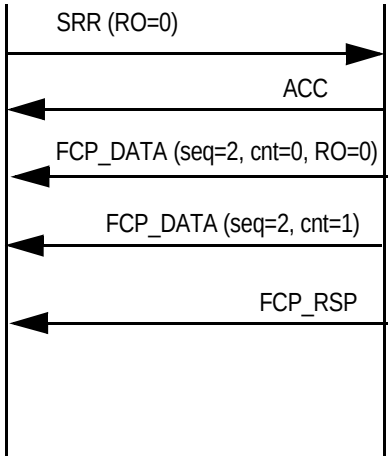
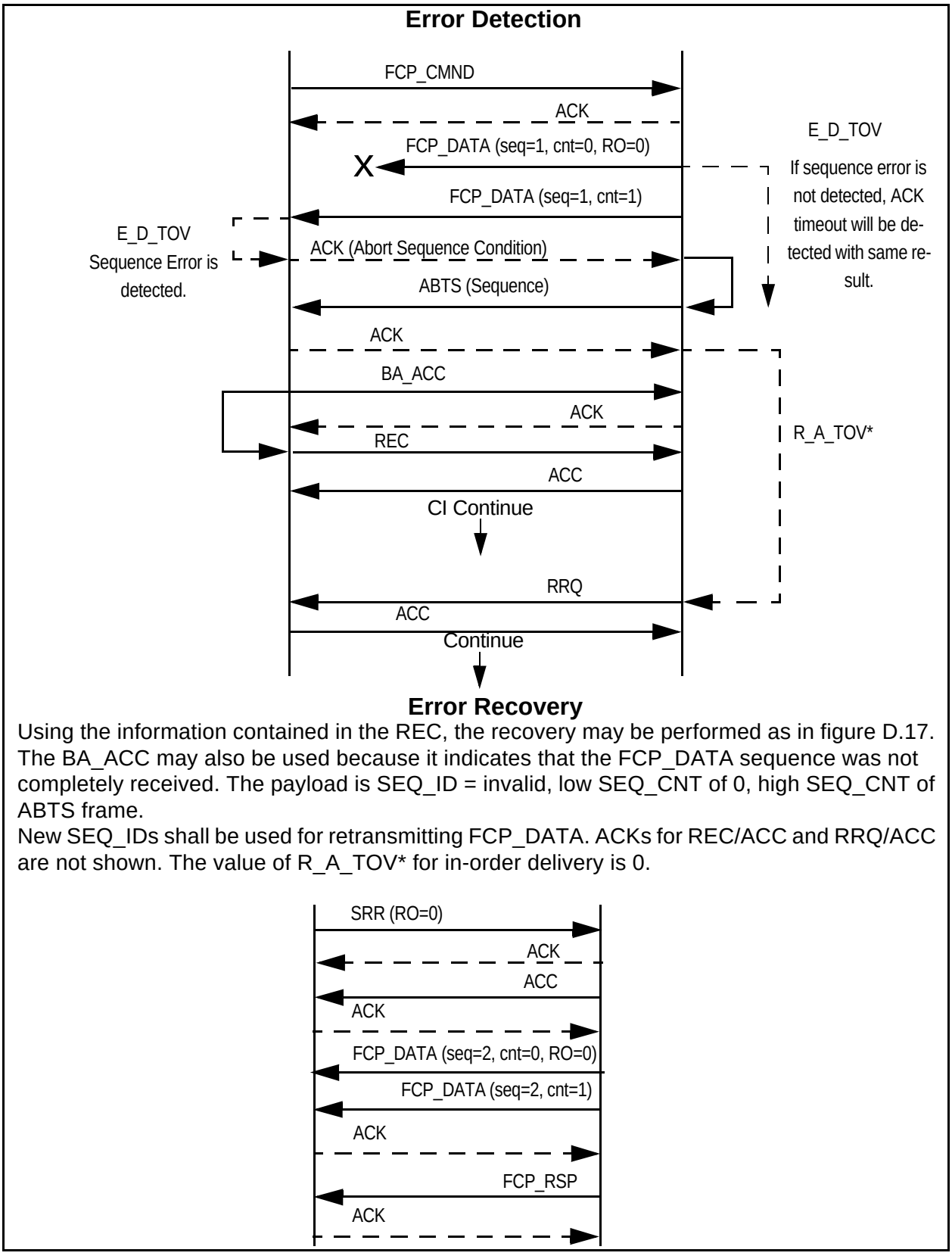


Figure D.20 - Lost Read Data, Not Last Frame of Sequence, Acknowledged Classes



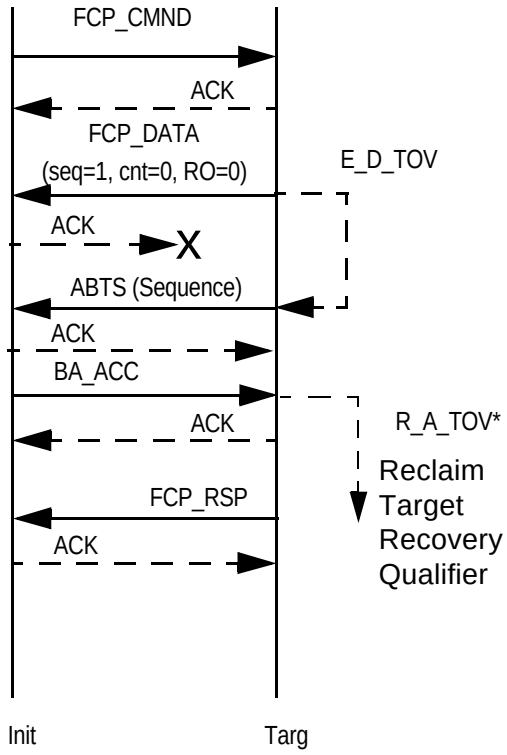
Using the information contained in the REC, the recovery may be performed as in figure D.17. The BA_ACC may also be used because it indicates that the FCP_DATA sequence was not completely received. The payload is SEQ_ID = invalid, low SEQ_CNT of 0, high SEQ_CNT of ABTS frame.

New SEQ_IDs shall be used for retransmitting FCP_DATA. ACKs for REC/ACC and RRQ/ACC are not shown. The value of R_A_TOV* for in-order delivery is 0.

Figure D.21 - ACK Lost on Read (Acknowledged Classes)

Error Detection

Acknowledged Classes



No Error Recovery

None:
The Initiator has received the FCP_DATA frame or sequence. No error recovery is required.

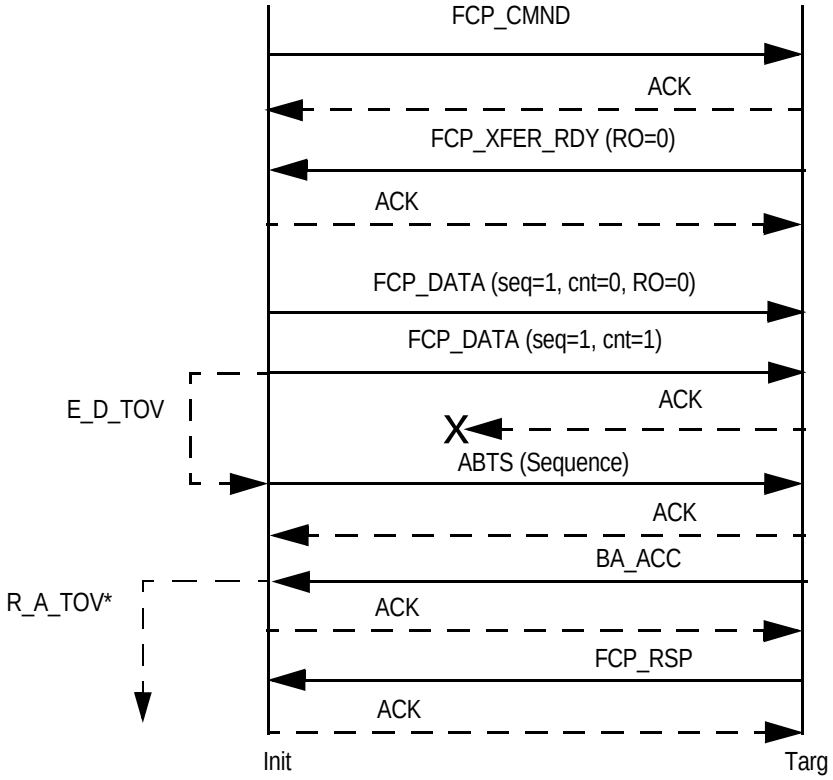
The BA_ACC indicates the FCP_DATA sequence was received (Payload is SEQ_ID valid, SEQ_ID value = SEQ_ID value of FCP_DATA sequence, low SEQ_CNT = high SEQ_CNT = SEQ_CNT of ABTS frame).

The target must establish its Recovery Qualifier. The resources associated with the Recovery Qualifier can be reclaimed after R_A_TOV. For in-order delivery, the value of R_A_TOV* is 0.

The issuance of RRQ is optional as no Recovery Qualifier was established by the initiator in this case. FCP_RSP can be received anytime after the transmission of FCP_CMND due to out of order delivery.

Figure D.22 - ACK Lost on Write (Acknowledged Classes)

Error Detection

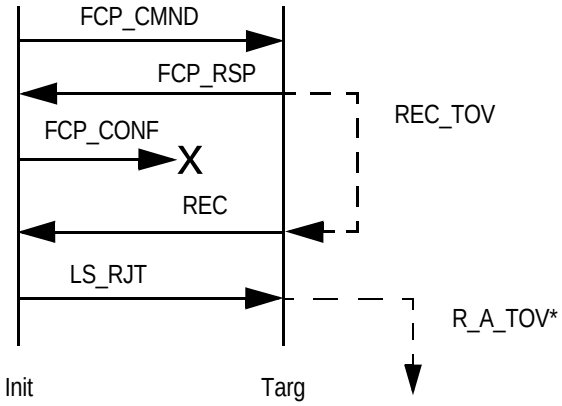


Error Recovery

None:
The Target has received the FCP_DATA sequence. No error recovery is required. The BA_ACC indicates the data sequence was received. The payload is set to SEQ_ID valid, SEQ_ID value = SEQ_ID value of the FCP_DATA Sequence, low SEQ_CNT = high SEQ_CNT = SEQ_CNT of ABTS frame. The target and initiator continue the Exchange. The initiator must establish its Recovery Qualifier. The resources associated with the Recovery Qualifier can be reclaimed after R_A_TOV. For in-order delivery, the value of R_A_TOV* = 0. The issuance of the RRQ is optional, as no Recovery Qualifier was established by the target. FCP_RSP can be received at any time after the last FCP_DATA frame has been transmitted.

Figure D.23 - FCP_CONF Lost, Unacknowledged Classes

Error Detection



No Error Recovery

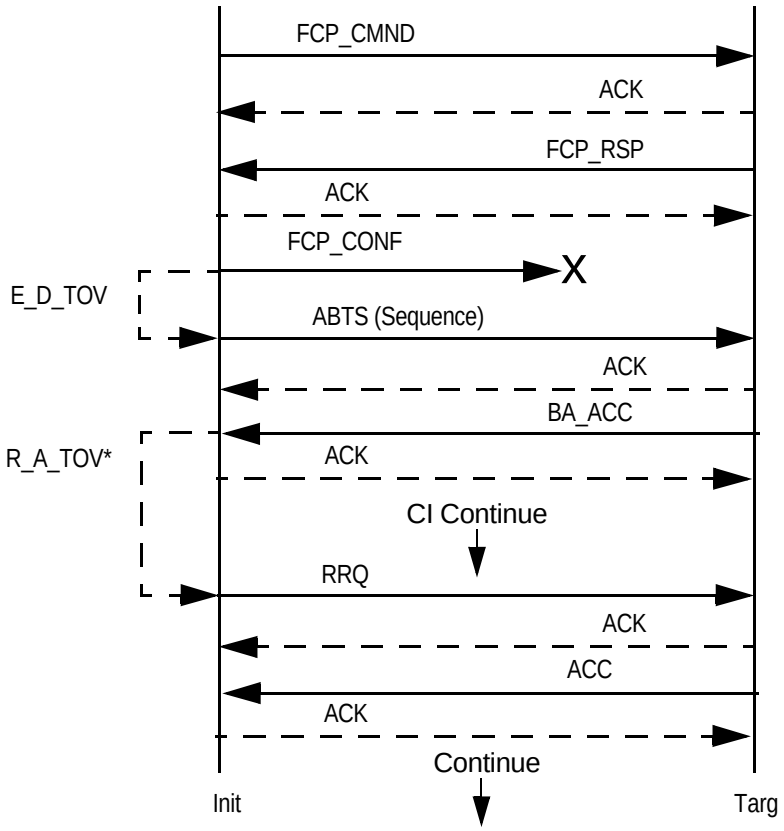
None.

The LS_RJT for the REC indicates the Exchange is unknown. That implicitly indicates that the initiator received FCP_RSP and sent FCP_CONF.

The context for the Exchange in the Target must be preserved for another R_A_TOV to prevent possible aliasing. If FCP_CONF is received after REC is transmitted, it is accepted and discarded and the context for the Exchange can be purged. For in-order delivery, the value of R_A_TOV* = 0.

Figure D.24 - FCP_CONF Lost, Acknowledged Classes

Error Detection



Error Recovery

The BA_ACC indicates that the FCP_CONF was not received. The BA_ACC payload is SEQ_ID valid, SEQ_ID value = SEQ_ID of FCP_CMND, Low SEQ_CNT = 0, High SEQ_CNT = 1). The value of R_A_TOV* for in-order delivery is zero. FCP_CONF must be retransmitted with a different SEQ_ID.

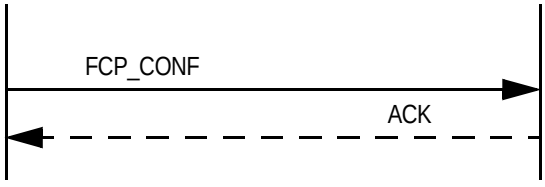
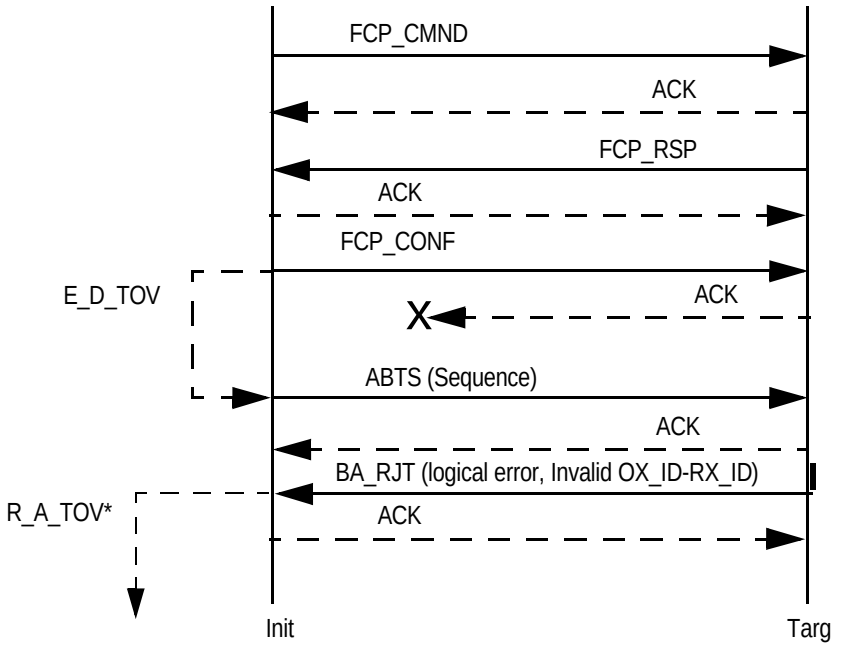


Figure D.25 - ACK lost on FCP_CONF, Acknowledged Classes

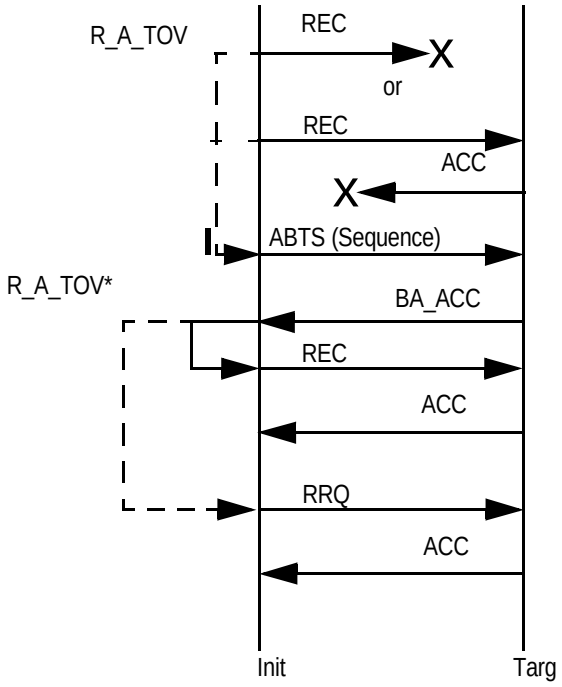
Error Detection



Error Recovery

None:
BA_RJT is the response to the ABTS, since no context exists for this Exchange and the ABTS was not issued on the first sequence of a new Exchange. The initiator must establish a Recovery Qualifier on receipt of the BA_RJT. The resources associated with the Recovery Qualifier can be reclaimed when R_A_TOV expires.
Issuance of the RRQ is optional, as no Recovery Qualifier was established by the target.
For in-order delivery, the value of R_A_TOV* = 0.

Figure D.26 - REC or REC Response Lost, Unacknowledged Classes

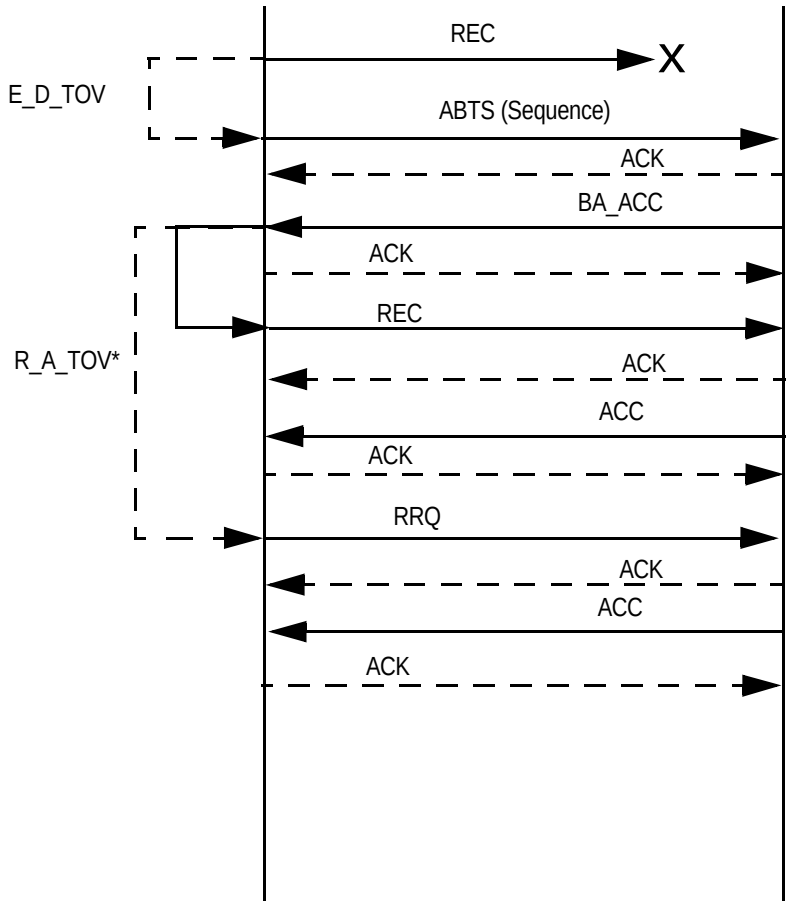


For the case of the REC never having been received, the BA_ACC payload will be SEQ_ID invalid, low SEQ_CNT = 0, high SEQ_CNT = SEQ_CNT of ABTS = 1.

For the case of the ACC response to REC never having been received, the target would view the ABTS as having been issued on a new Exchange. The BA_ACC payload will be SEQ_ID invalid, low SEQ_CNT = high SEQ_CNT = SEQ_CNT of ABTS.

In both cases, a Recovery Qualifier will be established. The second REC is issued in a new Exchange. For in-order delivery, the value of R_A_TOV* is 0

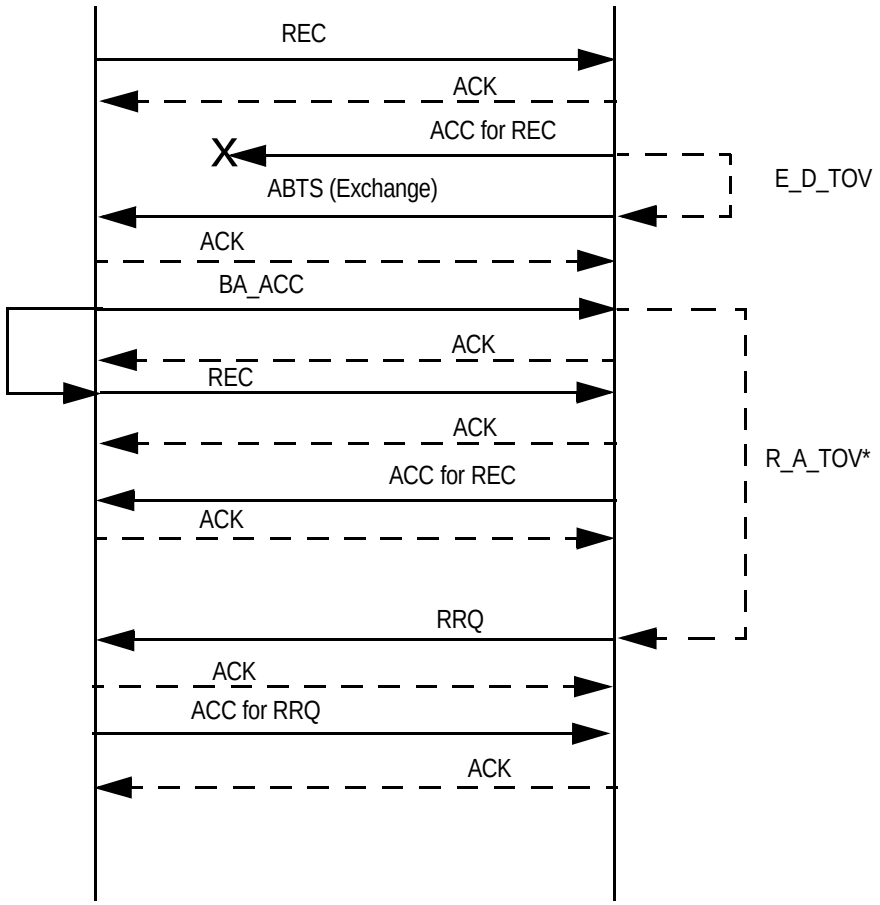
Figure D.27 - REC Lost, Acknowledged Classes



Error Recovery

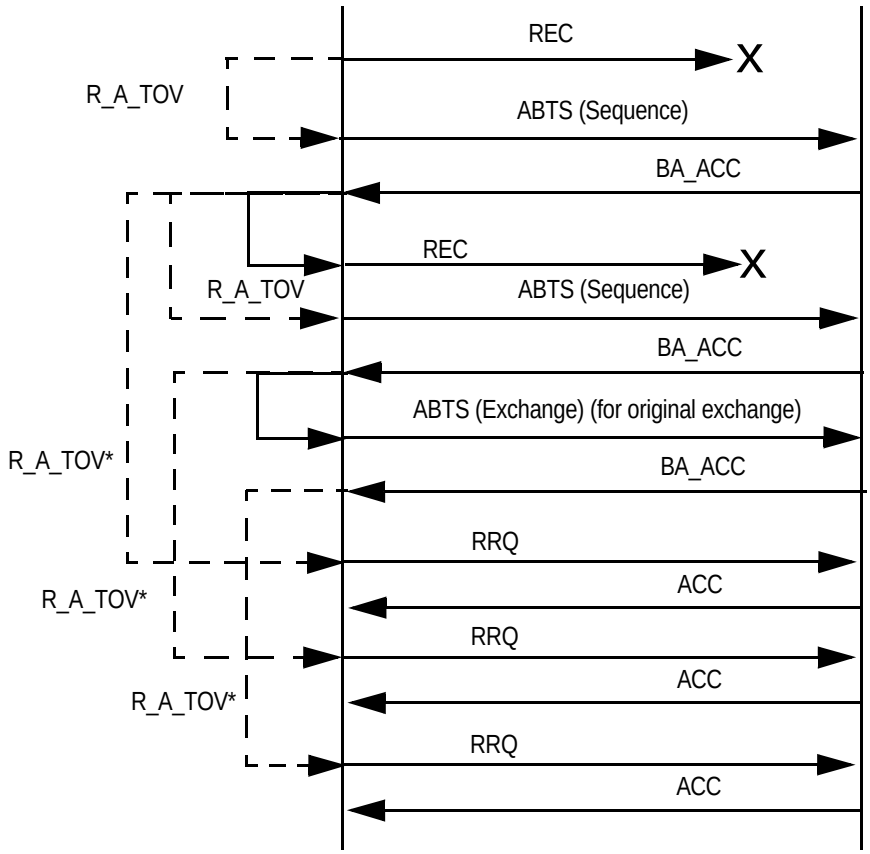
The BA_ACC payload indicates that the REC was never received by the target. The payload is SEQ_ID invalid, low SEQ_CNT = 0, high SEQ_CNT = SEQ_CNT in ABTS frame. Recovery Qualifiers are established on each side.
For in-order delivery, the value of R_A_TOV* is 0.
The second REC is issued using a new Exchange.

Figure D.28 - REC Response Lost, Acknowledged Classes



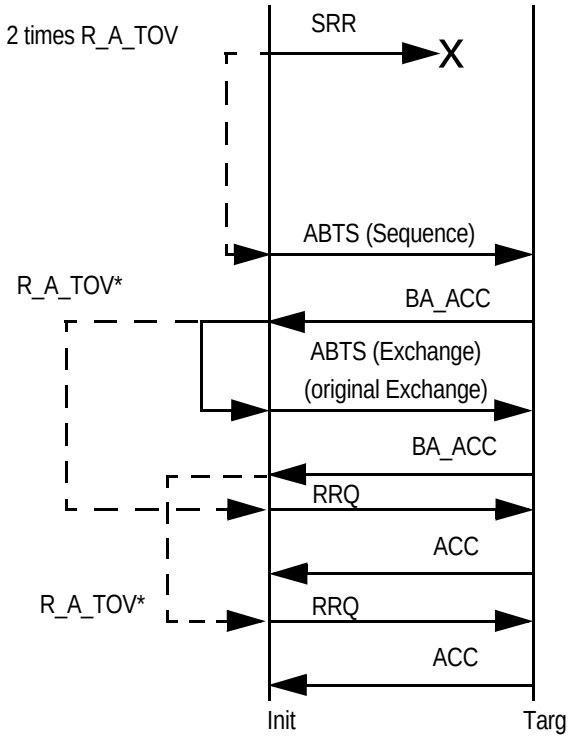
The BA_ACC payload indicates that the ACC was never received by the initiator. The payload is SEQ_ID invalid, low SEQ_CNT = 0, high SEQ_CNT = SEQ_CNT in ABTS frame.
After responding to the ABTS, the initiator reissues the REC in a new Exchange.
Recovery Qualifiers are established on each side.
For in-order delivery, the value of R_A_TOV* is 0.

Figure D.29 - Two RECs Lost, Unacknowledged Classes, Abort the original exchange



The failure of two RECs issued against the same Exchange causes all associated Exchanges to be aborted. The ABTS for the original Exchange uses the previous SEQ_ID and a SEQ_CNT one greater than the count used in the previous sequence and Bit 0 = 0 set in the Parameter field. The payload for the BA_ACC is SEQ_ID valid, SEQ_ID = SEQ_ID of last deliverable sequence received, low SEQ_CNT = 0 and high SEQ_CNT = FFFFh.
Recovery qualifiers are established on both sides for each Exchange.
For in-order delivery, the value of R_A_TOV* is 0.

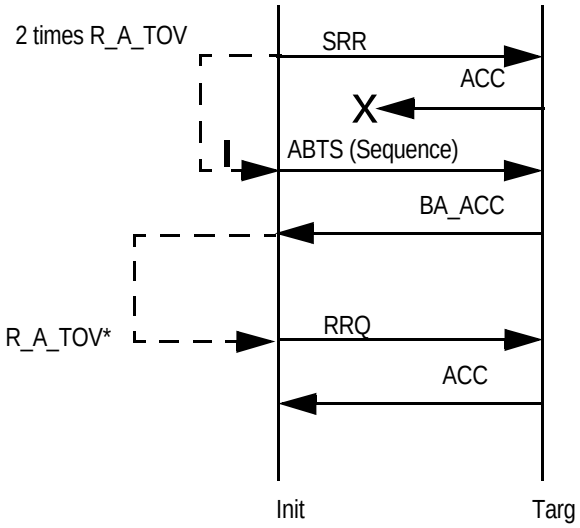
Figure D.30 - SRR Lost, Unacknowledged Classes, Abort original exchange



The payload for the BA_ACC associated with the ABTS of the SRR is SEQ_ID invalid, low SEQ_CNT = 0, high SEQ_CNT = SEQ_CNT of the ABTS frame.

The ABTS for the original Exchange uses the previous SEQ_ID and a SEQ_CNT one greater than the count used in the previous Sequence and Bit 0 = 0 in the Parameter field. The payload for the BA_ACC associated with the ABTS for the original Exchange is SEQ_ID valid, the SEQ_ID = SEQ_ID of the last deliverable Sequence of the original Exchange received, low SEQ_CNT = 0, and high SEQ_CNT = FFFFh.
Recovery Qualifiers are established on both sides for each Exchange. For in-order delivery, the value of R_A_TOV* is 0.

Figure D.31 - SRR Response Lost, Unacknowledged Classes



If the SRR Exchange is unknown to the Recipient, the Exchange was completed and the context purged. The payload for the BA_ACC is SEQ_ID invalid, low SEQ_CNT = 0, high SEQ_CNT = FFFFh. Recovery Qualifiers are established on both sides.

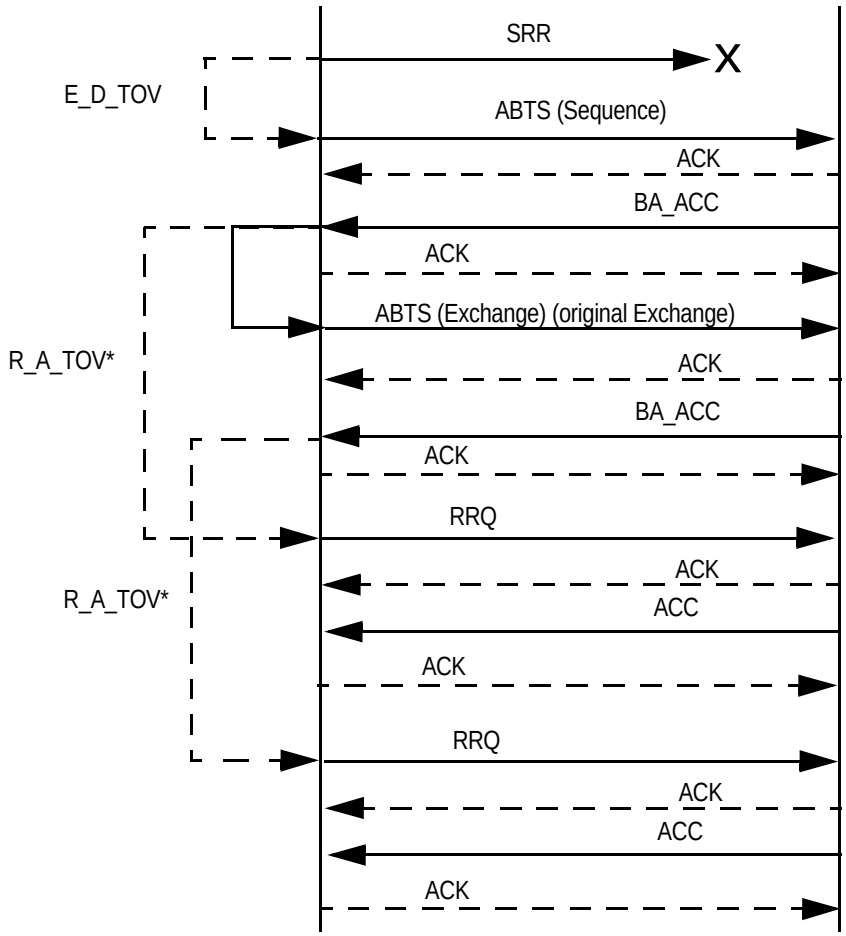
If the SRR Exchange is still known to the Recipient, the payload for the BA_ACC is SEQ_ID valid, SEQ_ID = SEQ_ID of the SRR, low SEQ_CNT = high SEQ_CNT = SEQ_CNT of the ABTS frame. Since no Recovery Qualifier is established, RRQ need not be issued. The Recovery Qualifier is established on the initiator side and must be timed out for R_A_TOV.

For in-order delivery, the value of R_A_TOV* is 0.

In either case, the original Exchange need not be aborted.

The RRQ references the exchange of the SRR.

Figure D.32 - SRR Lost, Acknowledged Classes

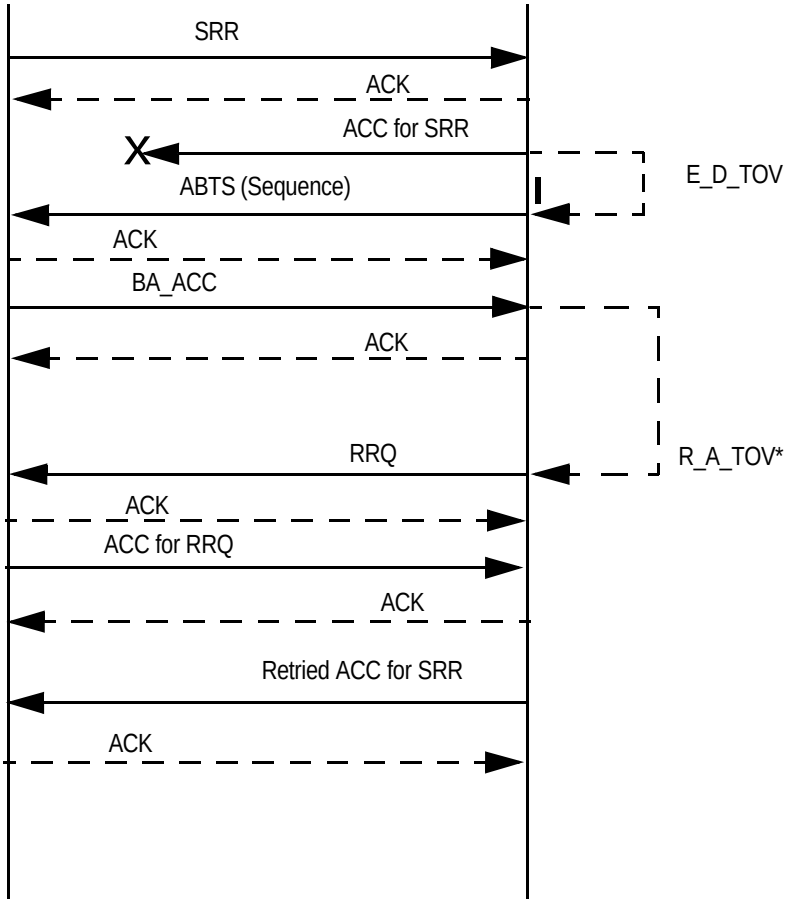


Since the ABTS on the SRR is on a new Exchange, Recovery Qualifiers must be established. The BA_ACC payload indicates SEQ_ID invalid, low SEQ_CNT = 0, and high SEQ_CNT = SEQ_CNT of the ABTS.

An error on an SRR is a second error and the original Exchange is also aborted. SRR is not retried.

For in-order delivery, the value of R_A_TOV* is 0.

Figure D.33 - SRR Response Lost, Acknowledged Classes



The BA_ACC of the ABTS associated with the SRR indicates that the ACC for the SRR was not received and will be discarded if it is later received. The BA_ACC payload indicates SEQ_ID invalid, low SEQ_CNT = 0, and high SEQ_CNT = SEQ_CNT of the ABTS.
The retry of the ACC for SRR is issued with a new SEQ_ID.
Recovery Qualifiers are established on each side.
For in-order delivery, the value of R_A_TOV* is 0.