FC-EE - Energy Efficient FIbre Channel
Feb. 3, 2013

1 Introduction

1.1 Overview

This proposal defines an energy efficient mode of operation for Fibre Channel. Much of the work in this proposal is based on Energy Efficient Ethernet (see IEEE std. 802.3-2012) with changes as necessary to support Fibre Channel concepts and operation.

For changes to existing Fibre Channel Standard text, the editorial notation is such: Text to be added is highlighted in blue, text to be removed is ~~crossed out in red.~~

1.2 Definitions

1.2.1 Low Power Idle (LPI): A primitive signal sent in place of Idle which indicates that the transmitter is operating in, or wishes to operate in Low Power mode (see 4).

1.2.2 LPI Mode: A link state in which the link is operating or wishing to operate in lower power mode by sending LPI (see 4.6).

2 Negotiation

2.1 FC-LS-3 Changes

{Editor: Make the following indicated changes to FC-LS-3 Table 140.}

Table 140 – Common Service Parameter applicability

Service Parameter	Word	Bits	Default Login Value	PLOGI and PLOGI LS_ACC Parameter applicability		FLOGI Parameter applicability		FLOGI LS_ACC Parameter applicability	
				Class		Class		Class	
				2	3	2	3	2	3
. . .									
Reserved	2	31-25	-	-	-	-	-	-	-
Energy Efficient Operation Supported	2	24	0	y	y	y	y	y	y
. . .									

Legend:
“y” indicates yes, applicable (i.e., has meaning);
“n” indicates no, not applicable (i.e., has no meaning);
“v” indicate the definition is vendor specific.

[a] E_D_TOV resolution and the corresponding value are only meaningful in a point-to-point topology or when doing PLOGI with an NL_Port on the same loop.

[b] The Common Service Parameter applicability is specified in FC-SP.

[c] Default buffer-to-buffer credit = 1 for all ports but an L_Port, and Buffer-to-buffer credit=0 for an L_Port.

[d] N_Port/F_Port=0 for an N_Port, and N_Port/F_Port=1 for an F_Port.

[e] BB_Credit Management=0 for an N_Port or F_Port, BB_Credit_Management=1 for an L_Port

6.6.2.7 Energy Efficient Operation Supported
0 = Nx_Port does not support Energy Efficient Operation
1 = Nx_Port supports Energy Efficient Operation

The Energy Efficient Operation Supported bit (Word 2, Bit 24), indicates support for Energy Efficient Operation. If set to one, the port supports Energy Efficient Operation, if set to zero, the port does not support Energy Efficient Operation (see FC-FS-4).

2.2 Changes for FC-SW-6

{Editor: Make the following changes to FC-SW-5 Table 6, ELP Request Payload.}

Table 6 – ELP Request Payload

Item	Size Bytes
10000000h	4
Revision	1
Flags	2
BB_SC_N	1
R_A_TOV	4
E_D_TOV	4
Requester Interconnect_Port_Name	8
Requester Switch_Name	8
Fabric Controller Class F Service Parameters	16
Class 1 Interconnect_Port Parameters	4
Class 2 Interconnect_Port Parameters	4
Class 3 Interconnect_Port Parameters	4
Reserved	20
ISL Flow Control Mode	2
Flow Control Parameter Length (N)	2
Flow Control Parameters	N

Flags: This field contains flag bits that provide additional information about the ELP. The following flag bit is defined.

Bit 15, the Bridge Port bit, shall indicate whether the sending port is a B_Port. If bit 15 is zero, the sending port is an E_Port and not a B_Port. If bit 15 is one, the sending port is a B_port.

Bit 14, the Bridge Virtual Fabrics bit, is meaningful only for a B_Port and shall indicate whether the sending B_Port supports Virtual Fabric Tagging. If bit 14 is zero, the sending B_Port does not support the passing through of VFT tagged frames (see FC-FS-4). If bit 14 is one, the sending B_port supports the passing through of VFT tagged frames.

Bit 13-12, <Used by FC-BB-6.>

Bit 11, the Energy Efficient Operation Supported bit.

Bits 10-0 shall be reserved.

2.3 EEEP ELS

{Editor: Add the following ELS to FC-LS-3.}

4.2.47 Exchange Energy Efficient Parameters (EEEP)

4.2.47.1 Description

The Exchange Energy Efficient Parameters (EEEP) ELS may be used by an Nx_Port in order to set the values of some the Energy Efficient parameters to values other than the default. The EEP ELS is used by an Nx_Port to notify the remote Nx_Port of it's desired parameters. It is not required that both Nx_Ports on a link use the same set of parameters as set by EEEP.

4.2.47.2 Protocol

a) Exchange Energy Efficient Parameters (EEEP) Request Sequence; and
b) LS_ACC or LS_RJT Reply Sequence.

4.2.47.3 Request Sequence

Addressing: The S_ID designates the source port that is sending the EEEP Request. The D_ID shall be the F_Port Controller Well-known address (i.e., FFFFFEh).

Payload: The format of the EEEP Payload is shown in table 7.

Table 7 – EEEP Payload

Bits Word	31 .. 24 Byte 0	23 .. 16 Byte 1	15 .. 08 Byte 2	07 .. 00 Byte 3
0	EEEP (<TBD>h)	00h	00h	00h
1	Descriptor List Length = 20			
2	EEEP Descriptor			
3				
4				
5				
6				

Table 8 – EEEP Descriptor Format

Bits Word	31 .. 24 Byte 0	23 .. 16 Byte 1	15 .. 08 Byte 2	07 .. 00 Byte 3
0	Descriptor Type = <TBD>h			
1	Length = 4			
2	Transmit T_w		Receive T_w	

Transmit T_w: Sets the Transmit T_w parameter. Value is specified in µs.

Receive T_w: Sets the Receive T_w parameter. Value is specified in µs.

4.2.47.4 Reply Sequence

LS_RJT: LS_RJT signifies the rejection of the EEEP Request

LS_ACC: LS_ACC signifies acceptance of the EEEP Request and returns the requested exchange information. The format of the LS_ACC Payload is shown in table 9.

Table 9 – EEEP LS_ACC Payload

Bits Word	31 .. 24 Byte 0	23 .. 16 Byte 1	15 .. 08 Byte 2	07 .. 00 Byte 3
0	LS_ACC (02h)	00h	00h	00h
1	Descriptor List Length = 0			

2.4 EEEP SW_ILS

{Editor: Add the following SW_ILS to FC-SW-6.}

6.1.31 Exchange Energy Efficient Parameters (EEEP)

6.1.31.1 Exchange Energy Efficient Parameters (EEEP) Request Payload

The Exchange Energy Efficient Parameters (EEEP) SW_ILS may be used by an E_Port in order to set the values of some the Energy Efficient parameters to values other than the default. The EEP ELS is used by an E_Port to notify the remote E_Port of it's desired parameters. It is not required that both E_Ports on a link use the same set of parameters as set by EEEP.

Protocol:

Exchange Energy Efficient Parameters (EEEP) request Sequence
Accept (SW_ACC) Reply Sequence

Addressing: The S_ID field shall be set to FFFFFDh, indicating the Fabric Controller of the originating Switch. The D_ID field shall be set to FFFFFDh, indicating the Fabric Controller of the destination Switch.

Payload: The format of the EEEP request Payload is shown in table 10.

Table 10 – EEEP Request Payload

Item	Size Bytes
Command code = <TBD>h	4
Descriptor List Length = 20	4
EEEP Descriptor	see 6.1.31.2

6.1.31.2 Exchange Energy Efficient Parameters (EEEP) Descriptor

Table 11 shows the EEEP Descriptor.

Table 11 – EEEP Descriptor

Item	Size Bytes
Tag Value = <TBD>h	4
Length = 4	4
Transmit T_w	2
Receive T_w	2

Transmit T_w: Sets the Transmit T_w parameter. Value is specified in µs.

Receive T_w: Sets the Receive T_w parameter. Value is specified in µs.

6.1.31.3 Exchange Energy Efficient Parameters (EEEP) SW_ACC Payload

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
Signifies the rejection of the EEEP request
Accept (SW_ACC)
Signifies acceptance of the EEEP request.
– Accept Payload

Payload: The format of the EEEP SW_ACC Payload is shown in table 12.

Table 12 – EEEP SW_ACC Payload

Item	Size Bytes
Command code = 0200 0000h	4
Descriptor List Length = 0	4

3 Encoding

{Modify FC-FS-4 Clause 5.3.1 as indicated.}

5.3.1 Overview

If the FC_Ports on a link determine to use FEC, the streams of 64B/66B Transmission Words in both directions on the link shall be encoded as specified in 5.3 and then further encoded as specified in subclause 74.7 and subclause 74.10 of IEEE 802.3-2008 and IEEE P802.3az. If the FC_Ports on a link determine not to use FEC, the streams of 64B/66B Transmission Words in both directions on the link shall be encoded as specified in 5.3.

{Modify FC-FS-4 Clause 5.3.6 as indicated.}

5.3.6 Control Transmission Words

{Modify table 10 as indicated.}

Table 10 – Valid 64B/66B Transmission Word type values

Transmission Word type value	Transmission Word content	Reference
1Eh	Idle or LPI Special Function followed by Idle or LPI Special Function; or Receiver Error	5.3.6.1 5.3.6.10
33h	Idle Special Function followed by SOF Special Function	5.3.6.2
B4h	EOF Special Function followed by Idle or LPI Special Function	5.3.6.3
2Dh	Idle Special Function followed by other Special Function	5.3.6.4
4Bh	Other Special Function followed by Idle Special Function	5.3.6.5
55h	Other Special Function followed by other Special Function	5.3.6.6
66h	Other Special Function followed by SOF Special Function	5.3.6.7
78h	SOF Special Function followed by word of data	5.3.6.8
FFh	Word of data followed by EOF Special Function	5.3.6.9
any other value	Restricted for IEEE 802.3-2008, shall not be transmitted	IEEE 802.3-2008

{Modify paragraph 3 as indicated.}

Idle Special Functions and receiver detected errors shall be represented as a series of four 7-bit control codes (see table 11). FC_Ports compliant with this standard shall not encode the control codes other than the following into a transmission word:

a) Idle (i.e., 00h), or

b) LPI (i.e., 06h), if the FC_Port supports Energy Efficient Fibre Channel.

If a control code value other than Idle, or LPI if the FC_Port supports Energy Efficient Fibre Channel, is decoded, the Transmission Word shall cause a code violation to be reported and the restricted control code shall be decoded as an Idle control code.

To communicate LPI Mode (see 4), the LPI control code (i.e., 06h) is sent in place of the Idle control code (i.e., 00h)

{Modify table 11 as indicated.}

Table 11 - Valid control code values

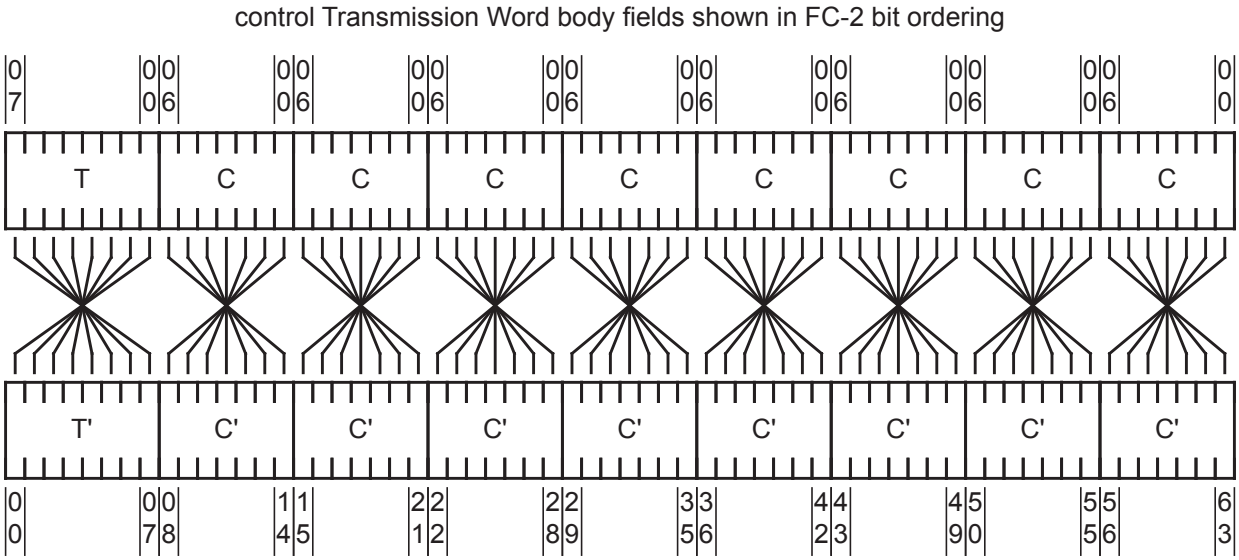
Value (least significant seven bits)	Meaning	Reference
00h	Idle	5.3.7.2
06h	LPI	4
1Eh	Error. This code shall be used only for receiver error reporting (see 5.3.6.10)	5.3.6.10
any other value	Restricted for IEEE 802.3-2008, shall not be transmitted	IEEE 802.3-2008

{Modify FC-FS-4 Clause 5.3.6.1 as indicated.}

5.3.6.1 **Idle** or LPI **/ Idle** or LPI

{General editorial comment. The notation used in this clause "<code> / <code>" is not defined. Either it needs to be defined, or we should drop the shortcut and say "<code> followed by <code>" in all cases.}

If the control Transmission Word represents transmission of an Idle or LPI Special Function followed by an Idle or LPI Special Function, the body of the control Transmission Word shall be composed as shown in figure 1. In each field, lower numbered bits represent less significant bits of the value than higher numbered bits.



Key:

T	Transmission Word type value set to 1Eh
C	7-bit control code set to zero (i.e., the Idle control code), or 06h (i.e., the LPI Control code)

NOTE Each field with a name marked by ' is shown in transmission bit order. It has the same numeric value as the field with the same unmarked name.

Figure 1 - 64B/66B control Transmission Word body: Idle or LPI **/ Idle** or LPI

{Modify FC-FS-4 Clause 5.3.6.3 as indicated.}

5.3.6.3 **EOF / Idle** or LPI

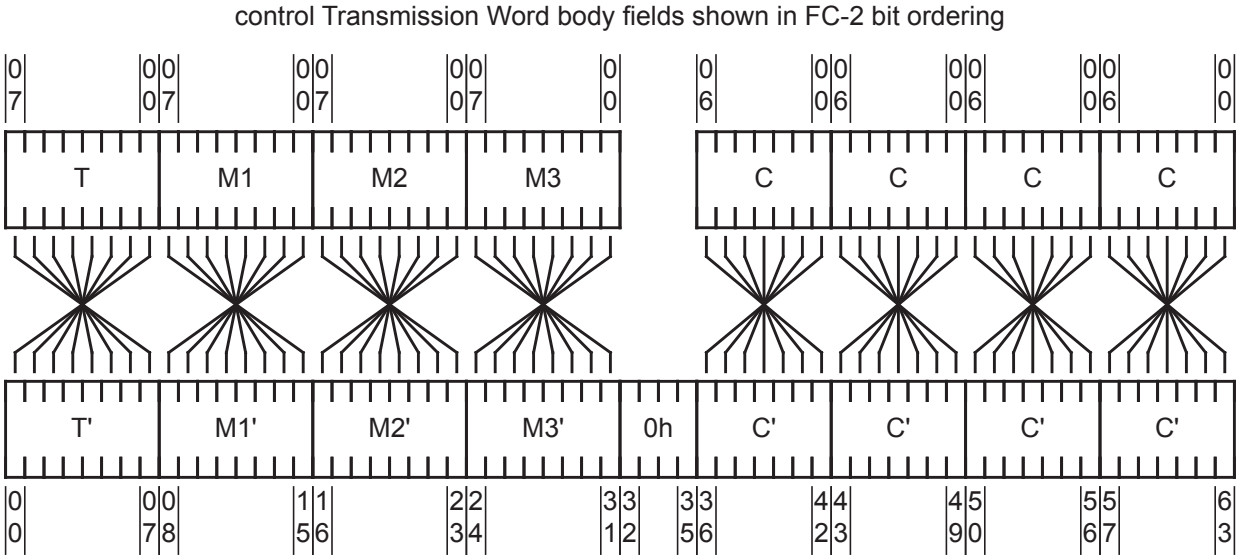
If the control Transmission Word represents transmission of an EOF Special Function followed by an Idle or LPI Special Function, the body of the control Transmission Word shall be composed as shown in figure 2. In each field, lower numbered bits represent less significant bits of the value than higher numbered bits.

If the Transmission Word received following receiving an EOF / Idle or LPI Transmission Word:

a) is a data Transmission Word;
b) is any transmission word containing an EOF; or
c) causes a coding violation to be reported,

then the EOF / Idle or LPI Transmission Word shall cause a code violation to be reported and shall be decoded as two Idle Special Functions.

NOTE 1 - This requires lookahead on encountering an EOF. The code violations based on the following Transmission Word reflect behavior required by the Receive state machine in IEEE 802.3-2008 subclause 49.2.13.3.



Key:

T	Transmission Word type value set to B4h
M1, M2, M3	Modifier bytes for EOF (see 5.3.7.1)
C	7-bit control code set to zero (i.e., the Idle control code), or 06h (i.e., the LPI Control code)

NOTE Each field with a name marked by ' is shown in transmission bit order. It has the same numeric value as the field with the same unmarked name.

Figure 2 - 64B/66B control Transmission Word body: EOF / Idle or LPI

4 FC-EE Operation

{Editors Note: The placement of this section in a final standard is TBD, but it is expected that the majority of the FC-EE Operation text will be placed in FC-FS-4.}

4.1 Overview

The Energy Efficient Fibre Channel capability provides a protocol and associated physical layer capabilities to allow a Fibre Channel link to operate at a lower power level. The goal of the Energy Efficient Fibre Channel is:

a) provide a protocol to allow transitions to and from a lower power level;

b) allow such transition to occur without changing the link status, dropping, or corrupting frames; and

c) provide a transition time that is small enough such that it is transparent to the upper level protocols (i.e., minimum impact on link bandwidth and latency).

Energy Efficient operation is negotiated per link using a login bit either in the FLOGI/PLOGI, for N_Ports, or in the ELP for E_Ports (see xxx). Wake parameters may be set using the EEEP ELS, for N_Ports (see xxx), or the EEEP SW_ILS, for E_Ports (see xxx).

Energy Efficient operation is achieved by entering the Low Power Idle (LPI) mode (see xxx). During Low Power Idle mode, the link is still active, but enters periods of lower power level operation. When one of the link partners has data to transmit, a wake-up signal is sent to indicate that the link should return to a full power operation.

Energy Efficient operation is not supported on NL_Ports.

4.2 FC-EE Negotiation

Energy Efficient operation is negotiated during FC_Port Login according to the following:

a) For N_Ports operating without a fabric, Word 2, Bit 24 of the Common Service Parameters in the PLOGI and PLOGI LS_ACC is set to 1 to indicate Energy Efficient operation support (see FC-LS-3);

b) For N_Ports connecting to a fabric, Word 2, Bit 24 of the Common Service Parameters in the FLOGI and FLOGI LS_ACC is set to 1 to indicate Energy Efficient operation support (see FC-LS-3). For N_Ports connected to a fabric, the Energy Efficient operation bit in any subsequent PLOGI or PLOGI LS_ACC shall be ignored; and

c) For E_Ports bit 12 in the Flags field of the ELP is set to 1 to indicate Energy Efficient operation support (see FC-SW-6).

In order for any particular link to support Energy Efficient operation both N_Ports or E_Ports of the link shall indicate support for Energy Efficient operation.

4.3 FC-EE and FEC

For FC_Ports which support FEC (see 9), a port implementing FC-EE shall implement the FEC rapid block synchronization as defined in clause 74 of IEEE P8023az.

4.4 Alert Signal

The Alert Signal is sent to indicate wake up from quiet mode. The Alert Signal is a repeating FF00h pattern.

4.5 Transmitter Turn Off

During the quiet cycle, some transceiver types may not be capable of turning off the transmitter/receiver. In this case, LPI shall be transmitted during the LPI Mode in order to indicate low power operation, this allows the port to turn off unused capabilities to save power.

{Should specify the transceiver types that are allowed to bypass turning off the transceiver.}

4.6 LPI Mode

4.6.1 Overview

Energy Efficient operation is accomplished by entering LPI Mode. LPI Mode operation is indicated by a set of Primitive Signals:

a) The transmitter indicates a request for LPI Mode by transmitting the LPI in place of Idle. The process by which this is accomplished depends on the link encoding (see FC-FS-4).

b) The receiver is notified of the link partner request for entry into LPI Mode by receipt of a LPI in place of Idle.

While in LPI Mode the data traffic transmission is disabled if the transmitter/receiver pair supports it (see 4.5, and the link operates in a quiet/refresh cycle until one of the link partners indicates a change back to full power operation by sending a wake signal for a predetermined amount of time. This wake signal consists of sending regular Idle (i.e., not an LPI) across the link. One reason for a return to full power operation would be the presence of data to transmit.

Figure 3 shows an overview of LPI Mode operation. In LPI Mode, after the sleep time (i.e., T_s, the link cycles between a Quiet time (i.e., T_q) and a refresh cycle (i.e. T_r). The wake time is defined as in TBD.

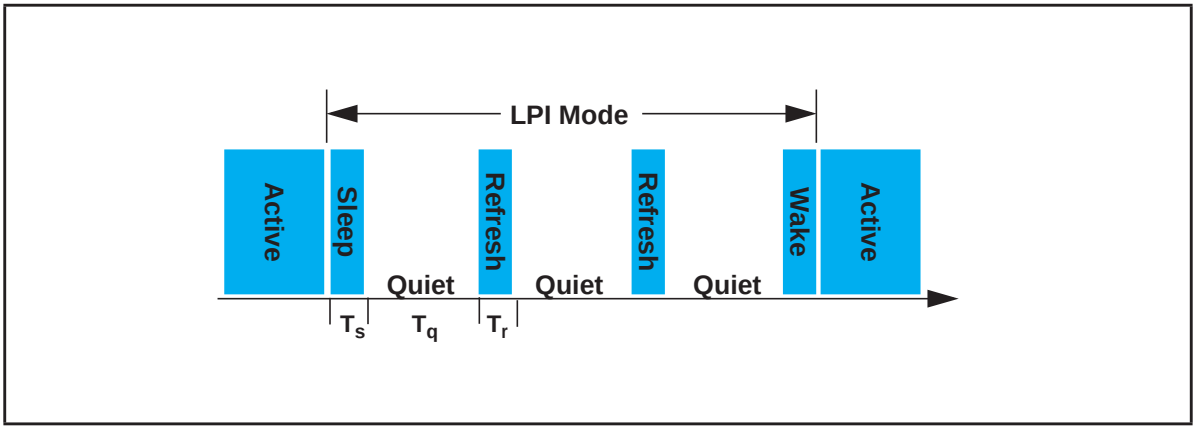


Figure 3 – Overview of LPI Mode operation

4.6.2 LPI Mode Entry

An FC_Port shall enter and exit LPI Mode from the Active State only (see FC-FS-4).

NOTE 1 – For 64B/66B encoding this means that the only valid code transitions for LPI are Idle to LPI, LPI to Idle, and EOF to LPI (see FC-FS-4).

4.6.3 LPI Mode Timing parameters

For LPI Mode operation, the Transmitter State Diagram timing parameters shall be as defined in table 11, and the Receiver State Diagram timing parameters shall be as defined in table 12:

Table 11 – Transmitter LPI timing parameters

Parameter	Description	16GFC		32GFC		Units
		Min	Max	Min	Max	
Alert_Timer	Time spent in the Alert state	1.1	1.3	1.1	1.3	µs
Bypass_Timer	Time spent in the SCR Bypass state	0.9	1.1	0.9	1.1	µs
T_s	Sleep Time from entering the Sleep state to when tx_mode is set to QUIET	4.9	5.1	0.9	1.1	µs
T_q	Quiet Time from when tx_mode is set to QUIET to entry into the Alert state	1.7	1.8	1.7	1.8	ms
T_w	Time spent in the Wake state	9.5	9.7	3.9	4.1	µs

Table 12 – Reciever LPI timing parameters

Parameter	Description	16GFC		32GFC		Units
		Min	Max	Min	Max	
T_q	The time the receiver waits for energy_detect to be set to TRUE while in the Sleep and Quiet states before asserting receive fault	2	3	2	3	ms
T_w	Time the receiver waits in the Wake state before indicating a wake time fault. (when scr_bypass_enable = FALSE)	10.1		N/A[a]		µs
T_w	Time the receiver waits in the Wake state before indicating a wake time fault. (when scr_bypass_enable = TRUE)	12.3		5.7		µs
T_{wtf}	Wake time fault recovery time	10		10		ms

[a] For 32GFC this timer has no meaning since FEC is required (i.e., scr_bypass_enable will never be FALSE).

4.6.4 FC-EE State Diagrams

4.6.4.1 LPI Mode Transmitter State Diagram

Figure 4 shows the state diagram for the transmitter in LPI Mode.

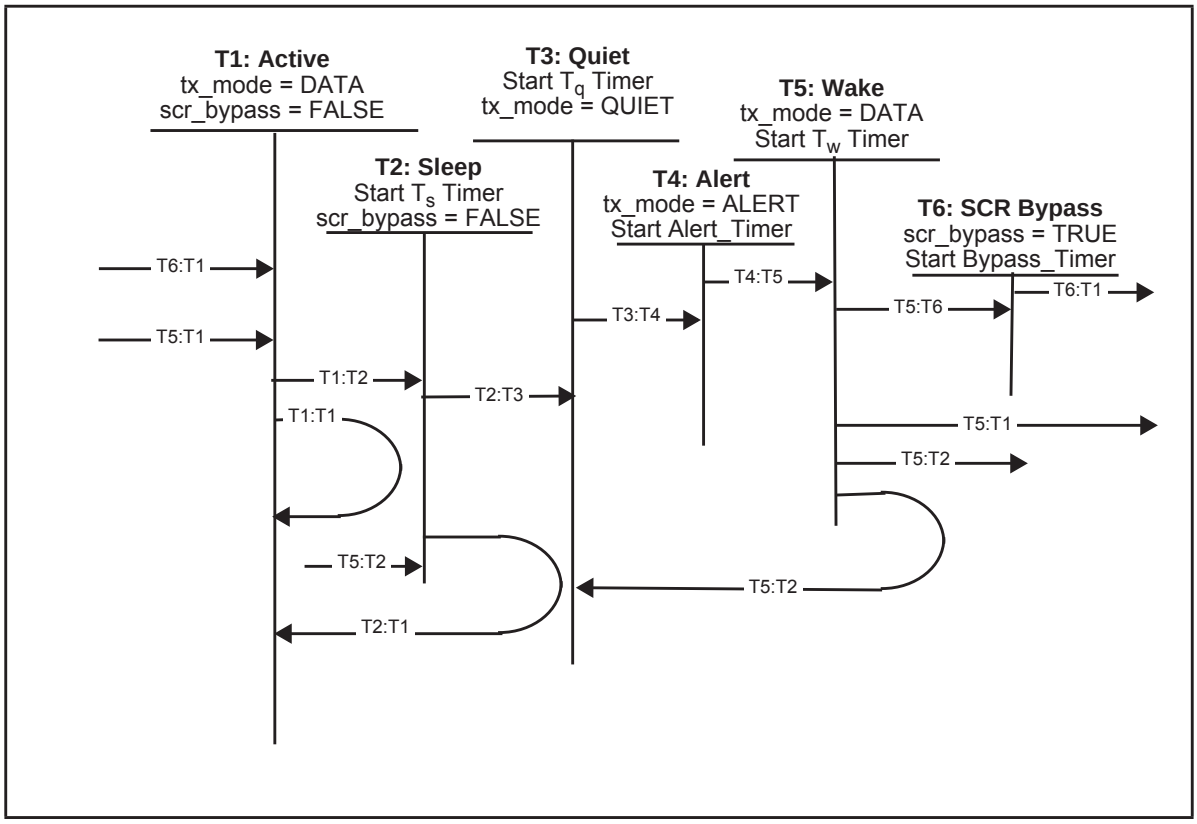


Figure 4 – LPI Mode Transmitter State Diagram

State T1 Active: This state is normal Fibre Channel operation. This state is entered upon entry into the Active state in the FC_Port state machine (see 7). The link is running at full power and data may be transmitted. tx_mode is set to DATA, and scr_bypass is set to FALSE.

Transition T1:T2: The port wishes to enter LPI Mode, LPI is transmitted and the T_s Timer is started.

Transition T1:T1: The port wishes to stay in Active mode, IDLE is transmitted.

State T2 Sleep: The port transmits LPI in place of Idle to indicate entry into LPI Mode. The T_s timer is started.

Transition T2:T1: The port transmits Idle instead of LPI, indicating exit from LPI Mode.

Transition T2:T3: The T_s timer has expired and LPI continues to be transmitted.

State T3 Quiet: Local port has entered Quiet mode. If the local port supports it (see 4.5), the Transmitter is turned off. The tx_mode variable is set to QUIET. T_q timer is started.

Transition T3:T4: T_q timer has expired.

State T4 Alert: Set tx_mode to ALERT and send Alert Signal (see 4.4) until Alert_Timer expires.

Transition T4:T5: The Alert_Timer has expired.

State T5 Wake: Start T_w timer.

Transition T5:T6: T_w timer has expired. Set scr_bypass to TRUE (i.e., disable 64B/66B scrambling). Start Bypass_Timer.

Transition T5:T1: T_w timer has expired and local port wishes to exit LPI Mode. Idle is transmitted to indicate to remote port that it should exit LPI Mode.

Transition T5:T2: T_w timer has expired and local port wishes to stay in LPI Mode. Return to Sleep mode.

State T6 SCR Bypass: Disable 64B/66B scrambling for 1 µs.

Transition T6:T5: The 1 µs timer has expired, re-enable 64B/66B scrambling.

Transition T6:T1: Local port wishes to exit LPI Mode. Idle is transmitted to indicate to remote port that it should exit LPI Mode.

4.6.4.2 LPI Mode Receiver State Diagram

Figure 5 shows the state diagram for the receiver in LPI Mode.

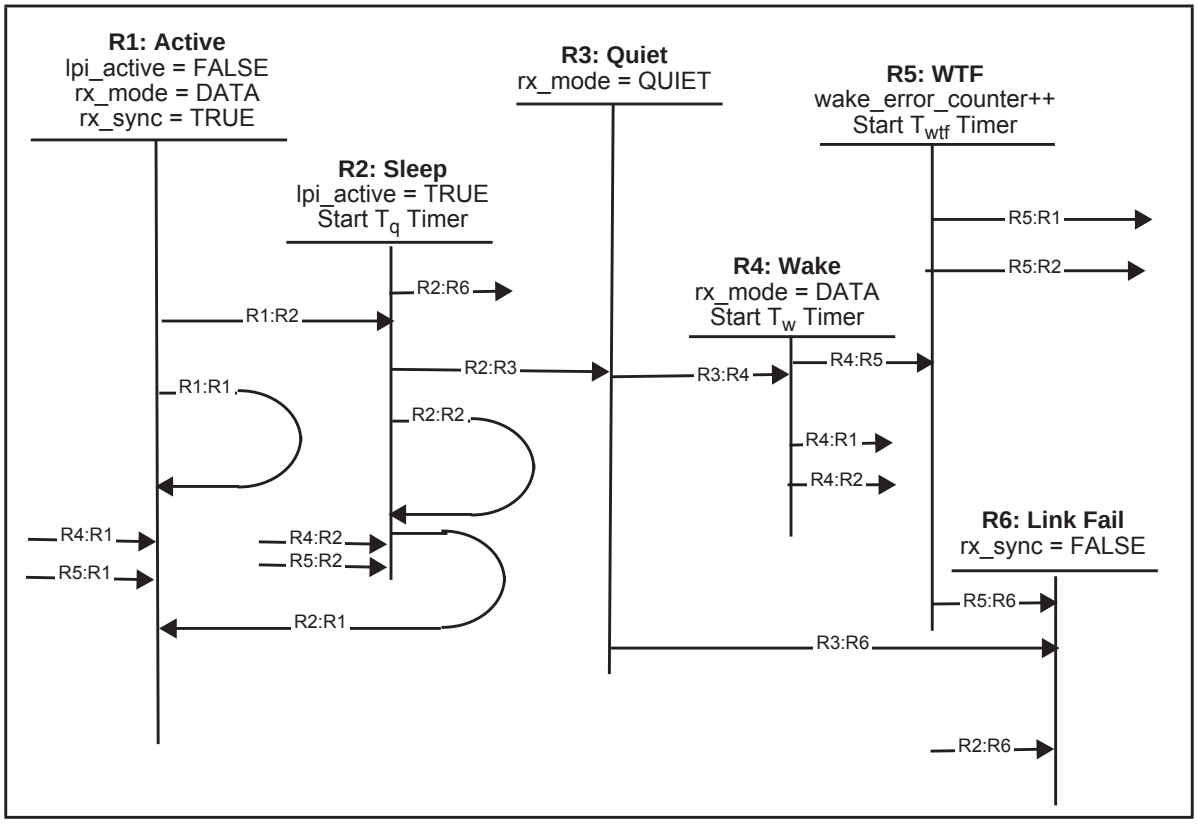


Figure 5 – LPI Mode Receiver State Diagram

State R1 Full Active: This state is normal Fibre Channel operation. This state is entered upon entry into the Active state in the FC_Port state machine (see 7). The link is running at full power and data may be received. The variables lpi_active is set to FALSE, rx_mode is set to DATA, and rx_sync is set to TRUE.

Transition R1:R1: If rx_sync is not equal to TRUE return to R1.

Transition R1:R3: The local port receives LPI in place of Idle. The variable lpi_active is set to TRUE. The Tq TImer is started.

State R2 Sleep: The local port has received LPI indicating that the remote port wishes to enter LPI Mode. Set lpi_actvie to TRUE. Start T_q timer.

Transition R2:R2: If rx_sync is TRUE and T_q Timer had not expired, and received LPI.

Transition R2:R1: if rx_sync is TRUE and T_q Timer has not expired, and received IDLE.

Transition R2:R3: T_q timer has expired and rx_sync is FALSE.

Transition R2:R6: Tq timer has expired.

State R3 Quiet: Local port has entered Quiet mode. Remote transmitter has been turned off, if the port supports it (see 4.5), and T_q timer is started. Rx_mode is set to QUIET.

Transition R3:R4: Energy detect on link.

Transition R3:R6: Energy not detected, and T_q timer has expired.

State R4 Wake: Remote transmitter is turned back on. Rx_mode is set to Data, Tw timer is started.

Transition R4:R1: T_w timer has not expired and rx_sync is TRUE and IDLE received.

Transition R4:R2: T_w timer has not expired and rx_sync is TRUE and LPI received.

Transition R4:R5: T_w timer has expired.

State R5 WTF: Increment wake error counter. Start T_{wf} timer.

Transition R5:R1: Twf timer has not expired and rx_sync is TRUE and IDLE received.

Transition R5:R2: Twf timer has not expired and rx_sync is TRUE and LPI received.

State R6 Link Fail: Port is in Link Failure and shall enter the link initialization state machine (see FC-FS-4).