



IBM Systems and Technology Group

***FC-EE Parameter Negotiation Proposal -
T11 Document 14-107v0***

19-Mar-2014

By: Adrian Butter
asbutter@us.ibm.com

Special Thanks To:

***Scott Kipp
Howard Johnson
Anil Mehta
Craig Carlson***

Agenda

- Recap of Energy Efficient Parameters Discussion at Feb T11 FC-FS-4 Meeting
- Introduce “Energy Efficient Operation Support Level” Service Parameter
- Clarify possible Fibre Channel Energy Efficient (FC-EE) operating states
 - ◆ Including identification of which states are “Default”, “Configured” or “Negotiated”
- Flow chart illustration of FC-EE operating state determination
- Proposed FC-LS-3 and FC-SW-6 text updates to cover “Energy Efficient Operation Support Level” Service Parameter

Summary of Energy Efficient Parameters Discussion from Feb 2014 FC-FS-4 Meeting

➤ “Fast Wake” Assessment (Item IBM-18 of 14-010v1)

- o Discussed benefits of introducing Fast Wake support into the LPI Transmit and Receive state diagrams
 - ◆ Provides a mechanism to save power when the transmitter and/or receiver connected by an FC link does not support the “Quiet” state
 - Consistent with current text description contained in FC-FS-4 Section 10.5...
 - ◆ For 32GFC operation, introducing Fast Wake support is a differentiating feature not included in 802.3 single link configurations
- o Per response to IBM-18 contained in 14-058v0, proposed solution dispositioned as “Accepted in Principle”

➤ Wake Time Negotiation (Item IBM-22 of 14-010v1)

- o Discussed the benefit of providing flexible Wake time support via a negotiation process
 - ◆ Provides a mechanism to maximize power savings when the transmitter and/or receiver connected by an FC link requires longer Refresh time periods between “Quiet” states
- o Proposed solution dispositioned as “More Work Required”
 - ◆ Wake Time Negotiation was previously proposed, but not included because no use cases required its use
 - ◆ Next steps are for old proposal to be refreshed based on the material contained in 14-010v1 and 14-036v0, and thereafter present to FC-FS-4 for debate and dispositioning (Owner: Craig Carlson)

o If “Accepted”, such negotiation must be optional

➤ Energy Efficient Parameter Negotiation (14-036v0)

- o Proposed an update to the current Exchange Energy Efficient Parameters descriptor contained in the latest FC-EE document (see Tables 8 and 11 of 13-011v1)
 - ◆ Negotiated parameters - Wake Time and Fast Wake

Proposed “Energy Efficient Level” Service Parameter

- 2-Bit field replacing “Energy Efficient Operation Supported” parameter defined in the latest FC-EE specification (13-011v1)
- Advantage - Provides one service parameter which captures allowed states of two key FC-EE parameters (“Fast Wake” and “Transmitter Wake Time”)
 - NOTE:** Since default range settings only apply to the Transmitter (per FC-FS-4 Table 24), only “Transmitter Wake Time” is subject to (optional) negotiation
- “Energy Efficient Level” definition
 - o '00' = Energy Efficient operation is not supported (default value prior to successful port Login)
 - o '01' = Minimal support for Energy Efficient operation
 - ◆ “Fast Wake” support only
 - o '10' = Full support for Energy Efficient operation
 - ◆ Either “Quiet” support with default Transmitter Wake Time (as defined by the Tw parameter in FC-FS-4 Table 24)...
 - ◆ ...or “Fast Wake” support
 - NOTE:** “Quiet” or “Fast Wake” mode usage determined by service parameter exchange
 - o '11' = Full support for Energy Efficient operation
 - ◆ Either “Quiet” support with negotiated Transmitter Wake Time...
 - ◆ ...or “Fast Wake” support as determined after service parameter exchange
 - NOTE:** “Quiet” or “Fast Wake” mode usage determined by service parameter exchange and (if applicable) Transmitter Wake Time negotiation

Proposed “Energy Efficient Level” Service Parameter (cont)

- Rules for selecting final “Energy Efficient Level” (EE_Lvl) based on Local (Local_EE_Lvl) and Remote (Remote_EE_Lvl) setting

NOTE: Local_EE_Lvl and Remote_EE_Lvl correspond to the “Energy Efficient Level” Service Parameter configured by the local and remote systems, respectively, during the port Login procedure.

- o If either side does not support Energy Efficient operation, then set EE_Lvl to '00' (for no EE operation support)
- o If both sides support Energy Efficient operation and there is a support level mismatch between the Local and Remote sides, then set EE_Lvl to '01' (for minimal level of EE operation support)
- o If both sides support Energy Efficient operation and the support level matches between the Local and Remote sides
 - ◆ If Local_EE_Lvl and Remote_EE_Lvl are '01' or '10', then set EE_Lvl to that value
 - ◆ If Local_EE_Lvl and Remote_EE_Lvl are '11' and Transmitter Wake Time negotiation is successful, then set EE_Lvl level to that value and use the negotiated Transmitter Wake Time
 - ◆ Otherwise, if Local_EE_Lvl and Remote_EE_Lvl are '11' and Transmitter Wake Time negotiation fails, then set EE_Lvl to '01' (for minimal level of EE operation support)

FC-EE Operating States

FC-EE State	EE State	Inactive Channel Transmission	Default/ Configured/ Negotiated	Entry Procedure	
				FC-LS-3	FC-SW-6
Not Supported	Not Supported	Idle	Default	PLOGI/FLOGI w/ EE_Lvl set to '00' (default value)	ELP Request w/ EE_Lvl set to '00' (default value)
Supported	"Fast Wake"	LPI	Configured OR Negotiated	PLOGI/FLOGI w/ EE_Lvl set to '01' OR PLOGI/FLOGI followed by EEEP w/ Local_TTw<>Remote_TTw => EE_Lvl set to '01'	ELP Request w/ EE_Lvl set to '01' OR ELP Request followed by EEEP w/ Local_TTw<>Remote_TTw => EE_Lvl set to '01'
	"Quiet" w/ Default Wake Time	No signal (Quiescent)	Configured	PLOGI/FLOGI w/ EE_Lvl set to '10'	ELP Request w/ EE_Lvl set to '10'
	"Quiet" w/ Negotiated Wake Time		Negotiated	PLOGI/FLOGI followed by EEEP w/ Local_TTw=Remote_TTw => EE_Lvl set to '11'	ELP Request followed by EEEP w/ Local_TTw=Remote_TTw => EE_Lvl set to '11'

Definitions:

Default = Energy Efficient Level value to be used prior to a successful Login

Configured = Energy Efficient Level value which may be selected after a successful Login based on extent of system EE support

Negotiated = Energy Efficient Level value which may be selected after a successful Login and successful Wake Time negotiation

PLOGI = N_Port Login; FLOGI = Fabric Login; ELP = Exchange Link Parameters; EEEP = Exchange Energy Efficient Parameters

Local_TTw = Local system Transmitter Wake Time parameter pertaining to a given link

Remote_TTw = Remote system Transmitter Wake Time parameter pertaining to a given link

LPI = Low Power Idle

Energy Efficient Level Translation

Energy Efficient Level Setting		Transmitter Wake Time		Energy Efficient Operating State	LPI Permitted?	lpi_fw Setting	Transmitter Wake Time Setting
Local	Remote						
00	xx	x	x	Not Supported	No	x	x
xx	00	x	x				
01	>00	x	x	"Fast Wake"	Yes	TRUE	x
>00	01	x	x				
10	10	Default	Default	"Quiet"	No	FALSE	Default
10	11	Default	TTw1	"Fast Wake"	Yes	TRUE	x
11	10	TTw1	Default				
11	11	TTw1	TTw1	"Quiet"	No	FALSE	TTw1
11	11	TTw1	TTw2	"Fast Wake"	Yes	TRUE	x

NOTES:

- "Energy Efficient Level Setting" and "Transmitter Wake Time" values are determined during the Port Login, Service Parameter Exchange and (optionally) Exchange Energy Efficient Parameters sequences
- "LPI Permitted?", "lpi_fw Setting" and "Transmitter Wake Time Setting" values control operation of the LPI Mode Transmitter and Receiver state diagrams contained in FC-FS-4
- LPI = Low Power Idle
- lpi_fw = Variable controlling the wake mode associated with the LPI Mode Transmitter and Receiver state diagrams contained in FC-FS-4; TRUE denotes "Fast Wake" mode, FALSE denotes "Quiet" mode (see 14-036v0 Slide 7)
- x = Don't Care
- "Default" refers to one of the Transmitter Tw ranges specified in FC-FS-4 Table 24 based on the operating mode (16GFC or 32GFC) of the FC link
- "TTw1" and "TTw2" denote two different Transmitter Wake Time values not contained within the relevant Transmitter Tw range specified in FC-FS-4 Table 24 based on the operating mode (16GFC or 32GFC) of the link

Updated FC-EE Common Service Parameter Definition for FC-LS-3

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	31-26	-	-	-	-	-	-	-
Energy Efficient Level	2	25-24	00	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.

6.6.2.7 Energy Efficient Level

00 = Nx_Port does not support Energy Efficient Operation (default Login value)

01 = Nx_Port provides minimal support for Energy Efficient Operation

 "Fast Wake" mode only

10 = Nx_Port provides full support for Energy Efficient Operation (see FC-FS-4)

 Either "Quiet" mode or "Fast Wake" mode is selected following Common Service Parameter exchange

11 = Nx_Port provides full support for Energy Efficient Operation with EEEP ELS (see FC-FS-4)

 Either "Quiet" mode or "Fast Wake" mode is selected following Common Service Parameter exchange and the Exchange Energy Efficient Parameters Extended Link Service

The Energy Efficient Level bits (Word 2, Bits 25-24) indicate the extent to which Energy Efficient Operation is supported. When set to 00, the port does not support Energy Efficient Operation; Otherwise, the port supports Energy Efficient Operation at one of three levels, and (as applicable) with or without the EEEP ELS. For details about how local and remote system settings of Energy Efficient Level determine the Energy Efficient operating state of the FC link, refer to FC-FS-4.

Updated FC-EE Extended Link Parameter Definition for FC-SW-6

Table 6 – ELP Request Payload

Item	Size Bytes
10000000h	4
Revision	1
Flags	2

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

...

Bits 11-10, the Energy Efficient Level bits, shall indicate the extent to which the port supports Energy Efficient Operation. If bits 11-10 are set to '00' (the default value), the port does not support Energy Efficient Operation. If bits 11-10 are otherwise set, the port supports Energy Efficient Operation at one of three levels, and (as applicable) with or without the Exchange Energy Efficient Parameters Switch Internal Link Service (EEEP SW_ILS). If set to '01', the port supports only "Fast Wake" mode. If set to '10', the port supports either "Quiet" mode without EEEP SW_ILS or "Fast Wake" mode. If set to '11', the port supports either "Quiet" mode with EEEP SW_ILS or "Fast Wake" mode. For the '10' setting, either "Quiet" mode or "Fast Wake" mode is selected following Extended Link Parameter (ELP) exchange. For the '11' setting, either "Quiet" mode or "Fast Wake" mode is selected following ELP exchange and EEEP SW_ILS. For details about how local and remote system settings of Energy Efficient Level determine the Energy Efficient operating state of the FC link, refer to FC-FS-4.

...

Summary

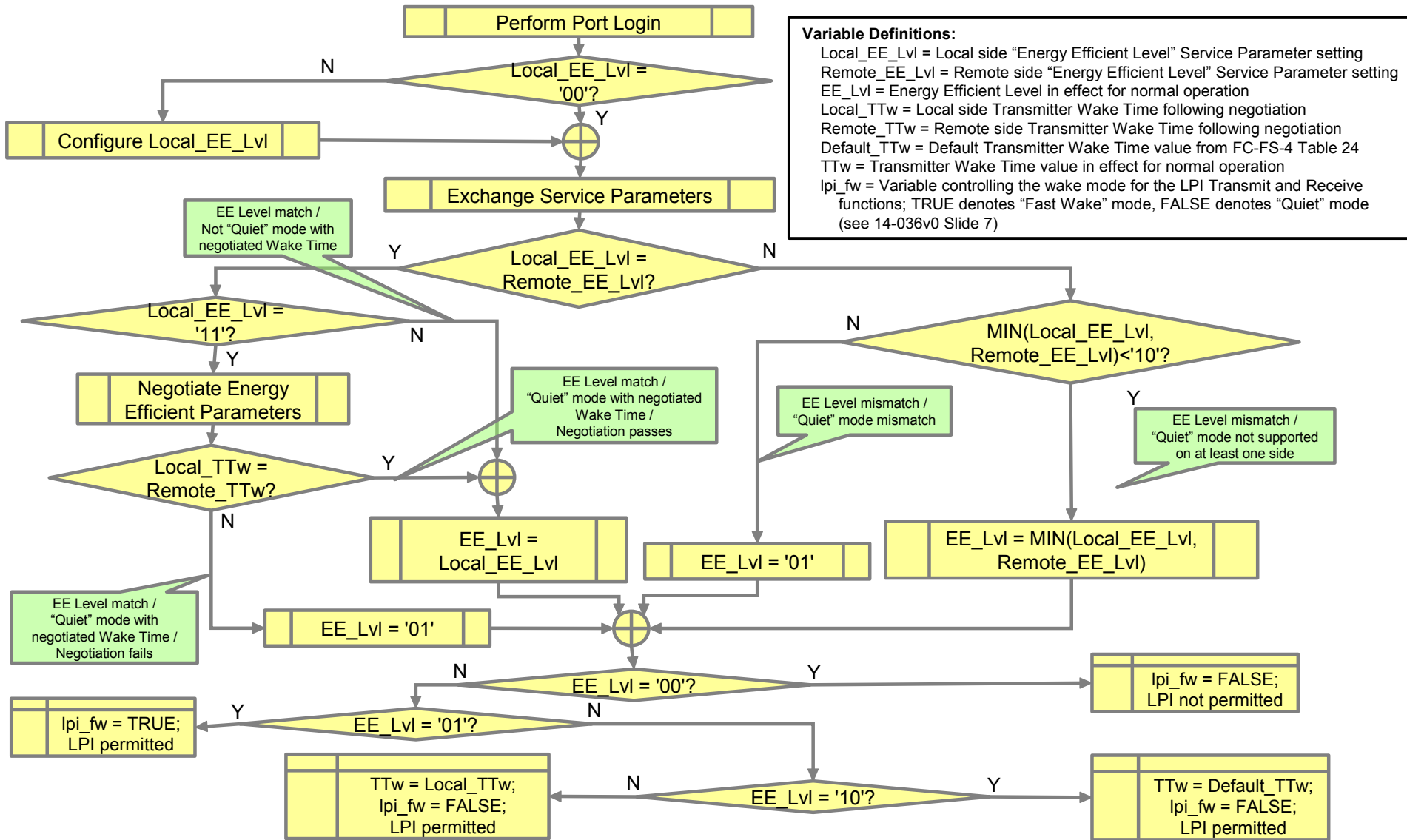
- Introduced “Energy Efficient Operation Support Level” service parameter
 - Encompasses key EE parameters used in both “Fast Wake” and “Quiet” mode
- Proposed Energy Efficient Parameter negotiation for Transmitter Wake Time
- Defined 4 Energy Efficient operating states for a FC link
 - Identified which states are “Default”, “Configured” or “Negotiated”
- Defined how the 4 Energy Efficient operating states are selected following the Port Login, Service Parameter Exchange and (optionally) Exchange Energy Efficient Parameters sequences.
 - ***Procedure ensures Energy Efficient Parameter negotiation is optional!!!***

Backup Material

FC-EE Operating State Determination Flow Chart

Variable Definitions:

Local_EE_Lvl = Local side "Energy Efficient Level" Service Parameter setting
 Remote_EE_Lvl = Remote side "Energy Efficient Level" Service Parameter setting
 EE_Lvl = Energy Efficient Level in effect for normal operation
 Local_TTW = Local side Transmitter Wake Time following negotiation
 Remote_TTW = Remote side Transmitter Wake Time following negotiation
 Default_TTW = Default Transmitter Wake Time value from FC-FS-4 Table 24
 TTW = Transmitter Wake Time value in effect for normal operation
 lpi_fw = Variable controlling the wake mode for the LPI Transmit and Receive functions; TRUE denotes "Fast Wake" mode, FALSE denotes "Quiet" mode (see 14-036v0 Slide 7)



***Backup Material from Feb T11 FC-FS-4 Meeting
(14-036v0)***

Fast Wake Proposal

➤ FC-FS-4 Section 10.5 mentions an LPI Mode use case in which the transmitter continually sends LPI...

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.

o However, current LPI Mode Transmitter state diagram does not support this use case...

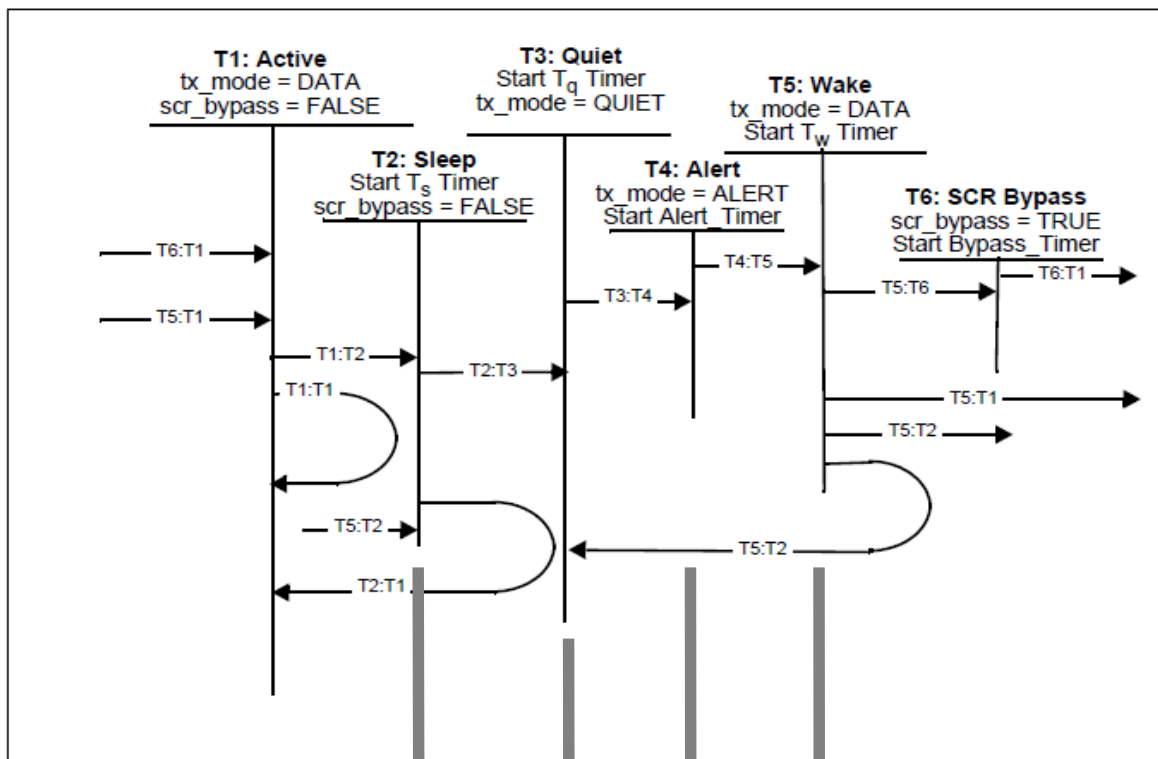


Figure 49 LPI Mode Transmitter State Diagram

Transmitted Frame: LPI LPI ALERT LPI (for Refresh sequence) or IDLE (for Wake sequence)

Fast Wake Proposal (cont)

- FC-FS-4 Section 10.5 mentions an LPI Mode use case in which the transmitter continually sends LPI...
 - ... and current LPI Mode Receiver state diagram does not support this use case:

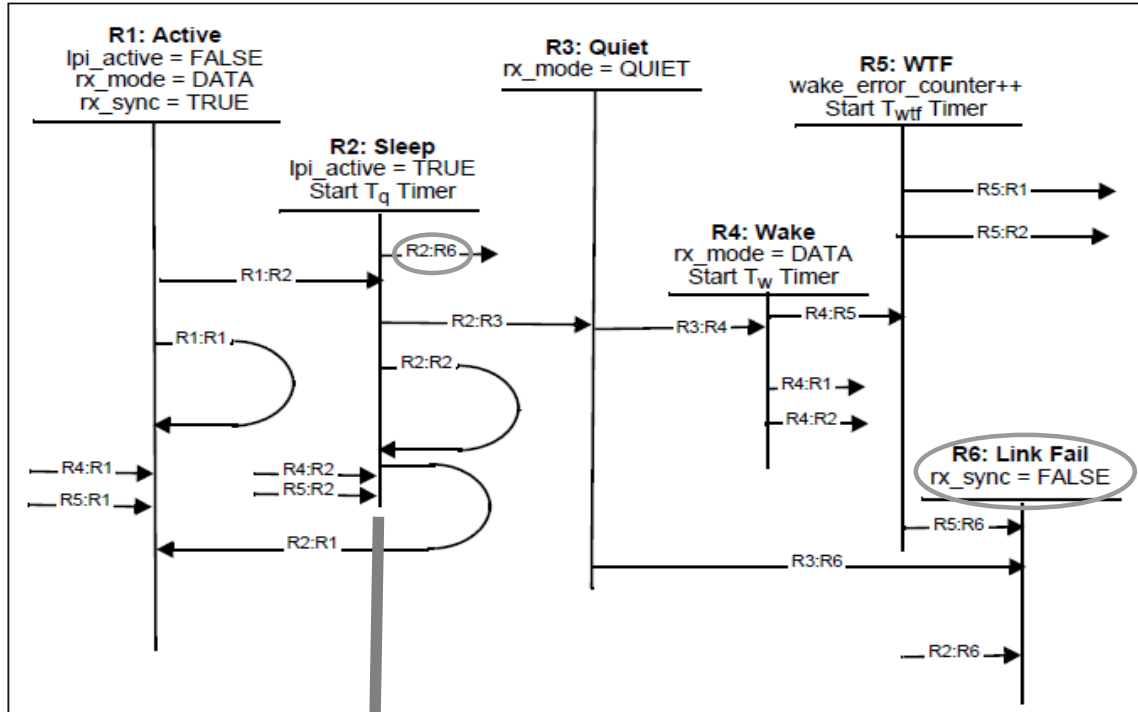


Figure 50 - LPI Mode Receiver State Diagram

Since transmitter continually sends frames, receiver maintains sync (rx_sync=TRUE) throughout R2:SLEEP:

If all frames are LPI, then only R2:R6 applies after Tq expires...

Fast Wake Proposal (cont)

- Proposed updates to FC-FS-4 to support the LPI Mode use case in which the transmitter continually sends LPI:
 - Define Transmitter and Receiver "Fast Wake" states (similar to 802.3bj Figure 82-16 & 82-17)

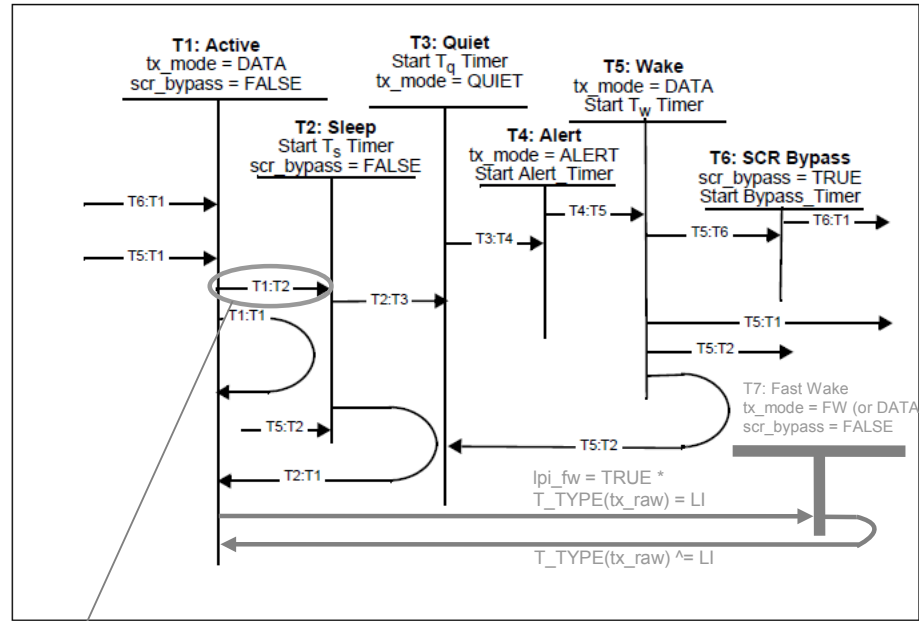


Figure 49 - LPI Mode Transmitter State Diagram

NOTE: Must also check for lpi_fw = FALSE to make T1:T2 transition...

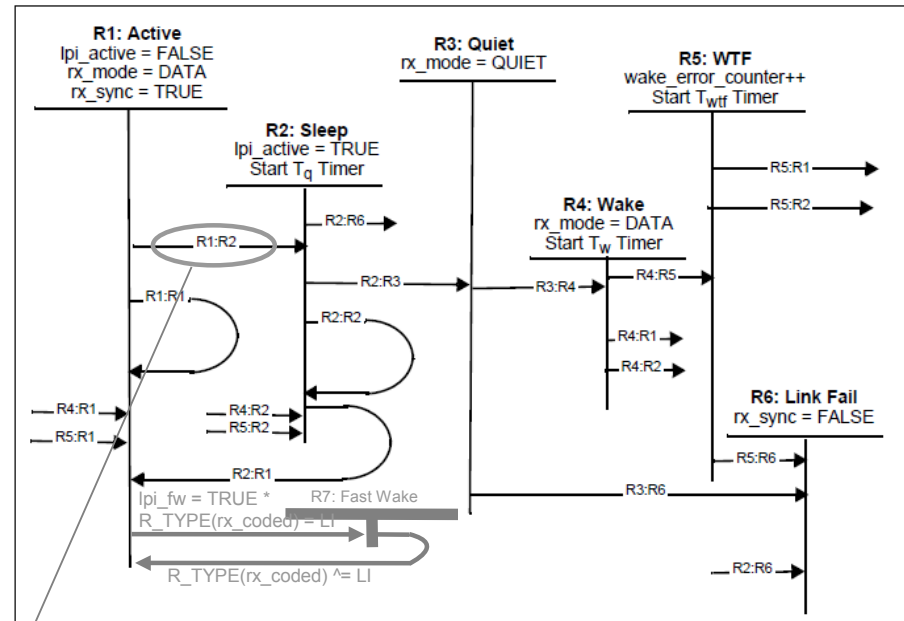


Figure 50 - LPI Mode Receiver State Diagram

NOTE: Must also check for lpi_fw = FALSE to make R1:R2 transition...

Fast Wake Proposal (cont)

- FC-FS-4 provides inadequate detail for FC-EE support @ 128G
- Proposed starting point for specifying FC-EE support @ 128G are LPI Mode state diagrams in 802.3bj Clause 82:

o Note proposed changes to LPI Receive state diagram RX_FW state for 128G FC-EE...

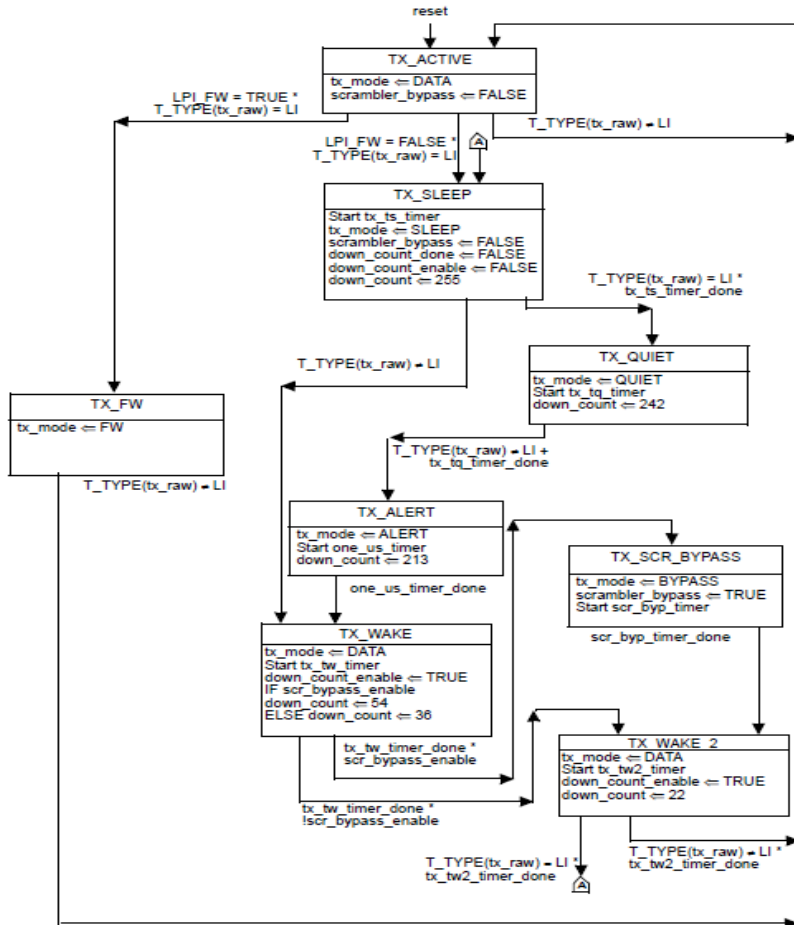


Figure 82-16—LPI Transmit state diagram

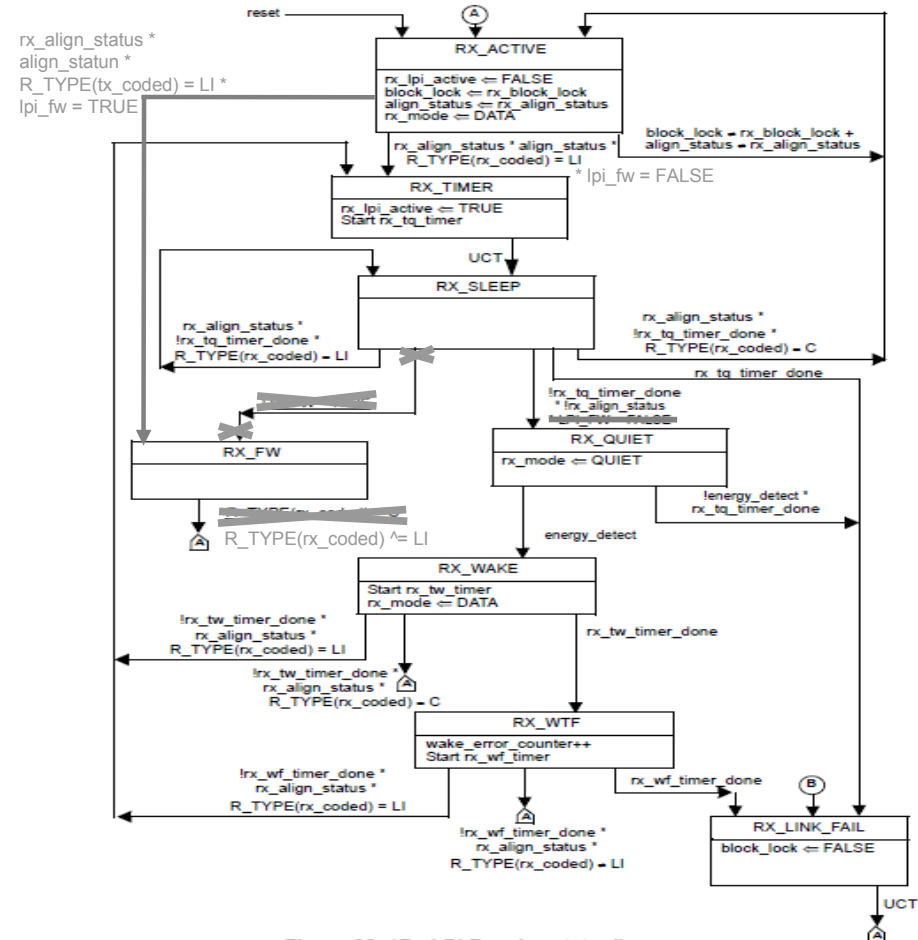


Figure 82-17—LPI Receive state diagram

Energy Efficient Parameter Exchange Descriptor Format

➤ Proposed Energy Efficient Parameter Exchange Descriptor Format (including Fast Wake):

o For FC-LS-3, propose adding Transmit and Receive Fast Wake (FW) to the Exchange Energy Efficient Parameters Extended Link Service (EEEE ELS) Descriptor field after the Transmit and Receive Wake time (Tw) fields:

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 = X 8			
2	Transmit T _w		Receive T _w	
3	Transmit FW	Echo Transmit FW	Receive FW	Echo Receive FW

o For FC-SW-6, propose adding those same parameters to the EEEP Switch Fabric Internal Link Service (SW_ILS) Descriptor field after the Transmit and Receive Wake time (Tw) fields:

Table 11 – EEEP Descriptor

Item	Size Bytes
Tag Value = <TBD>h	4
Length = X 8	4
Transmit T _w	2
Receive T _w	2
Transmit FW	1
Echo Transmit FW	1
Receive FW	1
Echo Receive FW	1