Energy Efficient Ethernet 10GASE-T LPI During Training

Presented by: Mike Grimwood, Broadcom

IEEE P802.3az Task Force

New Orleans, January 2010

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Overview

- This presentation addresses a potential issue caused by allowing transitions to LPI while the PHY is training.

- This presentation is submitted as a detailed suggested remedy to Comment #59 submitted against IEEE P802.3az/D2.2.

Problem: LPI During Training

- 10GBASE-T training can occur when link_status = OK. Since the training sequence is different from the normal signal, the XGMII input is ignored until PCS_Test.
- Once in PCS_Test, there is a problem if the LPI Client signals LPI over the XGMII interface since this can initiate a transition to QUIET before the Link Partner PHY is ready.
- The Transmit state diagrams need a check to prevent this from happening.

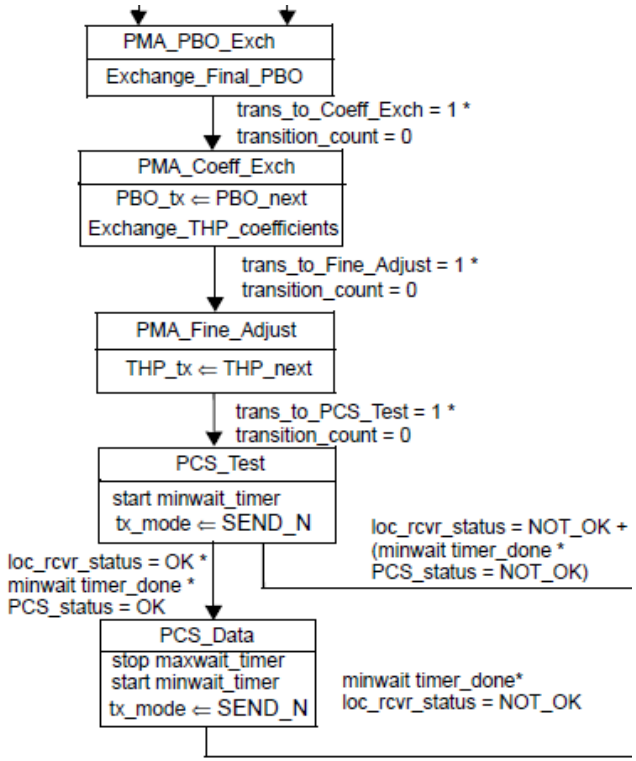


Figure 55–24—PHY Control state diagram

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Change Summary

- Introduce a new variable, *loc_lpi_en*.
- In PHY Control state diagram, set *loc_lpi_en* to FALSE in PCS_Test and set to TRUE in PCS_Data.
- In the Transmit PCS state diagram, inhibit transitions to LPI (TX_L) when *loc_lpi_en* is FALSE.
- Document the communications between the PHY Control block and the Transmit PCS block by updating the functional and reference diagrams and defining the PMA service primitive associated with the variable, *loc_lpi_en*.

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Create a New Variable, loc_lpi_en, and Modify PHY Control State Diagram

- At the end of section 55.4.5.1, introduce a new variable, *loc_lpi_en.*

  The *loc_lpi_en* variable is generated by the PMA PHY Control function and indicates whether or not the local PHY may transition to TX_L. It is passed to the PCS Transmit Control function via the PMA_LOCLPIEN.indication primitive. In the absence of the optional EEE capability, the PHY operates as if the value of this variable is FALSE.

  Values: TRUE or FALSE

- In the Fig 55-24 PHY Control state diagram, *loc_lpi_en* is set to FALSE upon entry into state PCS_Test and is set to TRUE upon entry into PCS_Data.

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Modify PCS 64B/65B Transmit PCS State Diagram

- In the Figure 55-15 PCS 64B/65B Transmit State Diagram, change the following transitions to the conditions noted:
  - TX_C to TX_L: (T_TYPE(tx_raw) = LI) * loc_lpi_en
  - TX_C to TX_C: (T_TYPE(tx_raw) = C + LII) + (T_TYPE(tx_raw) = LI) * !loc_lpi_en
  - TX_T to TX_L: (T_TYPE(tx_raw) = LI) * loc_lpi_en
  - TX_T to TX_C: (T_TYPE(tx_raw) = C + LII) + (T_TYPE(tx_raw) = LI) * !loc_lpi_en
  - TX_E to TX_L: (T_TYPE(tx_raw) = LI) * loc_lpi_en
  - TX_E to TX_C: (T_TYPE(tx_raw) = C + LII) + (T_TYPE(tx_raw) = LI) * !loc_lpi_en

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Add Service Primitive to Communicate loc_lpi_en

- At the end of 55.2.2 add the following PMA Service Primitive:

  PMA_LOCLPIEN.indication(loc_lpi_en)

- Add the following section to the end of 55.2.2:

  55.2.2.9 PMA_LOCLPIEN.indication

  This primitive indicates whether or not the Transmit PCS is able to transition to LPI. The loc_lpi_en variable is generated by the PMA PHY Control function and indicates whether or not the Transmit PCS may transition to TX_L. It is passed to the PCS Transmit Control function via the PMA_LOCLPIEN.indication primitive.

  Values: TRUE or FALSE

  55.2.2.9.1 Semantics of the primitive

  PMA_LOCLPIEN.indication (loc_lpi_en)

  55.2.2.9.2 When generated

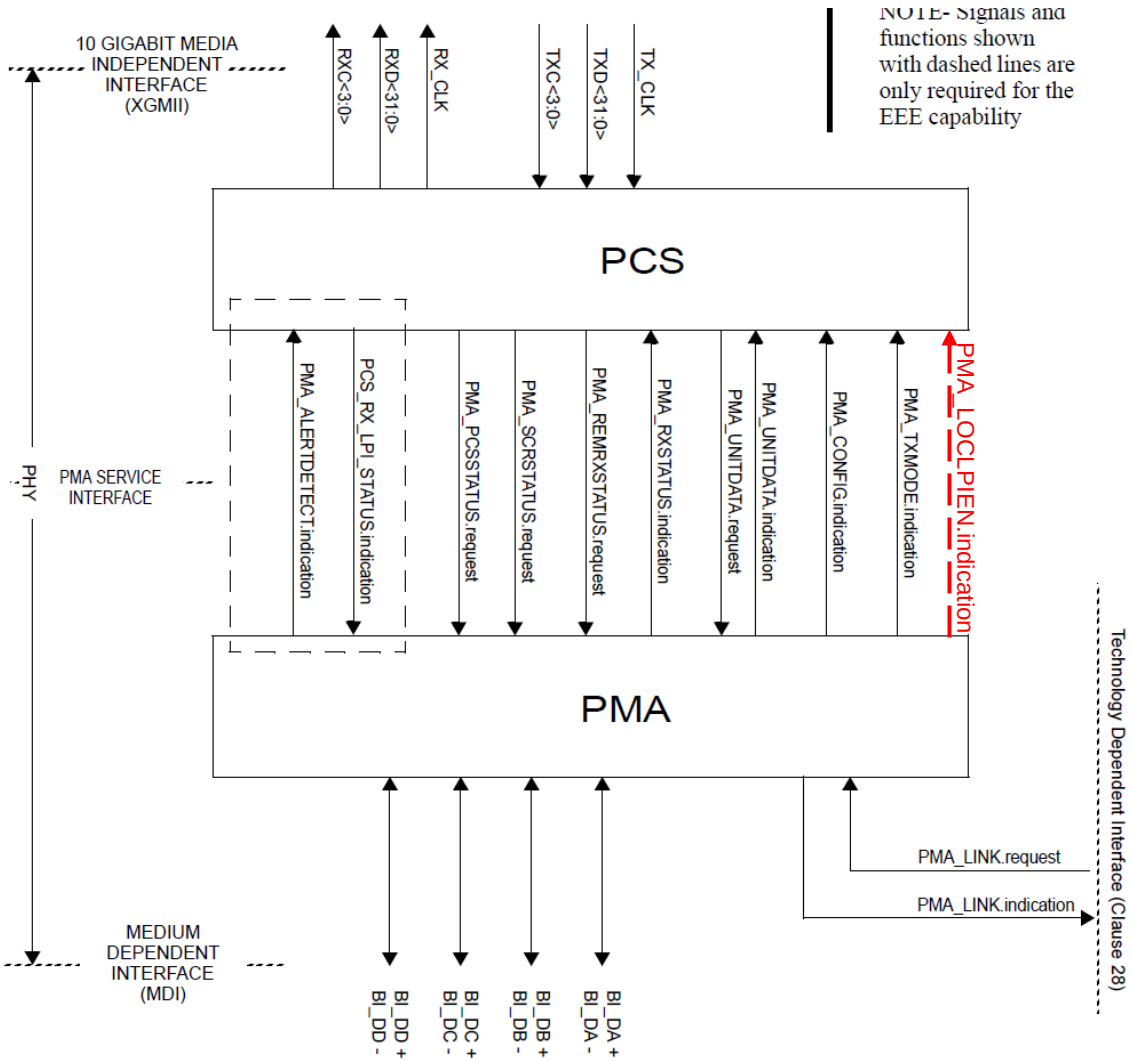
  The PMA PHY Control function generates PMA_LOCLPIEN.indication messages continuously.

  55.2.2.9.3 Effect of receipt

  Upon receipt of this primitive, the PCS performs its transmit function as described in 55.3.2.2.

Energy Efficient Ethernet

Figure 55-4 10GBASE-T Service Interfaces



Energy Efficien Ethernet

BROADCOM. Connecting everything®

Changes to Functional and Reference Diagrams

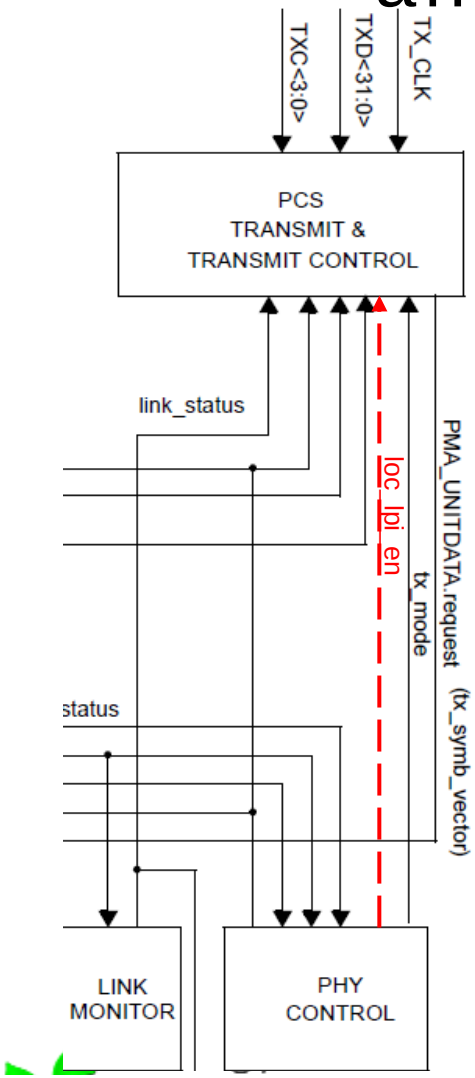


Figure 55-3 Functional Block Diagram

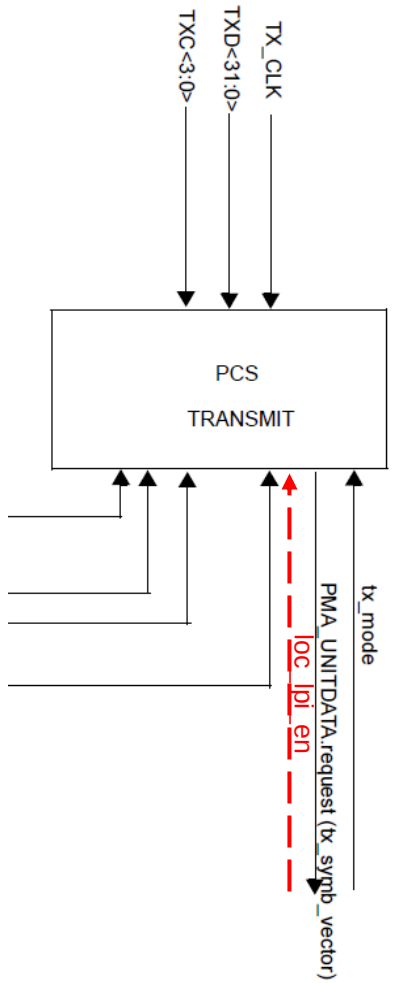


Figure 55-5 PCS Reference

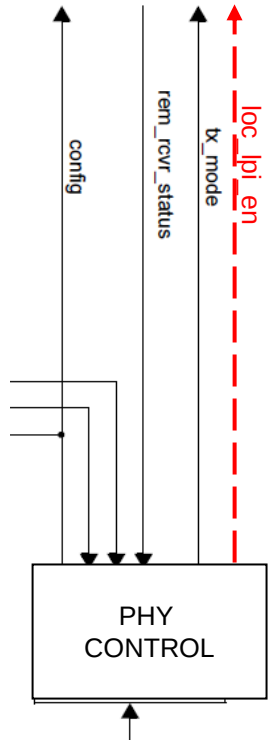


Figure 55-17 PMA Reference

Prevent LPI During Training - LPI Client Considerations

- There are two lpi_wake_time values: 4.48 usec if wake occurs after sleep and 7.36 usec if wake occurs during sleep.
- If the LPI Client signals LPI while the PHY is training, by the proposal in this document, the PHY will not transmit SLEEP until training has completed.
- If the LPI Client signals normal IDLE just after the PHY has begun transmitting SLEEP, a condition could arise in which it would appear that the system can use the 4.48 usec wake time, when it actually needs to use the longer 7.36 usec wake time (since the PHY must finish sending SLEEP).
- This issue can be remedied by noting that the 4.48 usec wake time does not apply if a PHY has indicated Local Fault during the last 10 msec.

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Prevent LPI during Training – Changes to Clause 78

- Comment #60: In Subclause 78.1.2.1.2

  The LPI_REQUEST parameter can take one of two values: ASSERT or DE-ASSERT. ASSERT initiates the signaling of LPI to the link partner. DE-ASSERT stops the signaling of LPI to the link partner. The effect of receipt of this primitive is undefined if link_status is not OK (see 28.2.6.1.1), ~~or~~ if LPI_REQUEST=ASSERT within 1 second of the change of link_status to OK, the PHY is indicating Local Fault, or the PHY is indicating Remote Fault.

- Comment #61: In Subclause 78.5 modify the Case-1 description for the 10GBASE-T wake time to the following:

  Case-1 of the 10GBASE-T PHY applies when the PHY is requested to transmit the Wake signal before transmission of the Sleep signal to the Link Partner is complete and if the PHY has not indicated Local Fault at any time during the previous 10 msec.

Thank you!