Link Failure Detection and Recovery
Comments on parnaby_03_1109.pdf

Mike Grimwood, Broadcom

IEEE P802.3az Task Force

Link Failure Detection and Recovery ad hoc

December 3, 2009

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Overview

- This presentation provides two comments on the proposed link failure detection and recovery scheme presented in parnaby_03_1109.pdf.
  - Issue caused by entering LPI during training.
  - State timer values.

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Problem: LPI During Training
(Applies to both normal and fast link recovery)

- 10GBASE-T training (including fast link recovery) 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 MAC signals LPI over the XGMII interface since this can initiate a transition to QUIET before the Link Partner PHY is ready.

- The TX state machines needs a check to prevent this from happening.

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Solution: Prevent LPI During Training
Create a New Variable

- Introduce a new variable, lpi_tx_inhibit
  - A Boolean set to TRUE to inhibit the PCS 64B/65B Transmit state machine from entering TX_L during training.
  - Values: TRUE or FALSE

- lpi_tx_inhibit is set to TRUE upon entry into state PCS_Test and is set to FALSE upon entry into PCS_Data.

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Solution: Prevent LPI During Training
Modify Transmit PCS State Diagram Changes

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

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Solution: Prevent LPI During Training
MAC 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).
- A note should be added to the specification identifying this potential situation.

Energy Efficient Ethernet

BROADCOM. Connecting everything®

State Timers

- PMA_Coeff_Exch has a proposed maximum time of 20 msec and PMA_Fine_Adjust has a proposed maximum time of 10 msec.
  - Is there a timing budget to justify these values?
  - Should the timers be separated into budgets for both the Master and Slave? The concern is that for whatever values are chosen, the timers may might not be feasible if a link partner consumes a large portion of the budget.
  - Recommend that these values be TBD pending additional analysis.

Energy Efficient Ethernet

BROADCOM. Connecting everything®

Thank you