SOLARFLARE Communications

Link failure detection and recovery in 10GBASE-T EEE

11/17/2009 Gavin Parnaby

Overview

- Not a proposal for a complete solution
- Describes some areas of 10GBASE-T EEE that could be improved by further work
- Some ideas for discussion

Link failure detection

- In the current draft the criteria used to drop the link during LPI is not specified – it is vendor dependent.
- Since PHYs can monitor link quality only during refreshes (and then only for 4 LDPC frames [~1.2us]) and since PHYs may choose not to wake for all refreshes, it can take multiple refresh cycles before the link drop is detected on both PHYs.
  - **Each refresh cycle of 512 LDPC frame periods is 0.16ms**
- In 802.3az draft 2.1 there is no specification for the maximum time period a PHY in rx-LPI is allowed to detect the link partner going completely silent (link drop), and restart the link.

Link recovery

- If the link environment is disturbed during LPI the ability of the PHY to recover is limited by the quiet-refresh signaling since only 4 LDPC frames out of 512 can be used for equalizer/echo training
  - **LPI uses heavily decimated adaptation compared to the normal operational mode**
- If the link fails there is no specified mechanism to notify the link partner, apart from breaking the link and waiting for the link partner to begin full link training (~2s PHY training time + autoneg in some circumstances)
- It would be advantageous to specify a means to recover the link more quickly

Requirements

- **1) Signal to the link partner that the link SNR is critical**
  - simple signaling to initiate link-drop and restart
  - establish synchronization between link partners for link failure recovery
- **2) Use a fast retrain protocol**
  - Skip some initial states to reduce training time and get back to normal operation mode asap
  - Go through the remaining training states more quickly
    - leverage existing receiver state
    - reduce handshaking overhead

Shorten link training

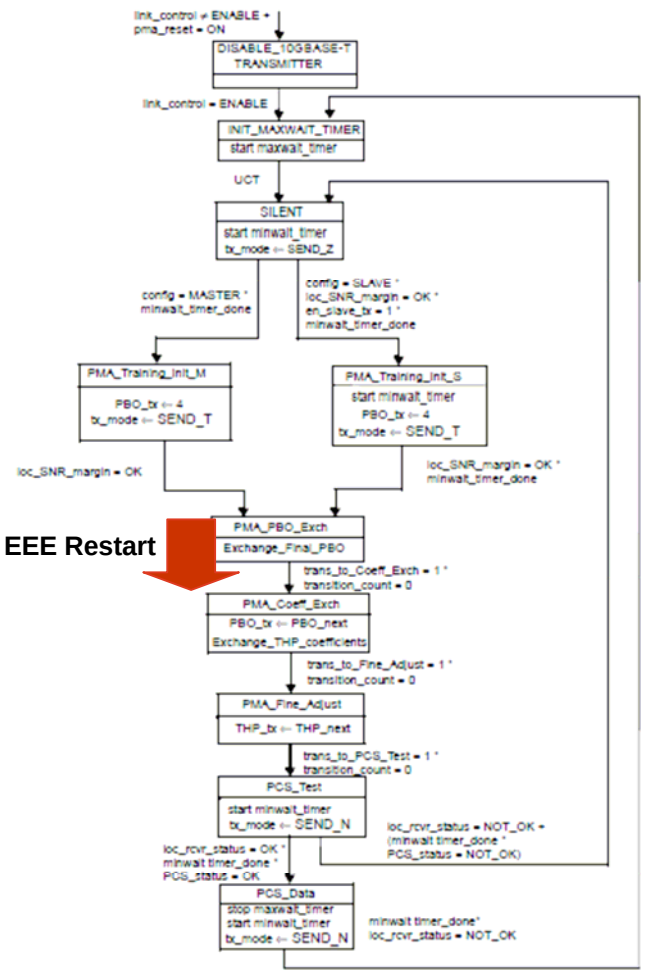

Figure 55–24—PHY Control state diagram

- **Shorten link training by re-entering PHY control state diagram at a later stage**
  - PMA_Coeff_Exch ?
  - Provides an opportunity to re-establish a robust link
- **Specify a faster restart**
  - Reduced training time improves link availability
  - We can reuse existing receiver state
  - Add a specific recommended sequence timing for LPI-restart
    - similar to Table 55-6

SOLARFLARE Communications

Discussion

- Is this something worth pursuing?
  - **If yes**
    - Form an ad hoc to produce results for integration into the draft following the January meeting