100GBASE-KP4 EEE synchronization and signaling

Comments #234 and #235

IEEE P802.3bj

July 2012, San Diego

Matt Brown – AppliedMicro

Adee Ran – Intel

Kent Lusted -- Intel

Introduction

- Addresses comments #234 and #235.
- Proposal for signals for:
  - Normal LPI, refresh and wake states
  - Fast-wake quiet state

Cl 94 *SC* 94.2.5 *P* 150 *L* 29 # 234

Matthew, Brown Applied Micro

Comment Type TR *Comment Status* X

For EEE operation, a signal structure and framing mechanism for allowing the receiver to quickly lock to the PMA frame signal.

SuggestedRemedy

A proposal will be provided at the July meeting.

Proposed Response *Response Status* O

Cl 94 *SC* 94.2.5 *P* 150 *L* 29 # 235

Matthew, Brown Applied Micro

Comment Type TR *Comment Status* X

For EEE operation, a signal structure and framing mechanism for allowing the PMA/PMD to remain operational during the fast wake.

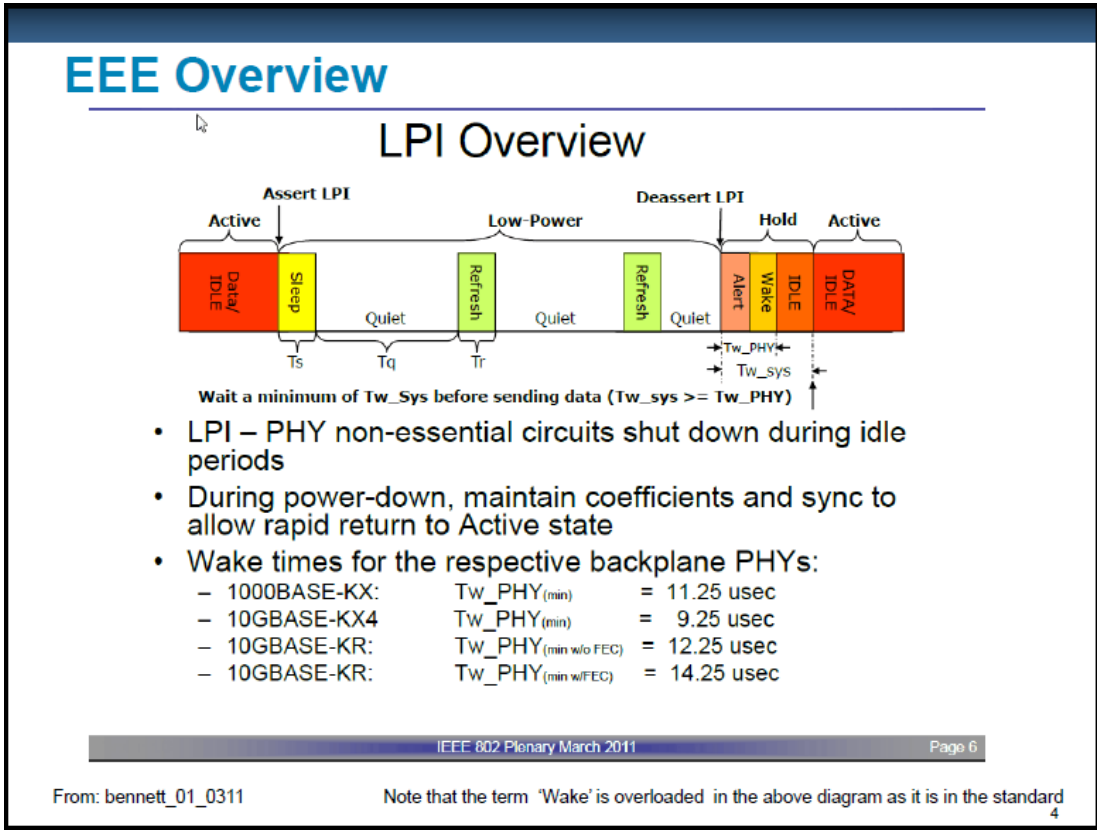
SuggestedRemedy

A proposal will be provided at the July meeting.

Proposed Response *Response Status* O

EEE overview

- EEE normal low power mode transitions to quiet state with occasional refresh states. Wake up and transition to normal data mode is targeted at 5 us.
- In EEE fast-wake low power mode, the PMA and PMD remain active and transmitting a signal so that only digital signals must be re-synchronized. Wake up and transition to data mode is targetted at 500 ns.



From gustlin_02_0112

Normal EEE Refresh and Wake Synchronization

- For EEE, it is necessary to synchronize very quickly to the signal on transitions from QUIET to WAKE or REFRESH.
- The PMA must synchronize in 3-4 us after receiving the Wake or Refresh signal.
  - Orders of magnitude faster than for initial synchronization.
  - The total budget is 5 us, but this must be allocated among transmitter, power up, equalization settling/convergence, etc.
- Once synchronized the FEC sub-layer must also synchronize.

Normal EEE Synchronization challenges

- PAM4 PHY is not able to easily make use of the PCS alignment markers for synchronization.
  - The markers are mixed up by the insertion of overhead, insertion of termination bits, conversion to 4 levels, and precoding.
- The PHY receiver may be relying heavily upon the PMA termination bits for effective data recovery.
  - Before synchronization, the data may not be reliably detected.
- Even with effective equalization, without FEC synchronization the BER will be very high (~1E-5).
  - A soft match method is necessary.
- The wake or refresh signal must be reliably discernible from noise to initiate synchronization for a valid signal.
  - Cannot miss and WAKE/REFRESH signal.
  - Cannot falsely detect a WAKE/REFESH signal.

Fast-Wake EEE synchronization.

- In fast-wake LPI, the PMA remains operational and continues to send a sending from transmitter to receiver.
- Recovery from LPI to active state is expected to occur is less than 500 ns.
- Since the upper layers are powered down, surrogate signal must be sent in place of the PCS and FEC signal.
- During TX_FW state (fast-wake quiet), the PMA will remain locked to the PMA frame since the PMA is always running.
- Since the PMA frame is aligned with a FEC code word, the FEC is immediately locked.
- The alignment marker lock will follow very quickly since it is at the beginning of a code word and the FEC will correct errors.

EEE link state signaling

- Fast-wake LPI mode
  - A means to distinguish FW and WAKE required.
- Normal LPI mode
  - A means to distinguish Refresh and Wake is required.

Training frame

- A concurrent presentation (lusted_03_0712) proposes a training frame largely based upon the 10GBASE-KR training frame.
- The frame includes both a long-symbol-period, two-level frame marker and control channel, as well as line rate training pattern.
- The frame marker and control channel provide a means to quickly lock to the frame even without converged equalization and with high noise.
- The control channel provides 5 bits for EEE signaling to signal the current LPI state, if necessary.
- The relative phase between the training frame and PMA frame is maintained and a countdown field is provided so that after training is complete the receiver is already locked to the PMA frame.
- Since the FEC codeword is aligned with the PMA frame, the FEC may make use of the start of PMA frame to be in lock as well.

Proposal

- For the ALERT and RF_ALERT signals, use the training frame in lusted_03_0712.
  - Status report bit 14 indicates the training frame mode:
    0 = training (the link is in start-up training mode)
    1 = EEE (the link is in LPI mode) ← use this for EEE
  - Status report field bit 15:16 indicates the EEE state (see 802.3bj draft 1.0 80.3.3.4.1):
    0 = Wake, 1 = Refresh.
  - To transition to the PMA frame, use the countdown field to coordinate transition and the offset field to set the alignment.
- For FW state replace the FEC signal with a PRBS31 pattern bit-striped across the four lanes.
  - Bit striping creates 4 PRBS31 signals spaced by 2^29 bits apart.
  - If necessary to distinguish EEE states use the PMA overhead.
  - When FW is done, replace the PRBS31 signal with the PCS/FEC signal.
  - Use the PMA overhead to indicate the state: FW, DATA. Method TBD.

PMA signal WRT PCS LPI state diagram

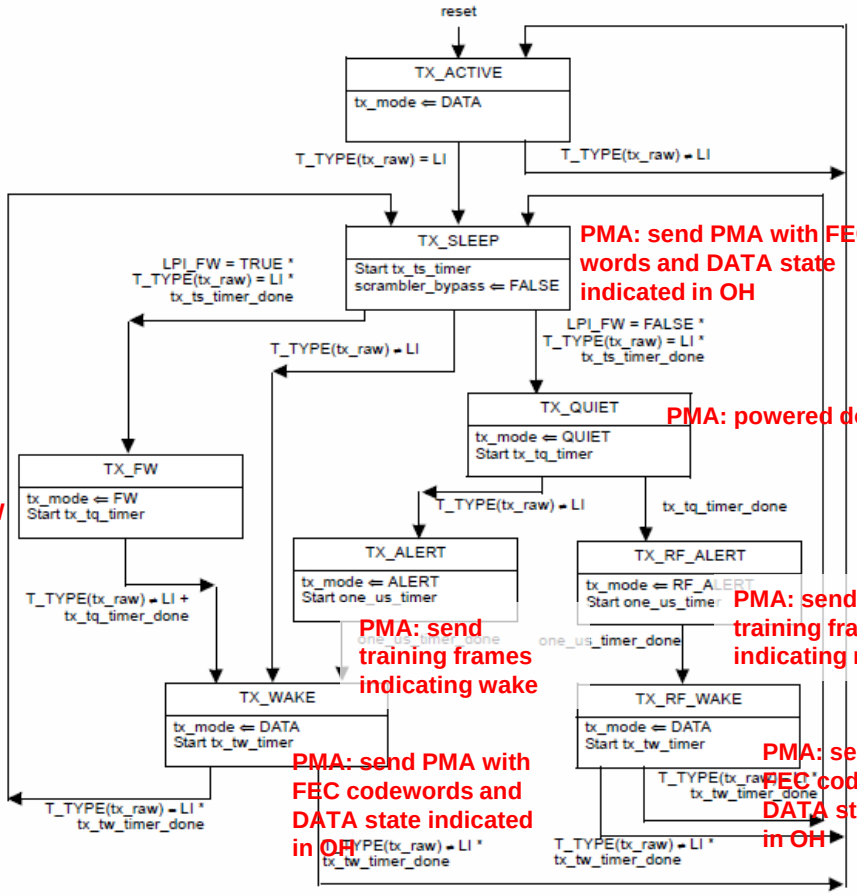


Figure 82–16—LPI Transmit state diagram