Energy Efficient Ethernet (EEE) for 802.3by

Arthur Marris - Cadence

Supporters

- Adee Ran — Intel
- Brad Booth — Microsoft
- Matt Brown — APM

Key features of Energy Efficient Ethernet (EEE)

- If a system has nothing to transmit it can power down its transmit path after it has indicated its intention to do so.
- The partner device can also power down its receive path when it detects that the path is going to be powered down.
- The link stays up while in low power mode and no frames are dropped.
- EEE is asymmetric. One direction can be powered up while the other is powered down.

How does EEE work?

- The client (i.e. system) requests Low Power Idle (LPI) from the reconciliation sublayer (RS).
- The RS then signals LPI on the MII (media independent interface).
- The transmit PCS (physical coding sublayer) encodes the LPI signal using a special symbol.
- The receive PCS decodes the LPI symbol and signals LPI on the receive MII.
- When the client ceases requesting LPI the RS continues inhibiting transmission for a fixed period to allow time for the local transmit path and remote receive path to recover from their low power modes.

What is “Fast Wake” mode of operation?

- In Fast Wake mode of operation the transmit and receive path remain active and continuously transmit and receive LPI when it is requested by the client. (However Clause 82 BIP is not calculated)
- Fast Wake is compulsory for 40G and 100G PHYs that support EEE
- Fast Wake is controlled by LLDP (Link Layer Discovery Protocol) rather than AN (auto-negotiation)
- Fast Wake is suitable for optical PHYs that are unable to power up and down quickly and do not support AN.

What is “Deep Sleep” mode of operation?

- In Deep Sleep mode the transmit path stops transmitting but periodically sends refresh indication while LPI is requested.
- In Deep Sleep mode the receiver checks for the occurrence of refresh indication and will assert link down if refresh does not appear.
- Deep Sleep mode requires the receive PMA to resume operation within a determined time period.
- In Deep Sleep mode the transmit and receive PCS generate tx_quiet and rx_quiet signals to allow the PHY to periodically power down its transmit and receive path.
- Clause 83 defines a mechanism for sending the tx_quiet signal over the CAUI/XLAUI interface and for synthesizing the rx_quiet signal.
- When the Clause 74 BASE-R FEC is present transmit scrambler bypass is used to allow for rapid FEC alignment

What should be done for 802.3by?

- Follow the principle that 802.3bj EEE mechanisms shall be followed for the copper PHYs and 802.3bm for the MMF PHY. Therefore:
  1. Adopt Fast Wake mode of operation for 25GBASE-SR
  2. Adopt Fast Wake and Deep Sleep modes of operation for 25GBASE-KR and 25GBASE-CR

Changes to base standard to support EEE in 802.3by

- Add 25GBASE-KR, 25GBASE-CR and 25GBASE-SR to “Table 78-1 Clauses associated with each PHY or interface type” and indicate that 25GBASE-SR does not support deep sleep with the “b” suffix.
- Add 25GBASE-KR and 25GBASE-CR to “Table 78-2 Summary of the key EEE parameters for supported PHYs or interfaces” with the same time values as 100GBASE-KR4.
- Add 25GBASE-KR and 25GBASE-CR to “Table 78-4 Summary of the LPI timing parameters for supported PHYs or interfaces” with the same time values as 100GBASE-KR4.

Changes to base standard to support EEE in 802.3by

- Change last paragraph of 78.1.3.3.1 to read “Deep sleep support is optional for PHYs with an operating speed of 25~~40~~Gb/s or greater that implement EEE with the exception of the PHYs noted in Table 78–1 which do not support deep sleep. *Fast wake* refers to the mode for which the transmitter continues to transmit signals during Low Power Idle so that the receiver can resume operation with a shorter wake time (as shown in Figure 78–4). For transmit, other than the PCS encoding LPI, there is no difference between fast wake and normal operation. Fast wake support is mandatory for PHYs with an operating speed of 25~~40~~Gb/s or greater that implement EEE.

Additions to 802.3by to support EEE

- The PCS needs to include an LPI receive state diagram to detect link down due to lack of refresh and set rx_mode to QUIET or DATA
- The PCS needs to include an LPI transmit state diagram to set tx_mode to QUIET, ALERT or DATA
- The PCS needs to support encoding and decoding of LPI
- The transmit PCS needs to support scrambler bypass for when the Clause 74 BASE-R FEC is used.
- PMA/PMD clause needs to support energy detect and transmit quiet and alert signalling.
- The RS-FEC needs to support rapid alignment marker (RAM) generation. Tx_mode from the PCS signals DATA, QUIET or ALERT and provides enough information to control RAM generation.

Summary

- 802.3by can support EEE in a similar way to 802.3bj for the copper PHYs and 802.3bm for the MMF PHY
- Following this principle it is unambiguous what you need to do to support EEE in 802.3by