Energy Efficient Ethernet (EEE) for 802.3bs

Arthur Marris - Cadence

Supporters

- Mark Gustlin – Xilinx
- Andre Szczepanek – Inphi
- Dave Ofelt – Juniper
- Gary Nicholl - Cisco

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 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 failure if refresh does not appear.
- Deep Sleep mode requires the receive PMA and PCS 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.
- As only optical PHYs are included in the 802.3bs objectives, none of the PHYs specified in 802.3bs will support Deep Sleep operation, however the architecture should not preclude support for Deep Sleep in the future.

What should be done for 802.3bs?

- Adopt Fast Wake mode of operation for the 802.3bs PHY types.
- Add these PHY types to “Table 78-1 Clauses associated with each PHY or interface type” and indicate that they do not support deep sleep with the “b” suffix.
- The CDMII will need to be able to signal LPI and the RS will need to include a transmit LPI state machine to defer transmission for the wake time period after de-assertion of LPI.
- The PCS will need to be able to encode and decode LPI.

Further considerations for supporting EEE in 802.3bs

- Consider whether the receive PCS should have a “RX_FW” state that it goes into when receiving LPI from its link partner. In this state it could disable error checking.
- When specifying the CDAUI electrical interface consider defining a mechanism for transmitting alert and quiet signalling similar to how it is done in 83.5.11.1 to allow for future support of Deep Sleep, and also consider generation of an energy detect signal.
- To support Deep Sleep mode the PCS needs to achieve synchronization and alignment quickly. This will most likely be done through the use of “rapid alignment markers” so make sure the architecture does not preclude their use in the future.
- If Deep Sleep is supported in the future then the CDXI may need to be powered down. In this case it would be necessary to preserve the MII signalling of LPI to the PCS in the transmit direction and LPI to the MAC in the receive direction until the interface powers up. Also it will be necessary for the CDXI to power up quickly.

Summary

- 802.3bs will use EEE fast wake functionality which requires the CDMII to signal LPI and the PCS to encode and decode it
- To allow for future support of deep sleep functionality consideration needs to be given to how the PCS and electrical interfaces can resume operation quickly after power down