

EEE in 40/100G PMA

Hugh Barrass
IEEE P802.3bj
September, 2012

With thanks...

With thanks to -

all those who participated in
conference call reviews

Agenda

- ▶ **Options**
- ▶ Clause 74 FEC compatibility
- ▶ PMA/PMD function
- ▶ Changes required
- ▶ Questions...

D-1.0 – Comment #128

- ▶ A proposal to make the two EEE modes optional was rejected because there was insufficient argument in favor...
 - However, during discussions, 2 significant reasons later emerged
- ▶ A separated PMA/PMD might consume more energy inferring the EEE state than it saves during quiescent phase.
- ▶ The inclusion of .3ba PHYs opens the issue of backward compatibility with modules or devices that will not support quiescent state.

Optional behavior

- ▶ EEE is an optional function:
- ▶ Fast Wake should be the default behavior for 40G and above (also works for optics!).
- ▶ Line quiescence should be an option for all PHYs that can support it.
- ▶ Negotiation required for the option.
 - Handle using LLDP only – since dynamic changes needed (per baseline).

Comments #96, 66, 67, 97, 105, 110, 111, 119, 120, 122, 123, 124, 125, 126, 127, 128

Agenda

- ▶ Options
- ▶ **Clause 74 FEC compatibility**
- ▶ PMA/PMD function
- ▶ Changes required
- ▶ Questions...

Clause 74 FEC with 10GBASE-KR

- ▶ Only intended for integrated PHY implementation
 - Signals from PCS – tx_mode, rx_mode, LPI_active
 - Signal from PMA/PMD – energy_detect (not just signal_detect)

- ▶ FEC relies on scrambler_bypass for rapid lock
 - Pre-determined data contents for 1uS during wake
 - 74.7.4.8 FEC rapid block synchronization for EEE (optional)

Scrambler bypass?

- ▶ 40/100G should not need scrambler bypass
 - RAMs are not scrambled – may be used for rapid sync by FEC
- ▶ Some change to Clause 74 will be required
 - 74.7.4.4 – Force start of FEC frame on RAM – when tx_mode changes from QUIET – start frame RAM down_count divisible by 4
 - Wording of 74.7.4.8 needs to accommodate RAM behavior for rapid sync

Comments #68, 69

Logical service interfaces

- ▶ 40/100G service interfaces need to match 10G
 - Although functional behavior of 40/100G may be better...
 - ... because of RAMs
- ▶ Proposals for interface definitions depend on PMA/PMD
 - Look at energy detect & rx_mode

Agenda

- ▶ Options
- ▶ Clause 74 FEC compatability
- ▶ **PMA/PMD function**
- ▶ Changes required
- ▶ Questions...

Transmit direction

- ▶ Functionally, 40/100G very similar to .3az 10G
 - PMA/PMD controlled by tx_mode
 - Only QUIET, ALERT & DATA have any real effect (all others act the same as DATA for PMA/PMD)
- ▶ Use a note for Clause 74 FEC compatibility
 - (if FEC sublayer present, only convey the 3 states)

Comment #95

Receive direction

- ▶ Problems still unresolved in 802.3bj
 - “clean” solution (with detached PMA) not acceptable
 - 802.3az solution might be problematic with CAUI
- ▶ Compatibility with CI 74 FEC requires az-like signals
 - i.e. rx_mode/energy_detect – not compatible with CAUI
- ▶ Solution can define integrated PHY operation...
 - ... allow inference for detached PMA
 - Only required for optional quiescent line mode
 - “complex” detached parts (e.g. FEC/PMA/PMD) uses inference
 - “simple” detached parts (e.g. PMA/PMD) uses fast-mode only

Therefore!

- ▶ From the PCS “down” the PHY – rx_mode: QUIET or DATA
 - This allows PMA/PMD to know when to expect quiescence...
 - ... and therefore when to “arm” energy detect circuit
 - Also used to decide when to go into low power mode
- ▶ From PMA/PMD “up” the PHY – energy_detect: FAIL or OK
 - Set to FAIL when rx_mode changes to QUIET, set to OK when ALERT received.
- ▶ Reuse these definitions – PCS/PMA/FEC sublayers
 - No definition for how these cross CAUI
 - PCS can still use received_tx_mode (more flexible than .3az)

Agenda

- ▶ Options
- ▶ Clause 74 FEC compatability
- ▶ PMA/PMD function
- ▶ **Changes required**
- ▶ Questions...

Definitions

- ▶ Change direction of rx_mode (all instances)

- Only 2 states – DATA & QUIET
- Add energy_detect (from PMA to PCS – as for .3az)
 - (PMA energy_detect = PMD signal_detect)
- Add rx_lpi_active (for FEC only)

Comments #61, 62, 63, 64, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 88, 89, 91, 92, 99, 100, 101, 102, 103, 104, 129

- ▶ 85.7.4 Global PMD energy detect

- Use definition from Clause 72
- Also works for Clause 84

Comment #94

- ▶ Note used for tx_mode from PCS to allow Clause 74

- States DATA, SLEEP, FW, WAKE and RF_WAKE all map to DATA
- States ALERT and RF_ALERT both map to ALERT
- (LP gets fully functional received_tx_mode from RAMs)

Rx LPI state machine

- ▶ Assign rx_mode = DATA in states RX_ACTIVE & RX_WAKE
 - Assign rx_mode = QUIET in state RX_QUIET
- ▶ State transitions:
 - RX_SLEEP > RX_SLEEP; RX_SLEEP > RX_ACTIVE – replace rx_mode = DATA with rx_align_status
 - RX_SLEEP > RX_QUIET – replace rx_mode = QUIET with !rx_align_status
 - RX_QUIET > RX_LINK_FAIL – replace rx_mode = QUIET with !energy_detect
 - RX_QUIET > RX_WAKE – replace rx_mode != QUIET with energy_detect
 - RX_WAKE > RX_TIMER; RX_WAKE > RX_ACTIVE – replace rx_mode = DATA with rx_align_status
 - RX_WTF > RX_TIMER; RX_WTF > RX_ACTIVE – replace rx_mode = DATA with rx_align_status

Comments #84, 85, 86, 87

Agenda

- ▶ Options
- ▶ Clause 74 FEC compatability
- ▶ PMA/PMD function
- ▶ Changes required
- ▶ Questions...

Questions...

???



CISCO