



Energy Efficient Ethernet (EEE) proposal for GEPOF

Rubén Pérez-Aranda
(rubenpda@kdpof.com)
David Ortiz

Agenda



- Background and objectives
- 802.3 EEE introduction
- EEE proposal for GEPOF

Background & Objectives



- In 802.3bv TF interim meeting of January 2015, the FEC and modulation schemes proposed in [1] were adopted for the baseline together with the transmission scheme defined in [2]
- Based on the transmission scheme of [2], a proposal to implement Low Power Idle (LPI) mode is presented here



802.3 EEE introduction

802.3 EEE introduction (based on Clause 78)



- The optional EEE capability combines the MAC sublayer with a PHY defined to support the operation in the Low Power Idle (LPI) mode. When the LPI mode is enabled, systems on both sides of the link can save power during periods of low link utilization.
- EEE protocol is defined to coordinate transitions to or from low power and does this without changing the link status and without dropping or corrupting Ethernet frames.
- LPI signaling allows to LPI client to indicate to the PHY, and to the link partner, that a break in the data stream is expected, and the LPI client can use it to enter power-saving modes that require additional time to resume normal operation.
- PHY LPI TX operation:
 - The LPI client requests the PHY to transition to its low power state by encoding “Assert LPI” on the GMII, then the PHY signals sleep to its link partner to indicate that the local transmitter is entering LPI mode
 - After local PHY signals sleep during a time, the local PHY transmitter goes quiet
 - The transmit function of the local PHY is enabled periodically to transmit refresh signals that are used by the link partner to maintain link integrity (keep aligned timing, equalizer circuits)
 - The quiet-refresh cycle continues until the reception of normal interframe encoding on the GMII, which is communicated by the local PHY to its link partner with wake signal for a predefined period of time. Then the PHY enters in normal operation mode.

802.3 EEE introduction



- PHY LPI RX operation:
 - Entering the LPI mode is triggered by the reception of a sleep signal from the link partner. After sending the sleep signal, the link partner will cease transmission.
 - When the receiver detect sleep signal, the local PHY indicates “Assert LPI” in the GMII and the local receiver can disable some functionality to reduce power consumption.
 - Local PHY will use the refresh signals periodically sent by the link partner to update equalization and timing.
 - When wake signal sent by the link partner is detected by the receiver, the local PHY prepares for normal operation and transition from “Assert LPI” encoding to normal interframe encoding on the GMII.
 - After specified amount of time for recovery, the link supports the nominal operational data rate
- The EEE proposal for GEPOF presented here allows independent LPI mode operation of the two communication directions, being possible to independently save power as function of the link utilization of each direction

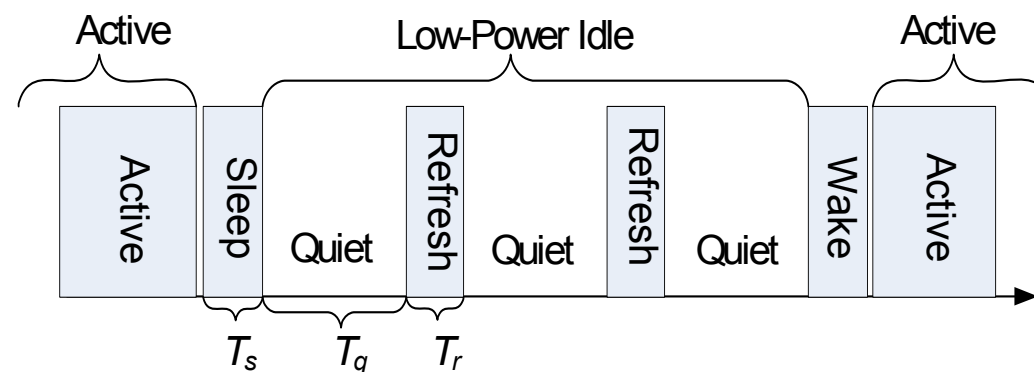


Figure 78–3—Overview of EEE LPI operation

802.3 EEE introduction - GMII encoding

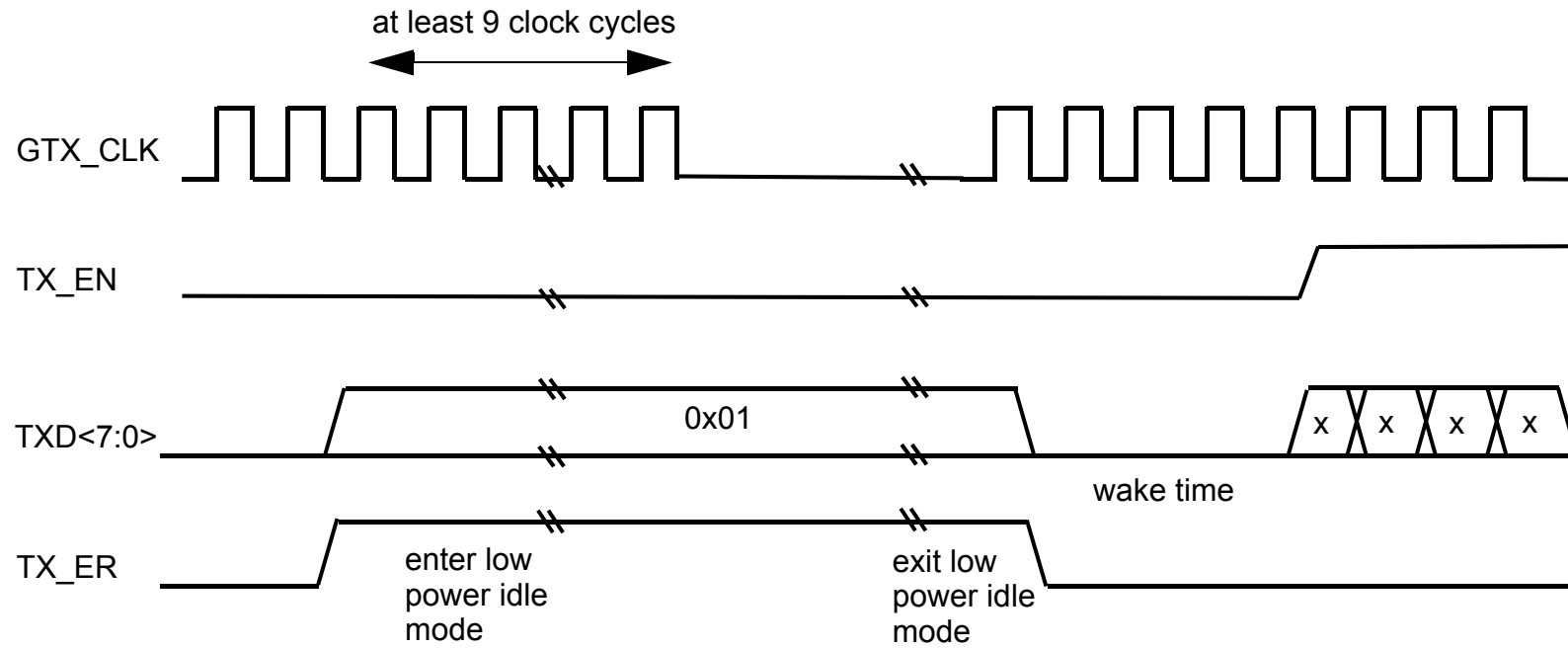


Figure 35-8—LPI transition

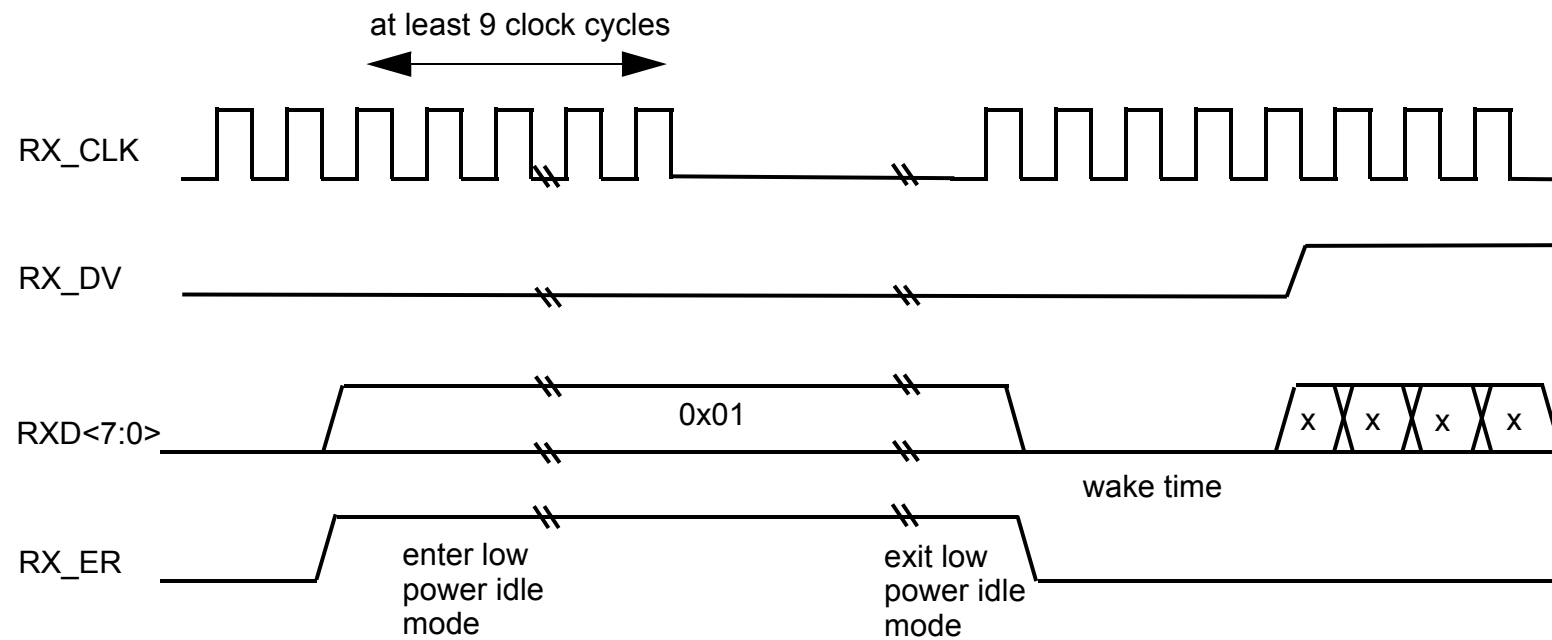


Figure 35-14—LPI transitions (receive)

802.3 EEE introduction - timing

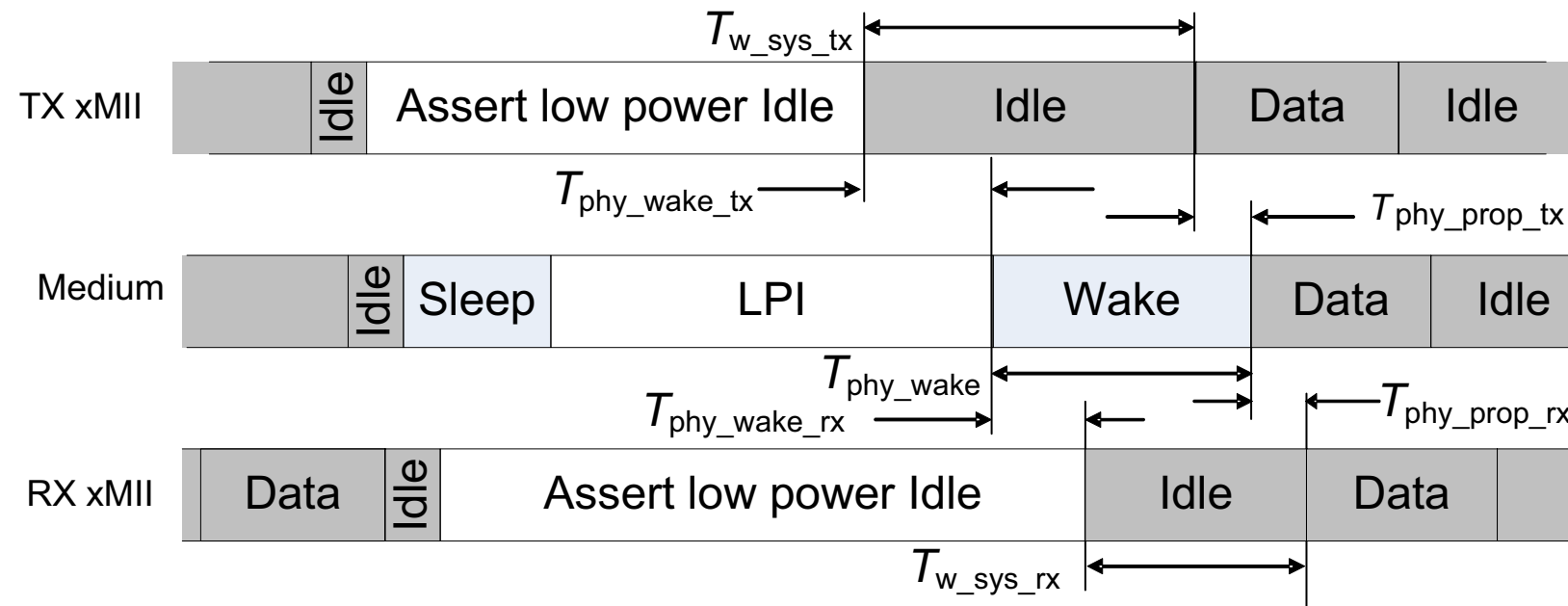


- For a technically complete EEE proposal for GEPOF, we need to define a set of timing parameters to be included within the Clause 78 (tables 78-2 and 78-4)

78.2 LPI mode timing parameters description

- T_s The period of time that the PHY transmits the sleep signal before turning all transmitters off
- T_q The period of time that the PHY remains quiet before sending the refresh signal
- T_r Duration of the refresh signal
- $T_{\text{phy_prop_tx}}$ The propagation delay of a given unit of data from the xMII to the MDI
- $T_{\text{phy_prop_rx}}$ The propagation delay of a given unit of data from the MDI to the xMII
- $T_{\text{phy_shrink_tx}}$ Transmitter shrinkage time, defined as the absolute time difference between the following two timing parameters:
- Delay between a transition from the “Assert LPI” to “Normal Idle” at the xMII and the corresponding start of the wake signal at the MDI
 - $T_{\text{phy_prop_tx}}$
- $T_{\text{phy_shrink_rx}}$ Receiver shrinkage time, defined as the absolute time difference between the following two timing parameters:
- Delay between start of the wake signal at the MDI and the corresponding transition from “Assert LPI” to “Normal Idle” at the xMII
 - $T_{\text{phy_prop_rx}}$
- T_{w_phy} Parameter employed by the system that corresponds to the behavior of the PHY. It is the period of time between reception of an IDLE signal on the xMII and when the first data code-words are permitted on the xMII. The wake time of a compliant PHY does not exceed T_{w_phy} (min).
- $T_{w_sys_tx}$ Parameter employed by the system that corresponds to its requirements. It is the longest period of time the system has to wait between a request to transmit and its readiness to transmit.
- $T_{w_sys_rx}$ Parameter employed by the system that corresponds to its requirements. It is the minimum time required by the system between a request to wake and its readiness to receive data.

802.3 EEE introduction - timing



$$T_{w_sys_tx}(\min) = T_{w_sys_rx}(\min) + T_{phy_shrink_tx}(\max) + T_{phy_shrink_rx}(\max)$$

$$T_{w_phy}(\min) = T_{phy_wake}(\min) + T_{phy_shrink_tx}$$

$T_{w_sys_res}(\min)$ is greater of $T_{w_sys_tx}(\min)$ and $T_{w_phy}(\min)$

$$T_{phy_shrink_tx}(\max) = (T_{phy_wake_tx}(\max) - T_{phy_prop_tx}(\min))$$

$$T_{phy_shrink_rx}(\max) = (T_{phy_wake_rx}(\max) - T_{phy_prop_rx}(\min))$$

where

- $T_{phy_wake_tx}$ = xMII start of wake to MDI start of wake delay
- $T_{phy_prop_tx}$ = xMII to MDI data propagation delay
- $T_{phy_wake_rx}$ = MDI start of wake to xMII start of wake delay
- $T_{phy_prop_rx}$ = MDI to xMII data propagation delay
- T_{phy_wake} = Minimum wake duration required by PHY

Figure 78–4—LPI mode timing parameters and their relationship to minimum system wake time



EEE proposal for GEPOF

LPI in GEPOF - introduction

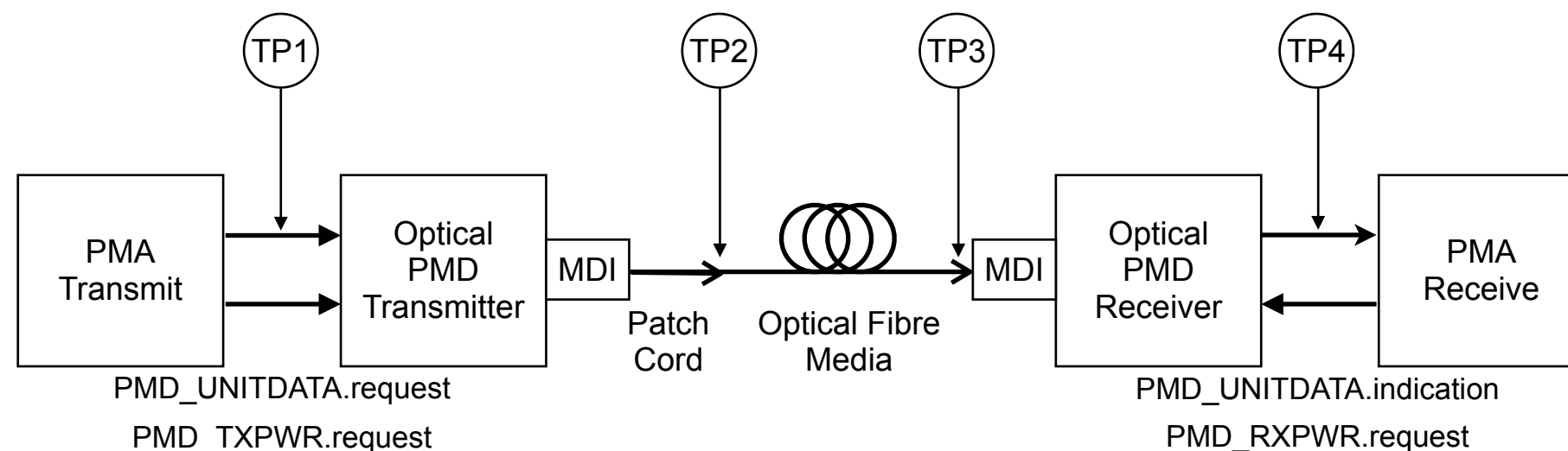


- Physical header PHD is used by both link partners to agree the capability to accept and generate transmission blocks implementing LPI mode during the startup (PHD.CAP.LPI [3])
- Basic idea behind LPI proposal for GEPOF:
 - The PHY partially turns off the PCS, PMA and PMD TX functionalities during the payload sub-blocks when LPI is asserted from LPI client, so that no optical power is injected into the fiber during these periods of time
 - On the other side, the receiver of the remote PHY, which has established the link and is aligned in timing with the local PHY, detects after every pilot and PHS sub-block, the state of its partner based on the amount of received optical power during the payload sub-blocks
 - Because both PHYs are synchronized, the receiver knows the timing for each element of the received signal, therefore is able to enable the PCS, PMA and PMD circuitry just for the reception of the pilot and PHS sub-blocks
- In both, normal and LPI operation modes, all the pilot and PHS sub-blocks are transmitted, although in LPI mode the transmission is switched off during the data sub-blocks
- The receiver is able to use pilots and PHS to keep aligned equalizers and timing
- The LPI mode always affects complete payload data sub-blocks, therefore it is not possible to stop or restart the transmission in the middle of a payload sub-block

LPI in GEPOF - PMD service interface



- PMD service interface needs to be defined since special control signaling is required to implement LPI.
- `PMD_UNITDATA.request(tx_signal)`: this primitive conveys to the PMD via the parameter `tx_signal` the value of the analog electrical signal to be converted by the PMD transmit function into optical power signal to be sent over the optical fiber.
- `PMD_UNITDATA.indication(rx_signal)`: this primitive conveys to the PMA via the parameter `rx_signal` the value of received electrical signal converted by the PMD receive function from the optical power signal received through the MDI.
- `PMD_TXPWR.request`: this primitive is generated by the PMA transmit function to signal to the PMD transmit function to turn off or on the optical power injection into the fiber, when LPI is used.
- `PMD_RXPWR.request`: this primitive is generated by the PMA receive function to signal to the PMD receive function to turn of or on the optical to electrical conversion, when LPI is used. By this indication the PMD receive function is able to:
 - Save the internal state of the AGC and DC restoration circuitry during the lack of optical power.
 - Resume the normal operating mode in a very short period of time.
- The timing of `PMD_TXPWR.request` and `PMD_RXPWR.request` primitives is implementation dependent



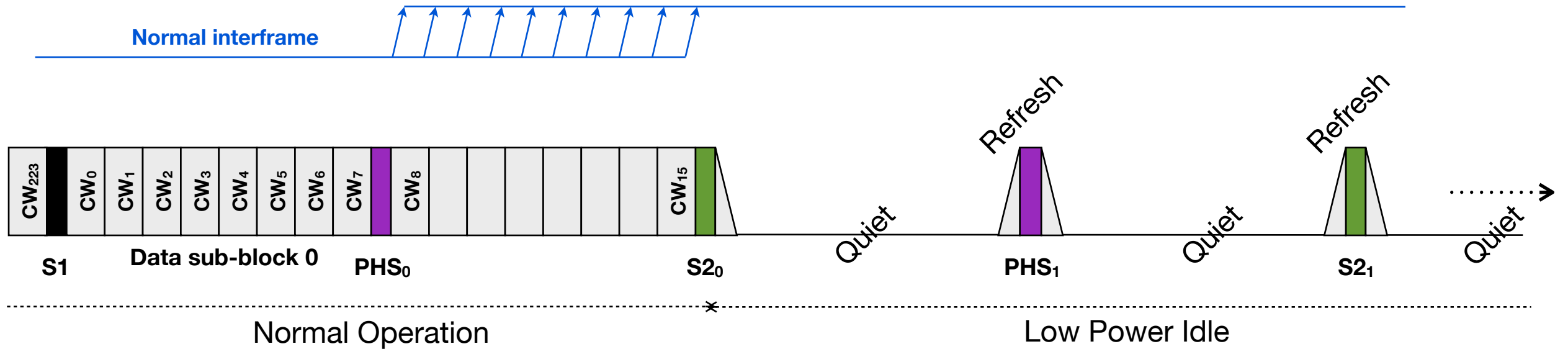
LPI in GEPOF - operation



EEE signaling:

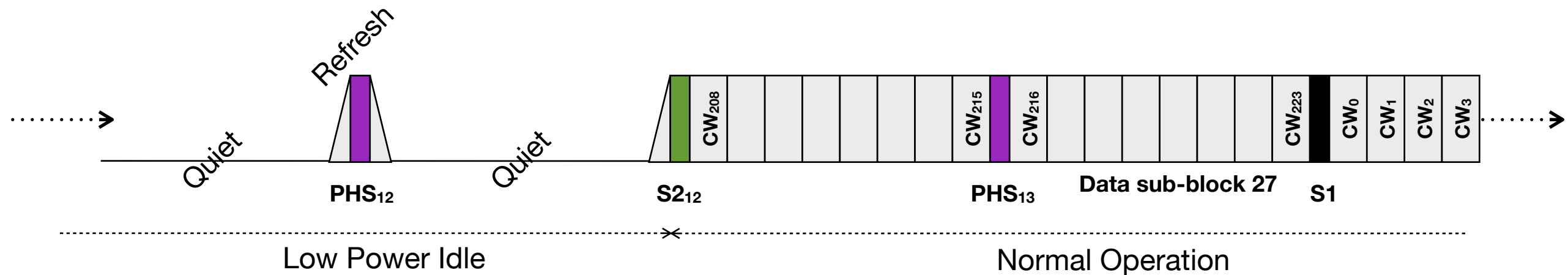
Assert LPI

Normal interframe



Assert LPI

Normal interframe and Ethernet packets



LPI in GEPOF - operation



- PHY LPI TX operation:

- The LPI client requests the PHY transmit function to transition to low power state by encoding “Assert LPI” on the GMII
- The PHY TX function encodes the “Assert LPI” signal on the PDB blocks as defined in [5]
- Once at least one “Assert LPI” encoded on a PDB has been transmitted to the link partner, the transmit function of the local PHY shall be able to go quiet state aligned to the beginning of the next data payload sub-block if no transition to normal inter-frame is signaled before the end of the current data payload sub-block
- Then, the transmit function of the local PHY is enabled periodically to transmit pilot and PHS sub-blocks (refresh signals), which are transmitted just in the same instants of time corresponding to normal operating mode
- The quiet-refresh cycle continues until the reception of normal interframe encoding on the GMII
- When normal interframe is signaled, the transmit function of the local PHY starts the normal operation aligned to the beginning of the next payload sub-block, therefore keeping the transmission enabled after the corresponding pilot or PHS sub-block
- It is important to note that neither sleep nor wake special signals exist during predefined periods of time in GEPOF; on the other hand, sleep and wake are events that are signaled to the link partner by the fact of existing, or not, of transmit signal in the established periods of time allocated to payload sub-blocks

LPI in GEPOF - operation



- PHY LPI RX operation:
 - Entering the LPI mode is triggered by the detection of no optical power received in the data payload sub-block after a pilot or PHS sub-block reception
 - When the receiver detects this event, the local PHY indicates “Assert LPI” on the GMII and the local receiver disable some functionality to reduce power consumption: partial PMD, PMA and PCS Rx functionalities
 - “Assert LPI” on the GMII can also be encoded due to the reception of PDBs containing LPI signaling from the link partner; this is the case of LPI assertion on the GMII Tx in the middle of a payload sub-block transmission
 - Local PHY will use the pilots and PHS sub-blocks (refresh signals) periodically sent by the link partner to update equalization and timing
 - When optical power received in the data payload sub-block after a pilot or PHS reception is detected, the local PHY starts normal operation and sends to the GMII the data contained on the PDBs received from the remote PHY, starting the PCS decoding aligned to the boundary of the first complete PDB received since the beginning of the payload sub-block
 - At this moment, the local PHY shall start to send normal interframe encoding on the GMII, since this is the information received from the remote PHY and the link is ready to provide nominal data rate

LPI in GEPOF - timing parameters



- Tables 78-2 and 78-4 timing parameters are based on:

- $T_{\text{pilot/PHS}} = (16 + 128 + 16)/325 = 0.4923 \mu\text{s}$
- $T_{\text{payload}} = (8 * 988)/325 = 24.32 \mu\text{s}$
- $T_{\text{pilot/PHS}} + T_{\text{payload}} = 24.8123 \mu\text{s}$

Table 78-2 – Summary of the key EEE parameters for supported PHYs or interfaces

PHY or interface type	T_s (μs)		T_q (μs)		T_r (μs)	
	Min	Max	Min	Max	Min	Max
1000BASE-RH	0	0	24.32	24.32	0.49	0.49

Table 78-4 – Summary of the LPI timing parameters for supported PHYs or interfaces

PHY or interface type	Case	$T_{w_sys_tx}$ (min) (μs)	T_{w_phy} (min) (μs)	$T_{phy_shrink_tx}$ (max) (μs)	$T_{phy_shrink_rx}$ (max) (μs)	$T_{w_sys_rx}$ (min) (μs)
1000BASE-RH		24.82	24.82	24.82	0	0

References



- [1] *Rubén Pérez-Aranda, et al., “High spectrally efficient coded 16-PAM scheme for GEPOF based on MLCC and BCH”, IEEE 802.3bv TF, Interim Meeting, January 2014*
- [2] *Rubén Pérez-Aranda, “Transmission scheme for GEPOF”, IEEE 802.3bv TF, Interim Meeting, January 2014*
- [3] *Rubén Pérez-Aranda, et al., “Physical Header Data content for PCS encoding, PHY control and OAM implementation in GEPOF”, IEEE 802.3bv TF, Plenary Meeting, March 2014*
- [4] *Rubén Pérez-Aranda, et al., “OAM channel proposal for GEPOF”, IEEE 802.3bv TF, Plenary Meeting, March 2015*
- [5] *Rubén Pérez-Aranda, et al., “64B/65B PCS encoding for GEPOF”, IEEE 802.3bv TF, Plenary Meeting, March 2015*



Questions?