802.3ch EEE Overview

KY-ANH TRAN

09/18/24

Supporters

- Alireza Razavi

AEONSEMI

Agenda

- Motivation

- 802.3ch EEE Overview
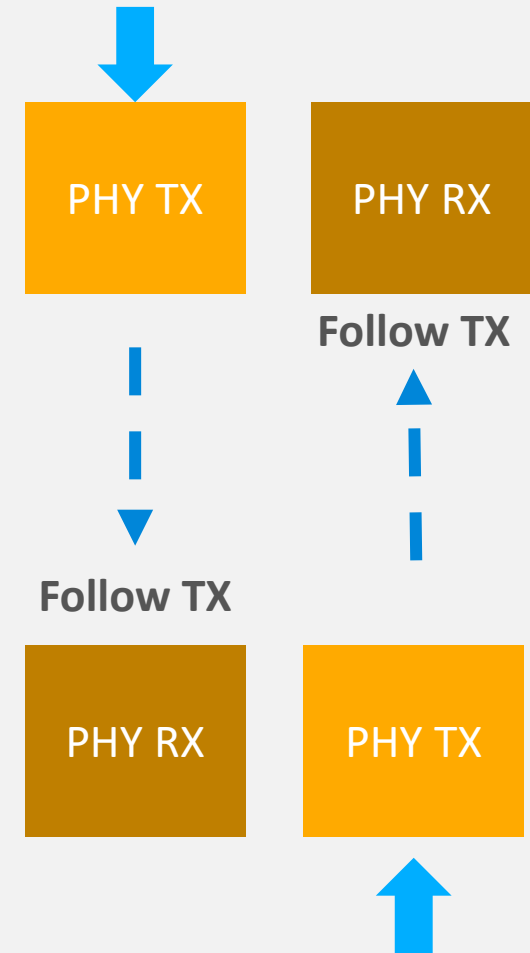
- LPI mode entry/exit

- Timing and Latency parameters

Motivation

- EEE has been discussed as power saving mechanism for 802.3ch Sensor PHY.
    - https://www.ieee802.org/3/cfi/0723_1/CFI_01_0723.pdf
        - "Can provide power saving with EEE (quiet-refresh cycling) based on traffic offered"
    - https://www.ieee802.org/3/cfi/0723_1/CFI_01_0723.pdf
        - "Can we adapt 802.3ch/quiet-refresh to meet the need?"

- There is interest to re-use existing specification if it can meet the needs of camera sensor.

- We will present here some information regarding the 802.3ch EEE spec which is relevant to 802.3dm

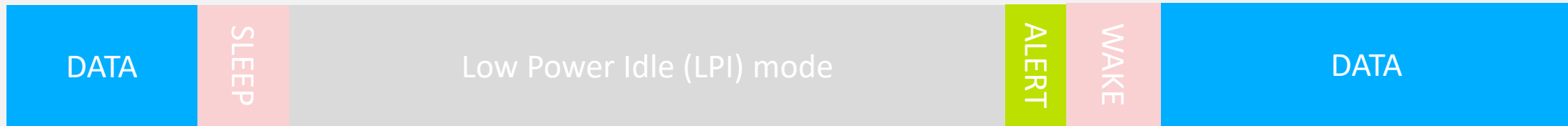AEONSEMI

802.3ch EEE Features

- EEE in 802.3ch provides a mechanism to lower power dynamically.

- The PHY TX initiates the power mode transition. The PHY RX follows.
  - The 2 traffic directions enter/exit low power mode independently.

- Heuristic when to enter/exit low power mode done in higher layer. PHY will handle requests according to spec.
  - Allows flexibility for higher layer to initiate low power mode based on traffic availability.

- Additional mechanism for link partner to request exit low power mode, through OAM channel.
  - Not discussed here.

Enter/exit low power mode requests

| PHY TX | PHY RX |

Follow TX

Follow TX

| PHY RX | PHY TX |

Enter/exit low power mode requests

AEONSEMI

802.3ch EEE Overview

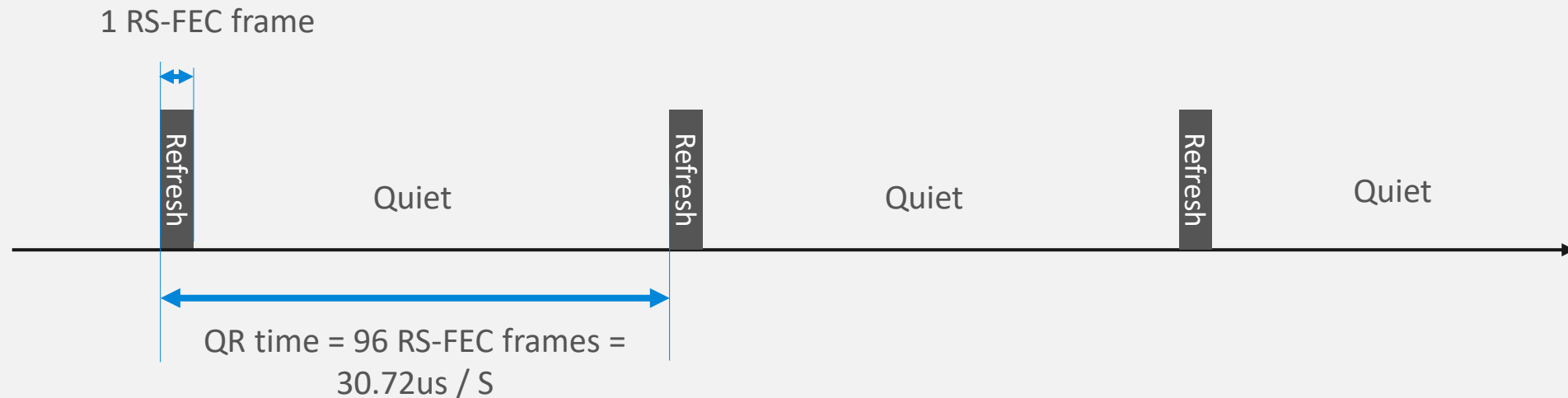| DATA	SLEEP	Low Power Idle (LPI) mode	ALERT	WAKE	DATA

- DATA mode:
  - Ethernet packets can flow through.

- ALERT:
  - ALERT reception starts the wake up process

- SLEEP/WAKE:
  - This is a transition state when entering/exiting low power

- Low Power Idle (LPI) mode:
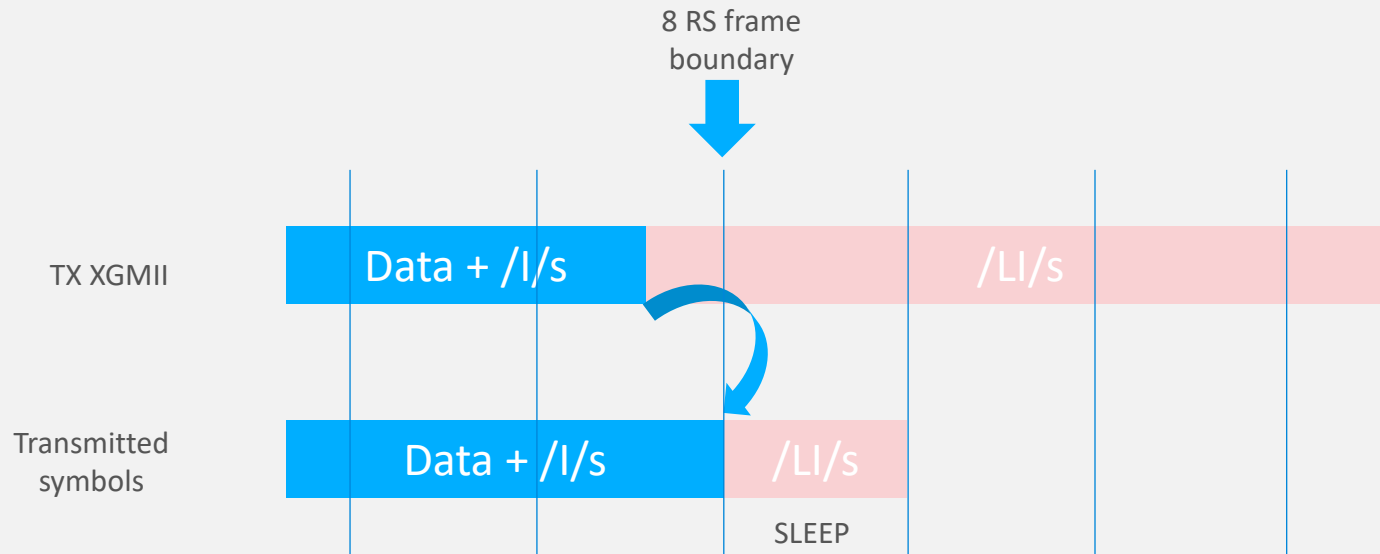 - Periodic refresh to keep clock/adaptation updated.

AEONSEMI

Low power Idle (LPI) mode

- In the low power state, the PHY TX periodically transmits "Refresh" signals to keep adaptation loops and clock recovery updated.
  - This is called the QR (quiet-refresh) cycle.
  - Refresh is 1/96 ~ 1% of the QR cycle

1 RS-FEC frame

| Refresh | Quiet | Refresh | Quiet | Refresh | Quiet |

QR time = 96 RS-FEC frames = 30.72us / S

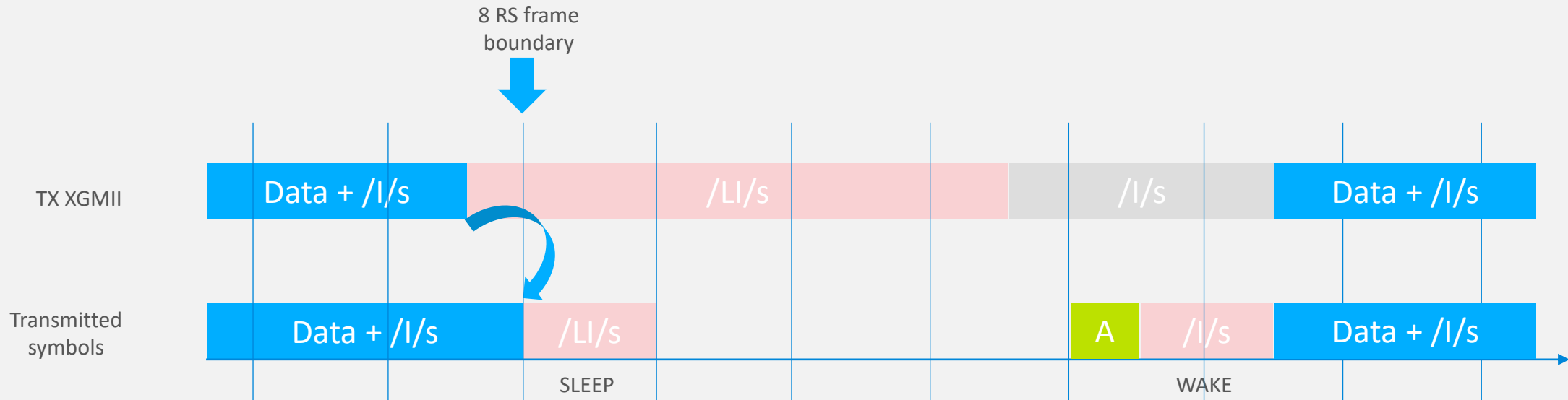ÆONSEMI

Going to LPI mode through /LI/ characters

- Upon receipt of /LI/ characters at the XGMII interface, the PCS TX initiates SLEEP at the next available 8 RS-FEC frame window.

- RX side enters LPI upon receipt of /LI/ characters.

8 RS frame
boundary

TX XGMII — Data + /I/s — /LI/s

Transmitted symbols — Data + /I/s — /LI/s

SLEEP

AEONSEMI

Exiting LPI mode through /I/ characters

- Upon receipt of non /LI/ characters at the XGMII interface, the PCS TX initiates wake up at the next alert opportunity.
  - Frequency of alert opportunities depend on **slow_wake**.

8 RS frame boundary

TX XGMII: Data + /I/s | /LI/s | /I/s | Data + /I/s

Transmitted symbols: Data + /I/s | /LI/s | A | /I/s | Data + /I/s

SLEEP   WAKE

AEONSEMI

TX Alert signal to wake up RX

- TX sends ALERT pattern to let RX know wake up sequence about to start.

- RX detects ALERT signal and starts wake up sequence.

- Alert is a known PAM2 sequence (same as SEND_S).
  - Baud rate is slow: 703.125MHz to ensure easy reception.
  - [1, 2, 4] * 4 * 0.320ns * 703.125MHz = [900, 1800, 3600] transmitted symbols for [10, 5, 2.5] Gb/s

- In 802.3ch, ALERT can be sent only at specific points in the QR cycle.

AEONSEMI
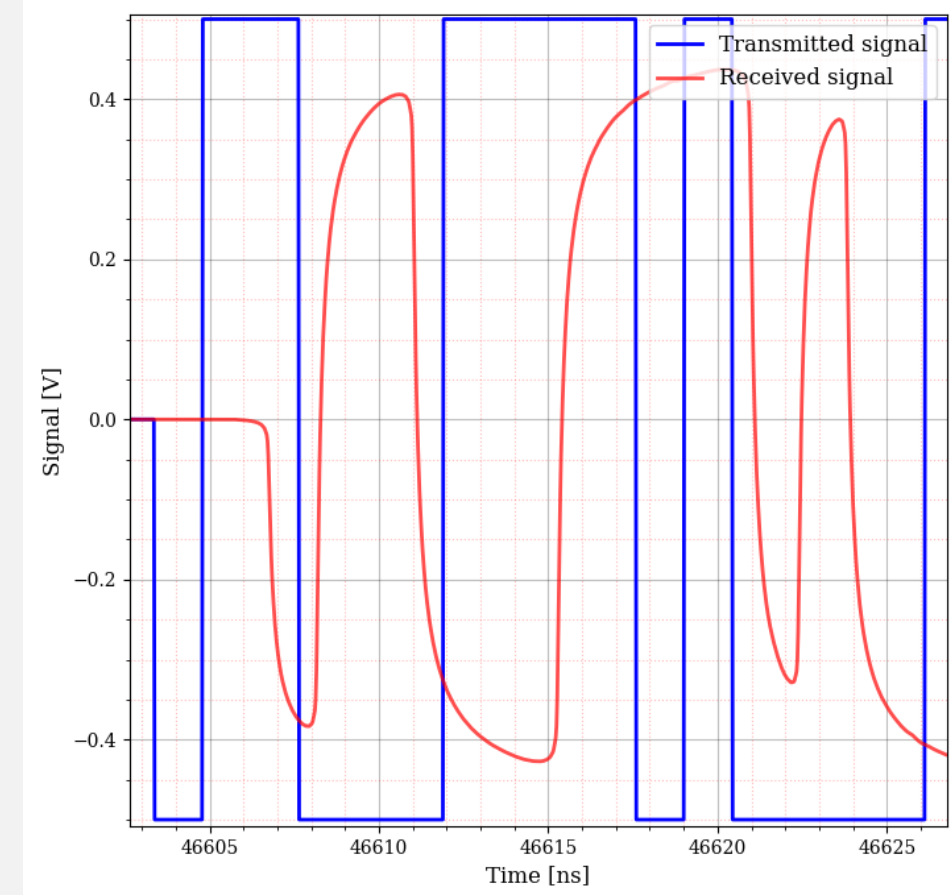
Timing parameter summary
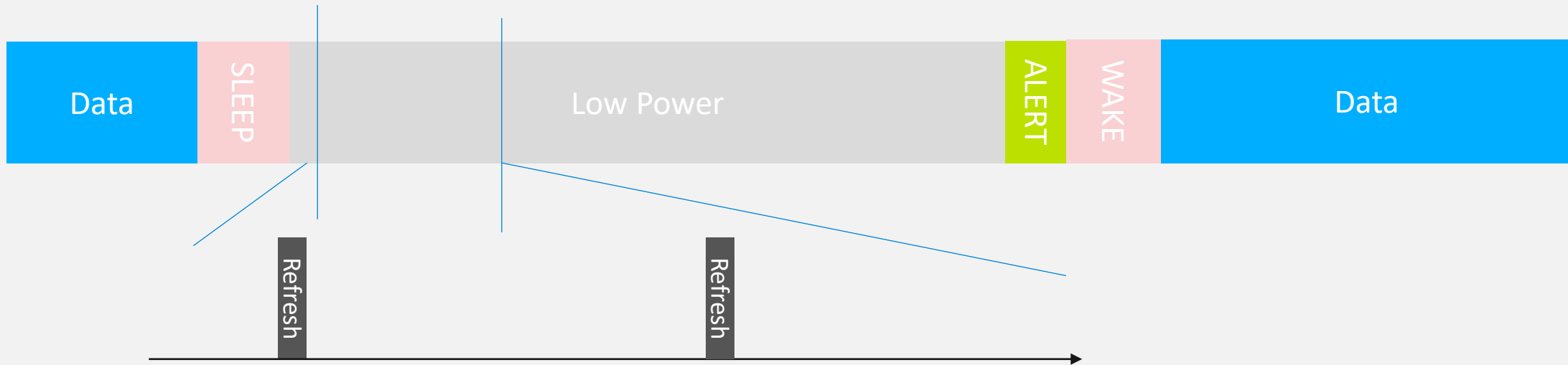
Parameter	Value
QR cycle length	30.72us / S
Refresh	0.32us / S
ALERT time	1.28us/ S
WAKE time	2.56us / S
SLEEP time	2.56us / S

- S = 0.25/0.5/1.0 for 2.5/5/10G

AEONSEMI

Latency Specification

- Key Question:
  - How long do I have to wait from the time I request wakeup at XGMII to the time I can transmit data at XGMII interface?
  - This number is bounded by spec under the name "**T_w_sys_tx**"
- Alert start opportunity sparsity determine how long before wake up sequence can start.

wakeup request

alert start opportunity

alert start opportunity

ALERT

WAKE

Latency incurred at XGMII interface

AEONSEMI
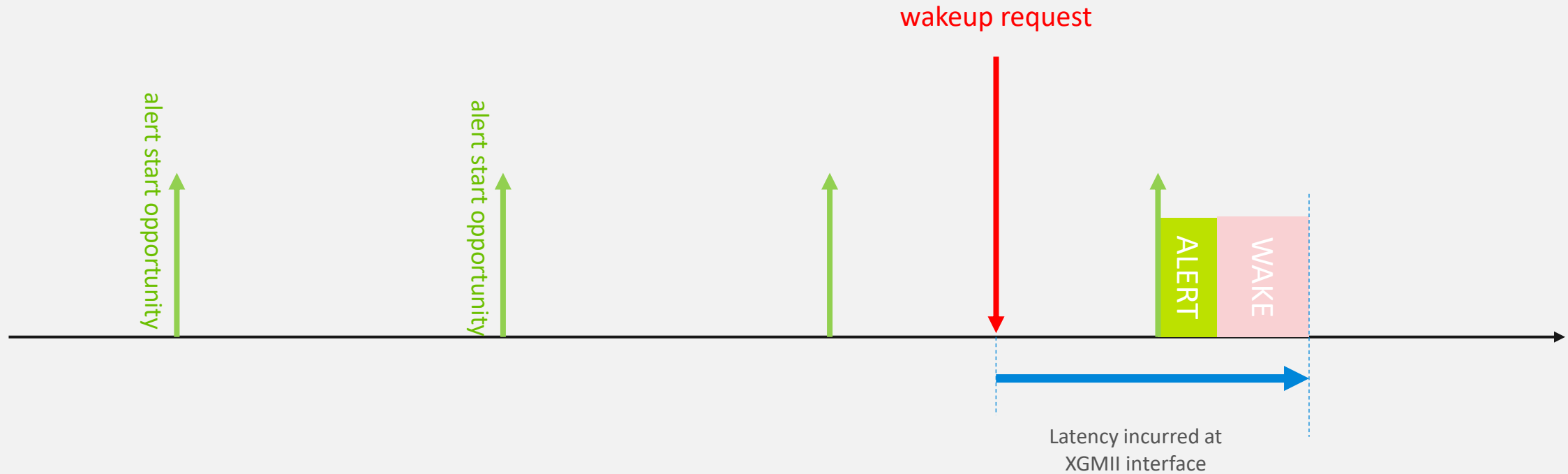
Latency Specification

- **T_w_sys_tx** depends on **slow_wake** mode.
  - In **slow_wake** = 0, alert opportunity is more frequent (every 8 RS-FEC frames).
  - In **slow_wake** = 1, it is every refresh (96 RS-FEC frame).

- **T_w_sys_tx** depends on whether sleep signal was completed when request to wake up was done.
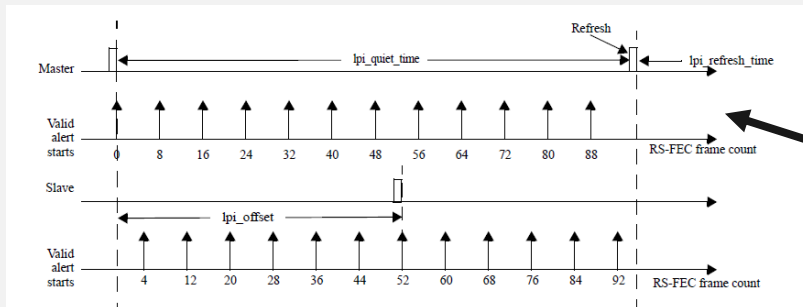

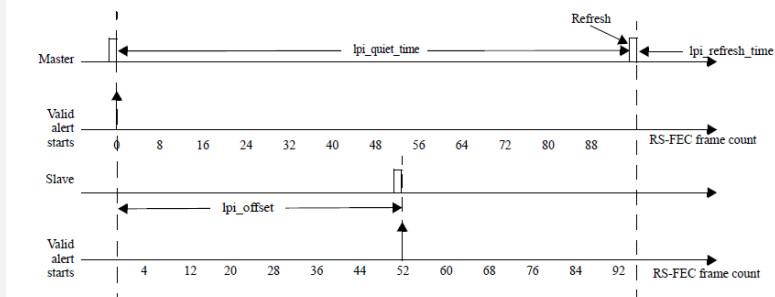Figure 149–13—Timing periods for LPI signals when Slow Wake not active

Figure 149–14—Timing periods for LPI signals when Slow Wake is active

More frequent alert opportunities when slow wake is not active

T_w_sys_tx	slow_wake = 0	slow_wake = 1
Before sleep complete	8.96us / S	37.12us / S
After sleep complete	6.4us / S	34.56us / S

AEONSEMI

Conclusion

- 802.3ch EEE provides a mechanism to dynamically save power based on traffic availability.

- 802.3ch EEE provisions for the 2 traffic directions to enter/exit LPI mode independently.
  - Asymmetric power states are allowed.

- In the low power state, periodic refresh maintains equalizer/clock recovery updates.
  - The refresh window is 1% duty-cycled.

- The worst case time between request to transmit and ability to transmit data is specified by spec to be 6.4us / S or 8.96us / S depending on whether SLEEP signal completed with **slow_wake**=0.

- This number is important when computing link latency for 802.3dm.

AEONSEMI

References

- "MultiGig EEE Overview", Jim Graba (TC16 Open alliance, April 28 2022)

- 802.3ch specification

AEONSEMI