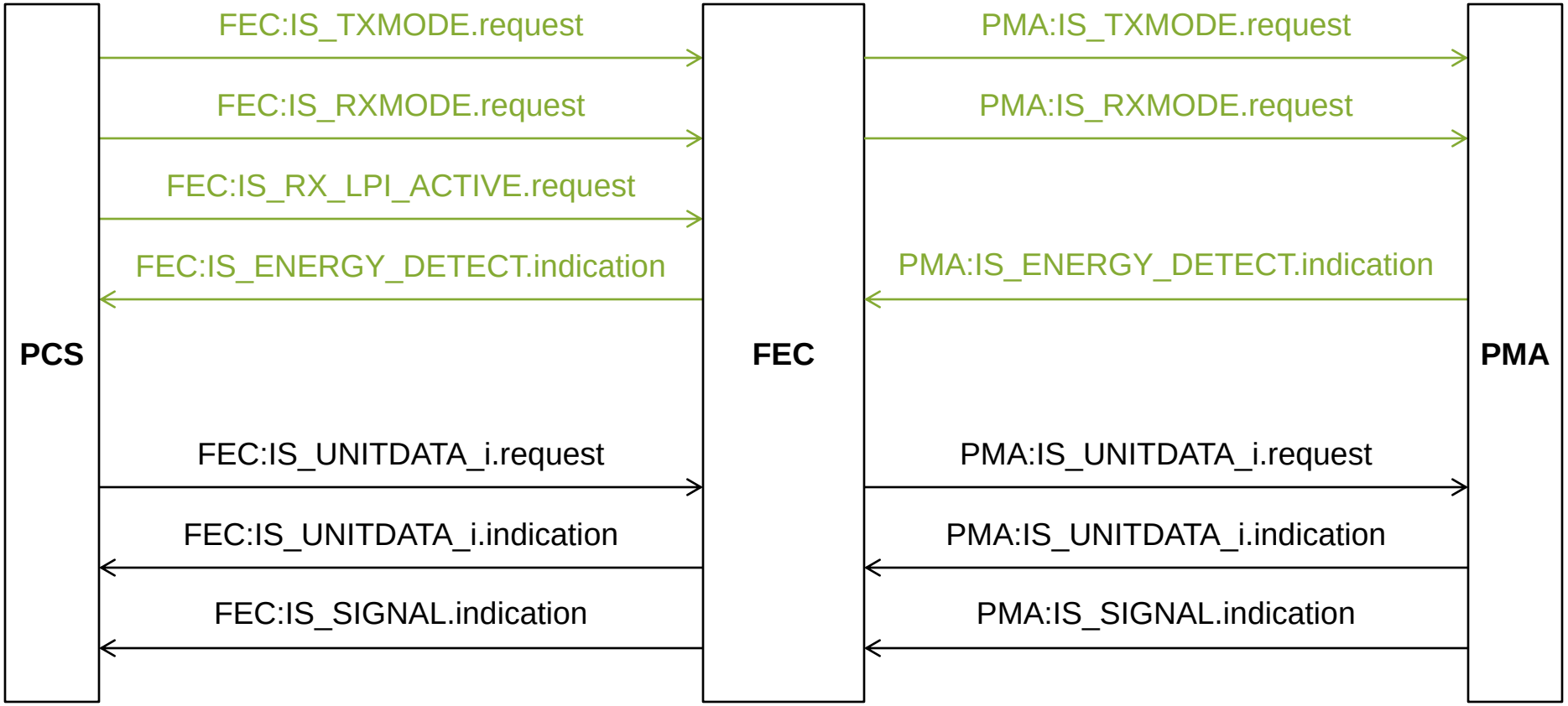
EEE deep sleep capability with XLAUI and CAUI (comment #110)

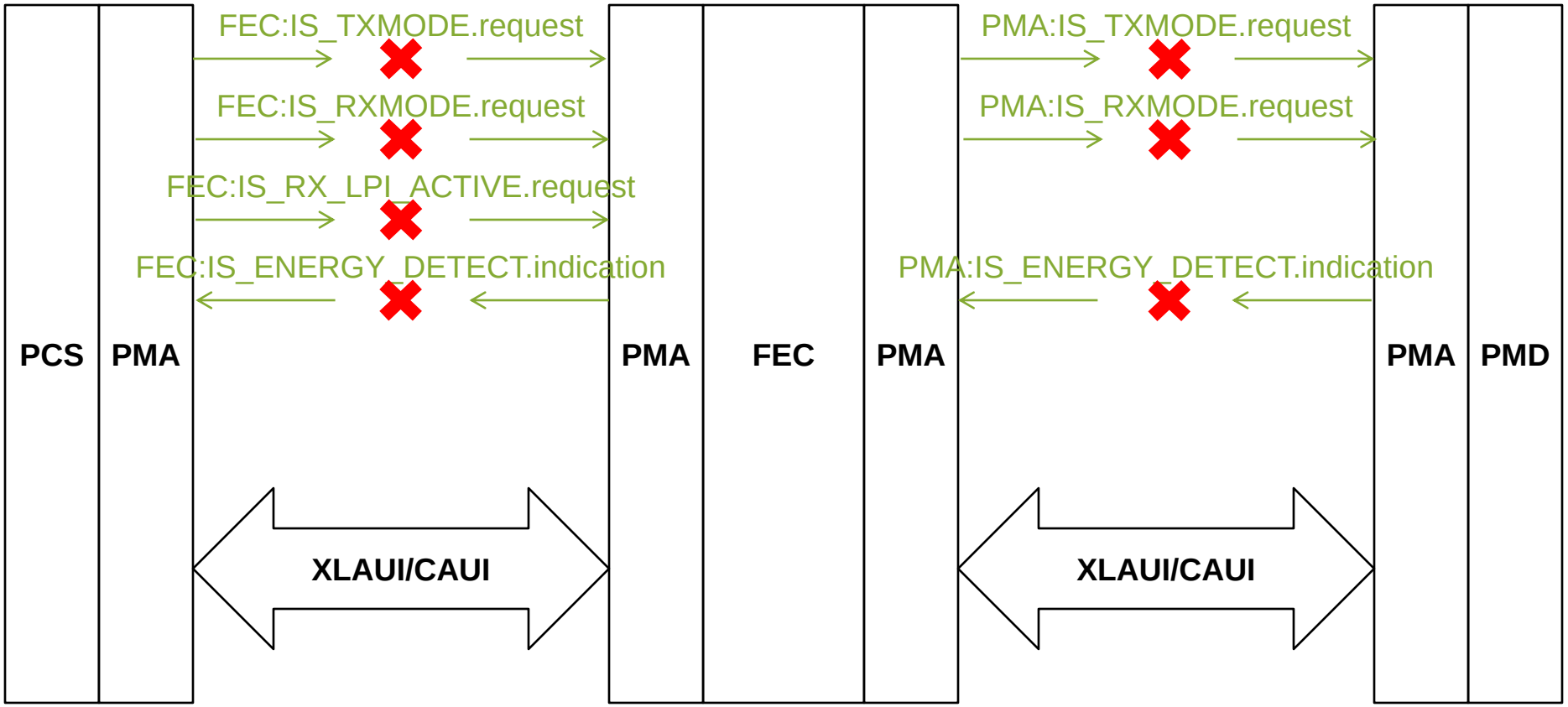
Adam Healey, LSI Corporation

IEEE P802.3bj Task Force
July 2013

FEC/PMA service interface



Include XLAUI or CAUI



XLAUI or CAUI does not provide a means to communicate the primitives required for correct operation of the optional EEE deep sleep capability

Problem statement

- nAUI does not provide a means to communicate the primitives required for correct operation of the optional EEE deep sleep capability
- Therefore, EEE deep sleep cannot be used with nAUI
- nAUI shutdown is incompatible with the optional EEE fast wake capability
- nAUI shutdown cannot be used at all

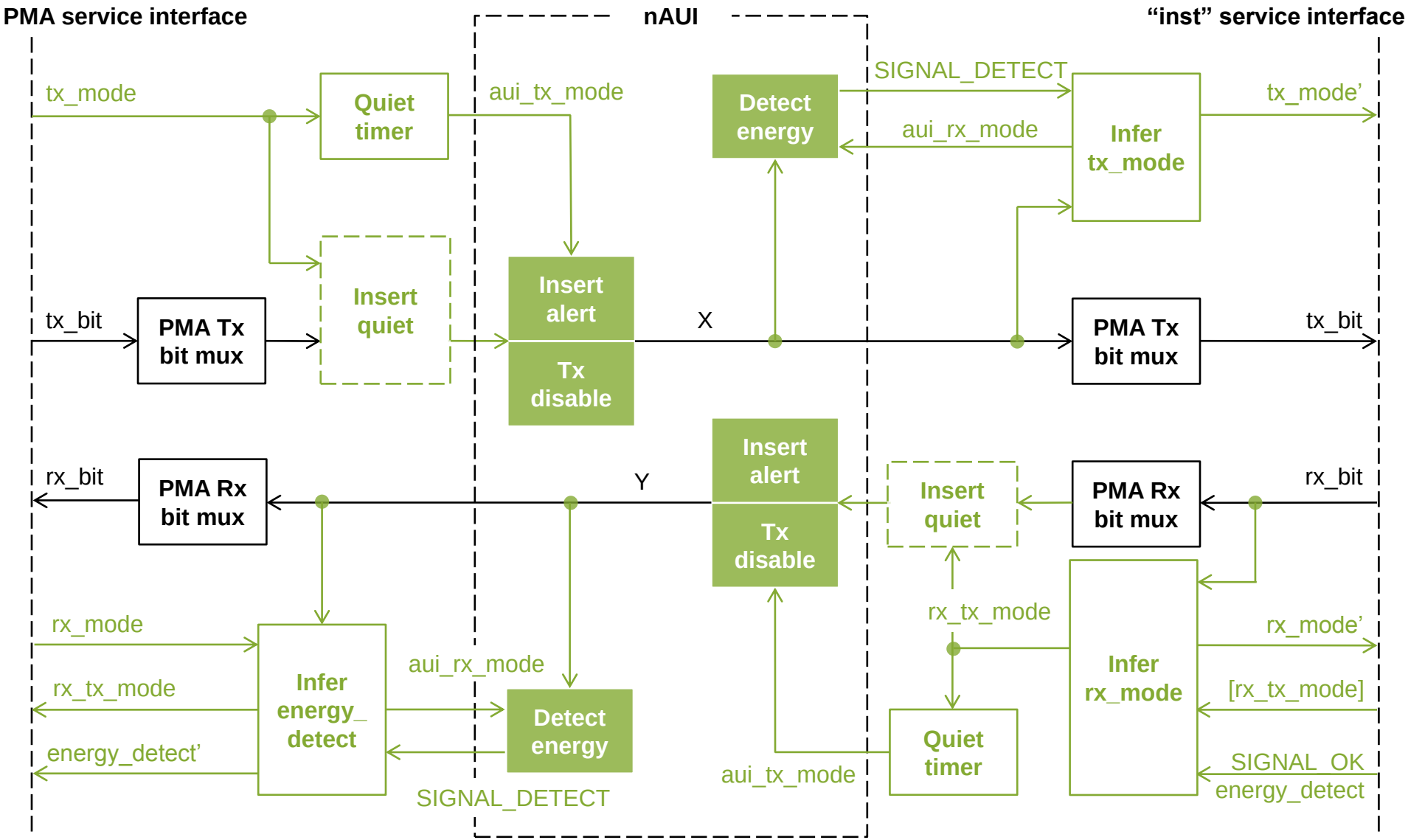
Objectives of this proposal

- Enable the use of nAUI with the optional EEE deep sleep capability
- Allow the user to choose whether or not to enable “nAUI shutdown” as part of an overall energy saving strategy
- Limit changes to the “responsible party” e.g. Clause 83 and Annexes
- Keep the PMA as simple as possible
- Define an architecture that can be extended to nAUI variants defined by future projects

General concepts

- Insert a PMA generated signal that represents tx_mode values QUIET and ALERT
- When possible, repeat the signal and detect it as it “passes by”
- Define a service interface primitive to convey the value across functions that corrupt the signal, e.g. bit multiplexing, transcoding
- When nAUI shutdown is allowed, define a minimum time to transmit “QUIET” that ensures detection before the transmitter is disabled
- When nAUI shutdown is allowed, define a “hold-off” time for signal detect so that it doesn’t accidentally trigger on the “QUIET” transmission

Proposed PMA functional diagram including nAUI

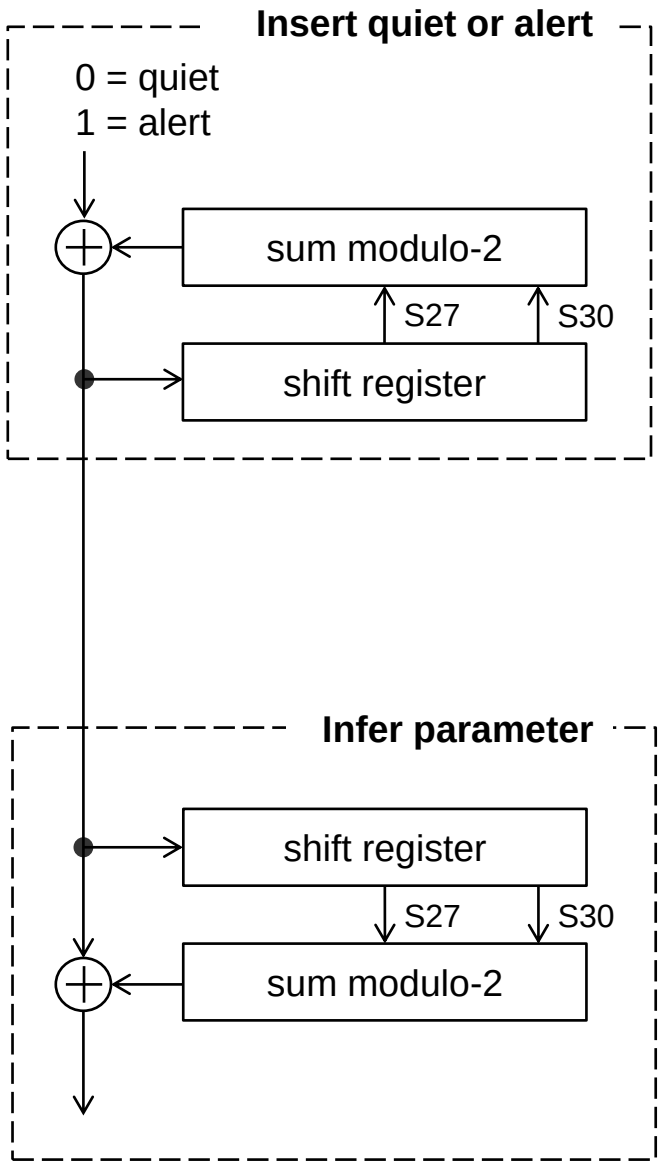


XX Functions present in D2.1

XX Functions added by this proposal

PMA quiet and alert signal definition

- Means to relay the value of tx_mode between PMA sublayers separated by nAUI or PMD-to-PMD link
- Simple to generate and quickly detect
- Scrambled pattern friendly to timing recovery and adaptive equalization
- Correlation between lanes can be controlled using initial shift register states
- Can be transmitted continuously when nAUI shutdown is not allowed



New service interface primitive

- PMA quiet and alert signals will be corrupted by transcoding, Forward Error Correction, or bit multiplexing
- In the Tx direction, tx_mode may be used to relay deep sleep state across these functions
- However, there is no such capability in the Rx direction
- Define IS_RX_TX_MODE.indication(rx_tx_mode) for this purpose

Proposed changes to Clause 80 [1]

Insert the following subclause at the end of 80.3.3.

80.3.3.8 IS_RX_TX_MODE.indication

The IS_RX_TX_MODE.indication primitive communicates the rx_tx_mode parameter. This parameter indicates the value of tx_mode that the PMA sublayer has inferred from the received signal. Without EEE deep sleep capability, the primitive is never generated and the sublayers behave as if rx_tx_mode=DATA.

80.3.3.8.1 Semantics of the service primitive

IS_RX_TX_MODE.indication(rx_tx_mode)

The parameter rx_tx_mode is assigned one of the following values: DATA, QUIET, or ALERT.

Proposed changes to Clause 80 [2]

80.3.3.8.2 When generated

This primitive is generated whenever there is change in the value of the rx_tx_mode parameter.

80.3.3.8.3 Effect of receipt

The specific effect of receipt of this primitive is defined by the sublayer that receives it.

Specifically, it is proposed that the FEC service interface for both Clause 74 and Clause 91 be updated to include this primitive and specify that rx_tx_mode is assigned the value received on the PMA service interface.

Proposed changes to Clause 83 [1]

Insert the following subclause at the end of 83.5.

83.5.11 Energy Efficient Ethernet

When the optional Energy Efficient Ethernet (EEE) deep sleep capability is supported, additional functions are required when the PMA service interface is physically instantiated as XLAUI or CAUI. These functions enable the communication of service interface parameters that are essential to the operation of the EEE deep sleep capability.

83.5.11.1 PMA quiet and alert signals

The PMA quiet and alert signals are generated on each lane with a self-synchronizing scrambler. The scrambler shall produce the same result as the implementation shown in Figure 83–XX. This implements the scrambler polynomial defined by Equation (83–YY).

$$G(x) = 1 + x^{28} + x^{31} \qquad (83–YY)$$

To generate the PMA quiet signal the input to the scrambler shall be 0. To generate the PMA alert signal the input to the scrambler shall be 1.

The initial state the scrambler of a given lane of PMA service interface is chosen to minimize the correlation between lanes.

Proposed changes to Clause 83 [2]

83.5.11.2 Detection of PMA quiet and alert signals

Each lane detects the PMA quiet and alert signals at the output of a self-synchronizing descrambler that implements the polynomial defined in Equation (83–YY). The descrambler shall produce the same result as the implementation shown in Figure 83–ZZ.

The output of the descrambler is considered in consecutive, non-overlapping blocks of 256 bits. If the number of zeros detected in a given 256-bit block is greater than or equal 224, then the lane shall indicate that the PMA quiet signal is detected. If the number of ones detected in a given block is greater than or equal to 224, then the lane shall indicate that the PMA alert signal has been detected. Otherwise, the lane infers that normal data is being received and shall not indicate that either the PMA quiet or PMA alert signal has been detected.

The PMA shall indicate that the quiet signal is detected when all lanes of the PMA service interface have detected the quiet signal. The PMA shall indicate that the alert signal is detected when all lanes of the PMA service interface have detected the alert signal. Otherwise, the PMA infers that normal data is being received and shall not indicate that either the PMA quiet or PMA alert signal has been detected.

NOTE: The observation window and thresholds are based on XLAUI or CAUI lanes operating at 10.3125 Gb/s. Future projects that consider operation at higher per-lane signaling rates may want to revisit these values.

Proposed changes to Clause 83 [3]

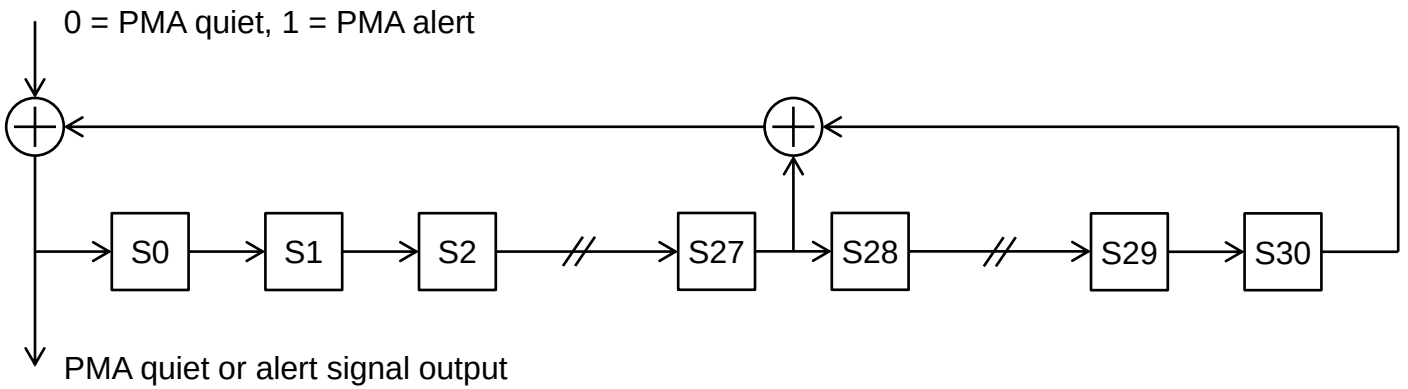


Figure 83–YY— Scrambler for PMA quiet and alert signals

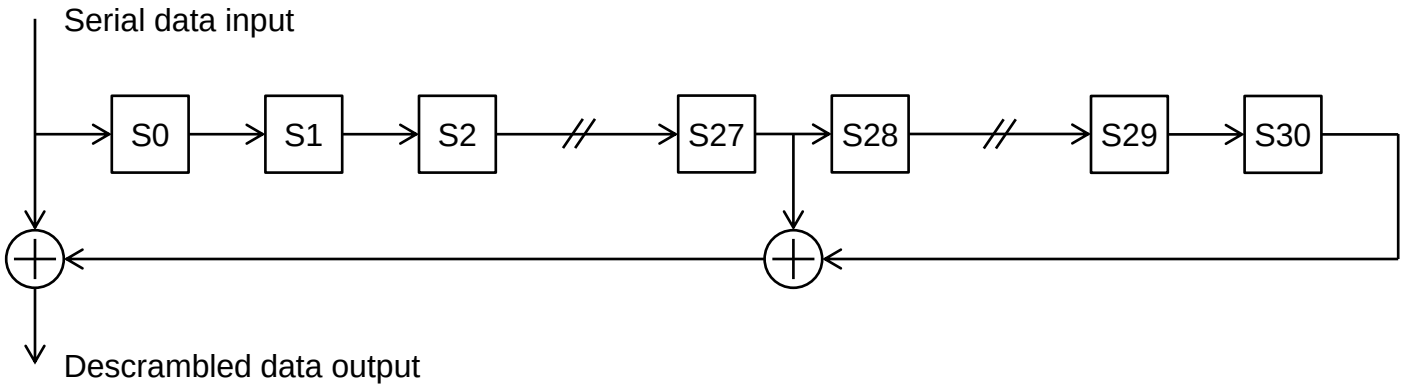


Figure 83–ZZ — Decrambler for PMA quiet and alert signals

Proposed changes to Clause 83 [4]

83.5.11.3 Additional transmit functions in the Tx direction

If the PMA client is the PCS, BASE-R FEC, or RS-FEC sublayer or is a PMA sublayer where the number of input lanes is not equal to the number of output lanes, then the PMA sublayer shall insert the PMA quiet and alert signals as follows. When the value of tx_mode is QUIET, the PMA inserts the PMA quiet signal defined in 83.5.11.1. When the value of tx_mode is ALERT, the PMA inserts the PMA alert signal defined in 83.5.11.1. For all other values of tx_mode, the PMA output is defined by the bit multiplexing function.

If XLAUI or CAUI is permitted to shut down (see 83.3), the variable aui_tx_mode shall be assigned the current value of tx_mode with the following exception. When tx_mode transitions from DATA to QUIET, the value of aui_tx_mode is held at DATA and the timer pma_quiet_timer (Tpq) is started. If tx_mode is QUIET when the timer expires, then aui_tx_mode is set to QUIET. If tx_mode is set to a value other than QUIET before the timer expires, then aui_tx_mode is set to DATA.

If XLAUI or CAUI is not permitted to shut down, aui_tx_mode shall be assigned the value DATA.

PMA functions in the Tx direction may be disabled in order to conserve energy while aui_tx_mode is QUIET.

Proposed changes to Clause 83 [5]

83.5.11.4 Additional receive functions in the Tx direction

The value of tx_mode shall be inferred as follows. If the PMA quiet signal is detected, the value of tx_mode is set to QUIET. If the PMA alert signal is detected, the value of tx_mode is set to ALERT. Otherwise, the value of tx_mode is DATA.

If XLAUI or CAUI is permitted to shut down (see 83.3), then the variable aui_rx_mode shall be assigned as follows. The variable aui_rx_mode is initialized to DATA upon PMA power on or reset. When the PMA quiet signal is detected, the timer hold_off_timer (Tho) is started. If the PMA alert signal is not detected before the timer expires, then aui_rx_mode is set to QUIET. While aui_rx_mode is QUIET, it will be set to DATA when SIGNAL_DETECT transitions from FAIL to OK. The value of tx_mode is inferred to be ALERT and the timer alert_timer (Ta) started upon a transition of aui_rx_mode from QUIET to DATA. The value of ALERT will be held until alert_timer expires after which the value of tx_mode will be set to DATA.

If XLAUI or CAUI is not permitted to shut down, aui_rx_mode shall be assigned the value DATA.

PMA functions in the Tx direction may be disabled in order to conserve energy while aui_rx_mode is QUIET.

Proposed changes to Clause 83 [6]

83.5.11.5 Additional transmit functions in the Rx direction

The value of rx_mode shall be inferred as follows. The value of rx_mode is initialized to DATA upon PMA power on or reset. When the PMA quiet signal is detected, the timer hold_off_timer (Tho) is started. If the PMA alert signal is not detected before the timer expires, then rx_mode is set to QUIET. While rx_mode is QUIET, it will be set to DATA when the PMA alert signal is detected or energy_detect (or SIGNAL_OK) transitions from false to true.

The value of rx_tx_mode may be passed via the PMA:IS_RX_TX_MODE.indication primitive otherwise it shall be inferred as follows. If the PMA quiet signal is detected, the value of rx_tx_mode is set to QUIET. The value of rx_tx_mode is set to be ALERT and the timer alert_timer (Ta) started upon a transition of the value of rx_mode from QUIET to DATA. The value of ALERT will be held until alert_timer expires after which the value of rx_tx_mode will be set to DATA.

If XLAUI or CAUI is permitted to shut down (see 83.3), the variable aui_tx_mode shall be assigned the current value of rx_tx_mode with the following exception. When rx_tx_mode transitions from DATA to QUIET, the value of aui_tx_mode is held at DATA and the timer pma_quiet_timer (Tpq) is started. If rx_tx_mode is QUIET when the timer expires, then aui_tx_mode is set to QUIET. If rx_tx_mode is set to a value other than QUIET before the timer expires, then aui_tx_mode is set to DATA.

Proposed changes to Clause 83 [7]

83.5.11.5 Additional transmit functions in the Rx direction (continued)

If XLAUI or CAUI is not permitted to shut down, aui_tx_mode shall be assigned the value DATA.

PMA functions in the Rx direction may be disabled in order to conserve energy while aui_tx_mode is QUIET.

If the PMA is the client of the BASE-R FEC or RS-FEC sublayer or a PMA sublayer where the number of input lanes is not equal to the number of output lanes, then the PMA sublayer shall insert the PMA quiet and alert signals as follows. When the value of rx_tx_mode is QUIET, the PMA inserts the PMA quiet signal defined in 83.5.11.1. When the value of rx_tx_mode is ALERT, the PMA inserts the PMA alert signal defined in 83.5.11.1. For all other values of rx_tx_mode, the PMA output is defined by the bit multiplexing function.

Proposed changes to Clause 83 [8]

83.5.11.6 Additional receive functions in the Rx direction

The value of energy_detect shall be inferred as follows. The value of energy_detect is initialized to true upon PMA power on or reset. When the value of rx_mode is set to QUIET, the value of energy_detect is set to false. The value of energy_detect is set to true when the PMA alert signal is detected or SIGNAL_DETECT transitions from FAIL to OK.

The value of rx_tx_mode shall be inferred as follows. If the PMA quiet signal is detected, the value of rx_tx_mode is set to QUIET. The value of rx_tx_mode is set to be ALERT and the timer alert_timer (Ta) started upon a transition of the value of rx_mode from QUIET to DATA. The value of ALERT will be held until alert_timer expires after which the value of rx_tx_mode will be set to DATA.

If XLAUI or CAUI is permitted to shut down (see 83.3), then the variable aui_rx_mode shall be assigned as follows. The variable aui_rx_mode is initialized to DATA upon PMA power on or reset. When the PMA quiet signal is detected, the timer hold_off_timer (Tho) is started. If the PMA alert signal is not detected before the timer expires, then aui_rx_mode is set to QUIET. While aui_rx_mode is QUIET, it will be set to DATA when SIGNAL_DETECT transitions from FAIL to OK. The value of tx_mode is inferred to be ALERT and the timer alert_timer (Ta) started upon a transition of aui_rx_mode from QUIET to DATA. The value of ALERT will be held until alert_timer expires after which the value of tx_mode will be set to DATA.

Proposed changes to Clause 83 [9]

83.5.11.6 Additional receive functions in the Rx direction (continued)

If XLAUI or CAUI is not permitted to shut down, aui_rx_mode shall be assigned the value DATA.

PMA functions in the Rx direction may be disabled in order to conserve energy while aui_rx_mode is QUIET.

83.5.11.7 Support for BASE-R FEC

When the PMA is a client of the BASE-R FEC sublayer, the rx_lpi_active parameter of the IS_RX_LPI_ACTIVE.request primitive shall be defined as follows. The value of rx_lpi_active is initialized to false upon PMA power on or reset. The value of rx_lpi_active is set to true and the timer rx_lpi_active_timer (Tht) started upon a transition of the value of rx_mode from QUIET to DATA. When the timer expires, the value of rx_lpi_active is set to false.

Proposed changes to Annex 83A

Replace all instances of tx_mode with aui_tx_mode and all instances of rx_mode with aui_rx_mode.

Proposed changes to Clause 94

Clause 94 includes the definition of a PMA sublayer that is specific to 100GBASE-KP4 and the proposed changes to Clause 83 do not apply.

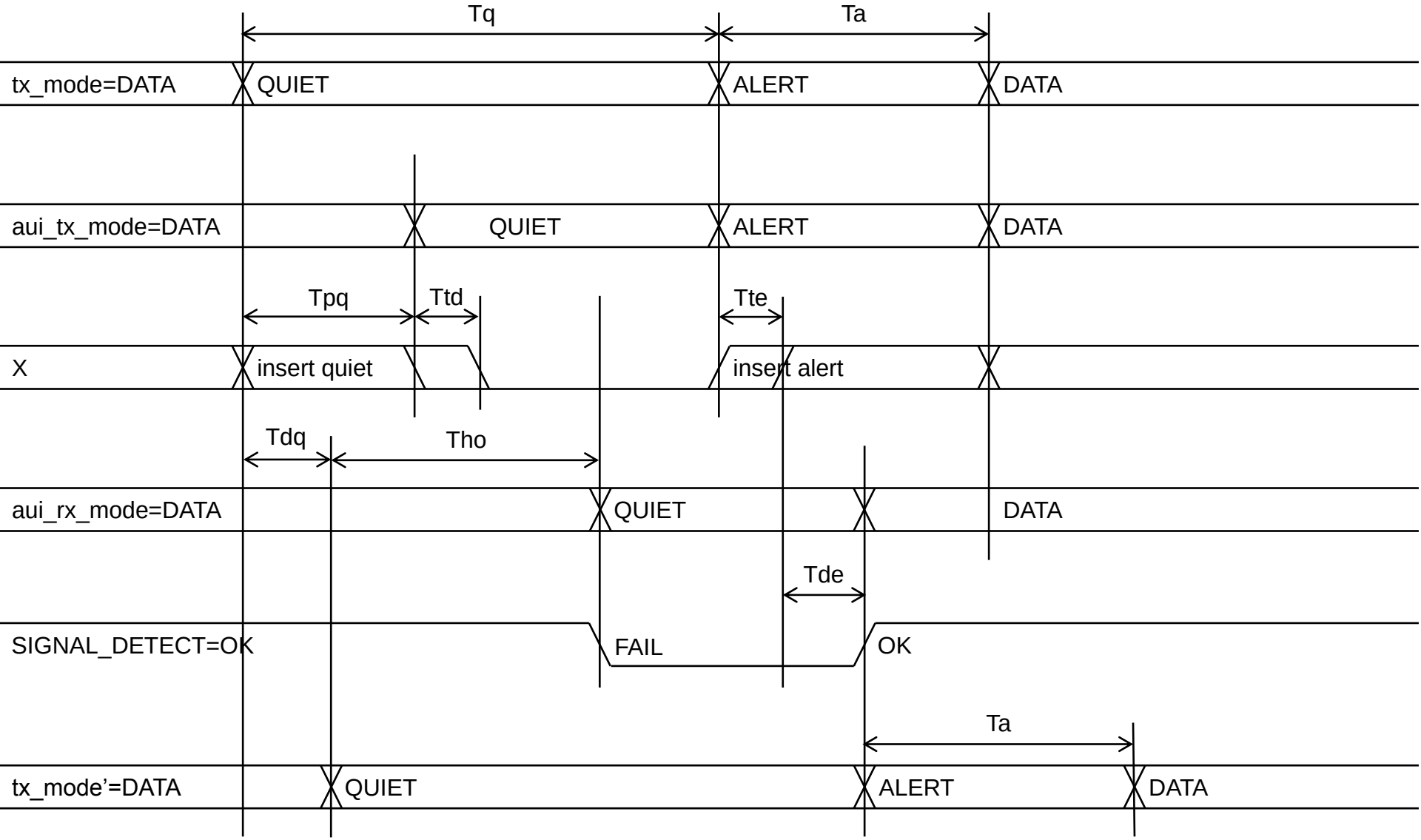
However, the service interface must include the PMA:IS_RX_TX_MODE.indication primitive and the value of rx_tx_mode must be defined. This will be passed through the RS-FEC sublayer to enable the CAUI implementation that could exist above.

Timing parameters

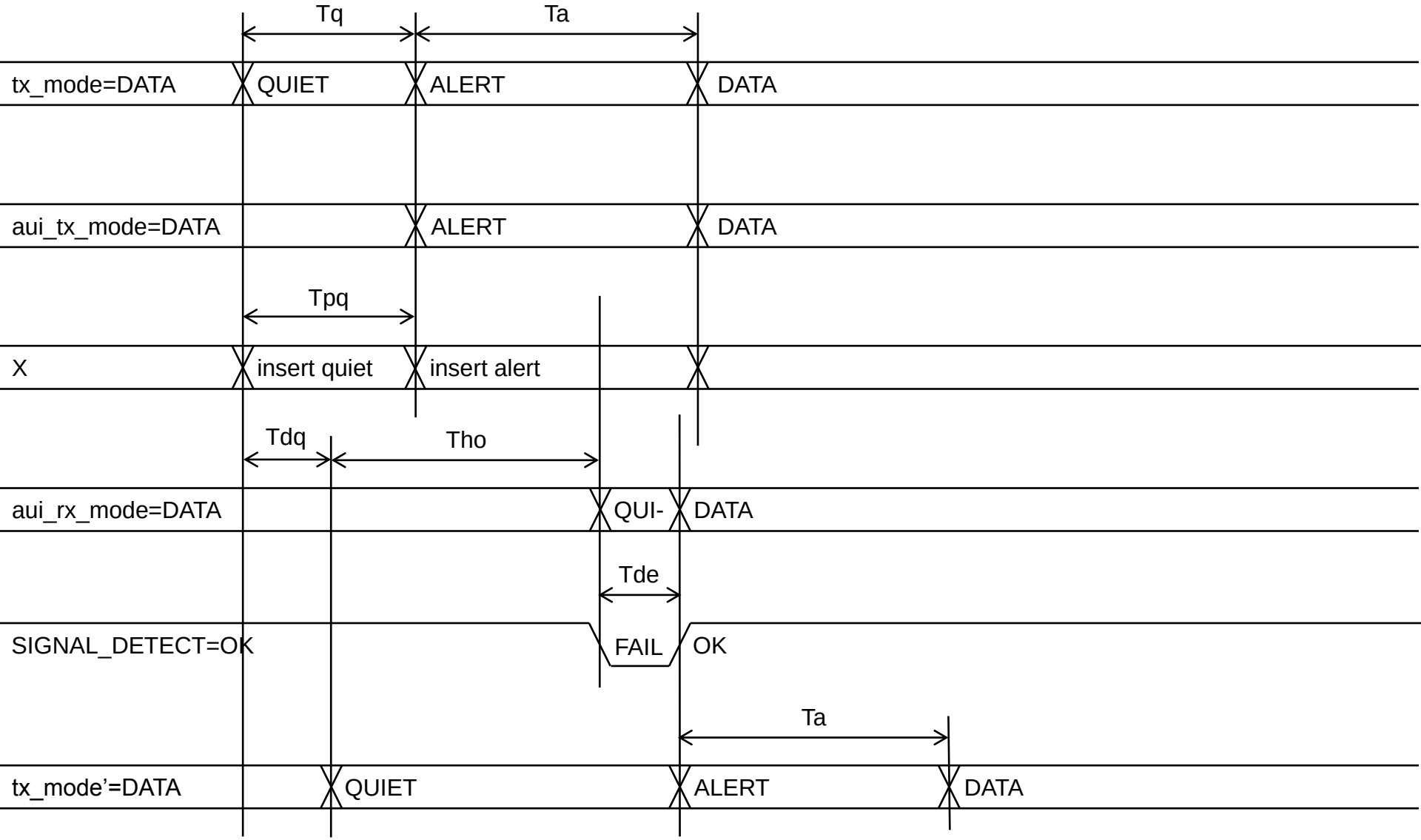
Timer	Symbol	Min.	Max.	Units
Quiet time	Tq	variable		
Alert signal duration	Ta	~~1100~~ 1150	1300	ns
Time to disable transmitter	Ttd	—	500	ns
Time to enable transmitter	Tte	—	500	ns
Time to assert energy detect	Tde	—	500	ns
PMA quiet signal duration	Tpq	200	225	ns
Energy detect hold-off time	Tho	750	800	ns
Time to assert PMA quiet detect	Tdq	25	50	ns
Time to assert PMA alert detect	Tda	—	25	ns
Time to hold rx_lpi_active=true	Tht	4000	5500	ns

Green text indicates timers that are added or modified by this proposal.

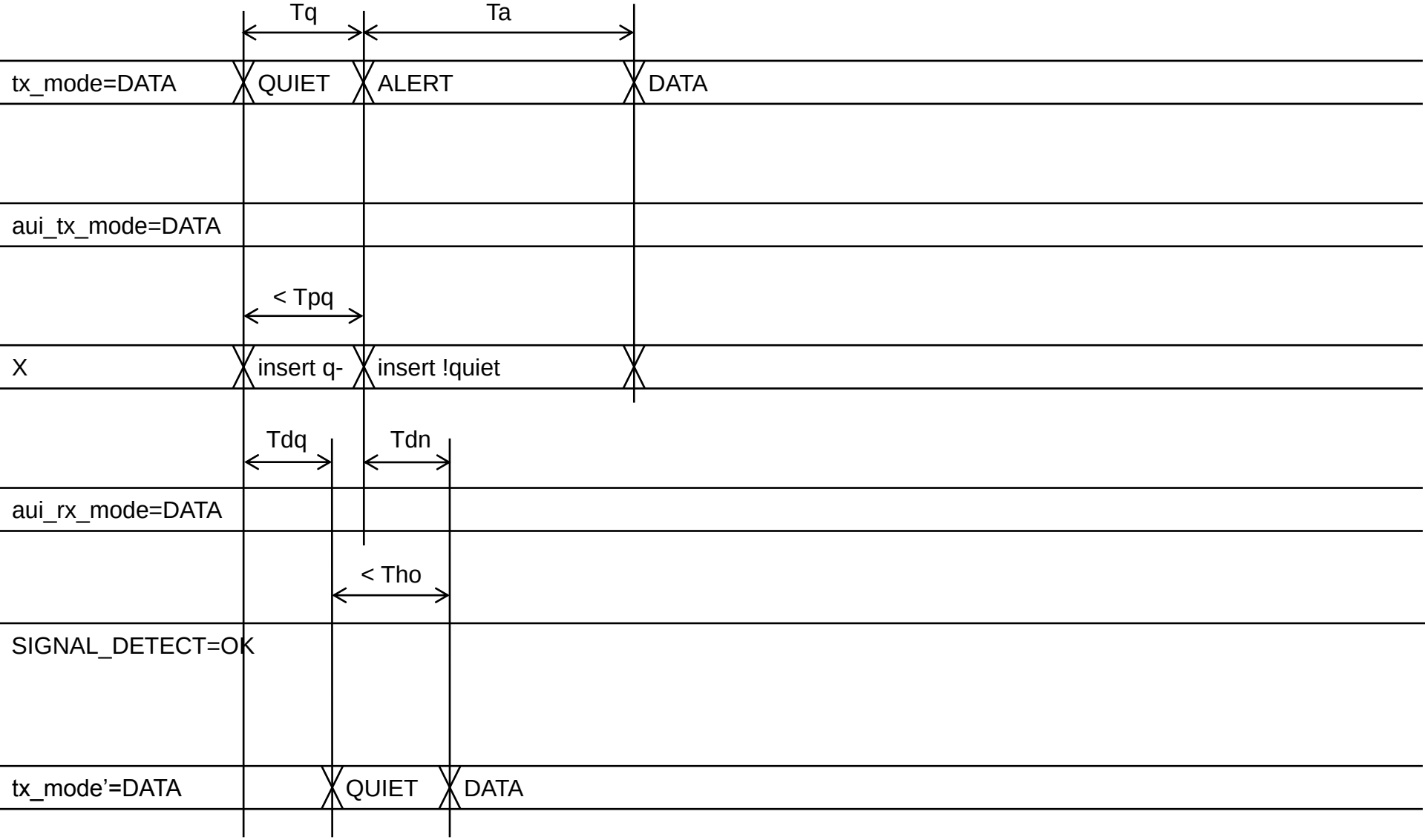
Tx direction timing: Tq > Tpd



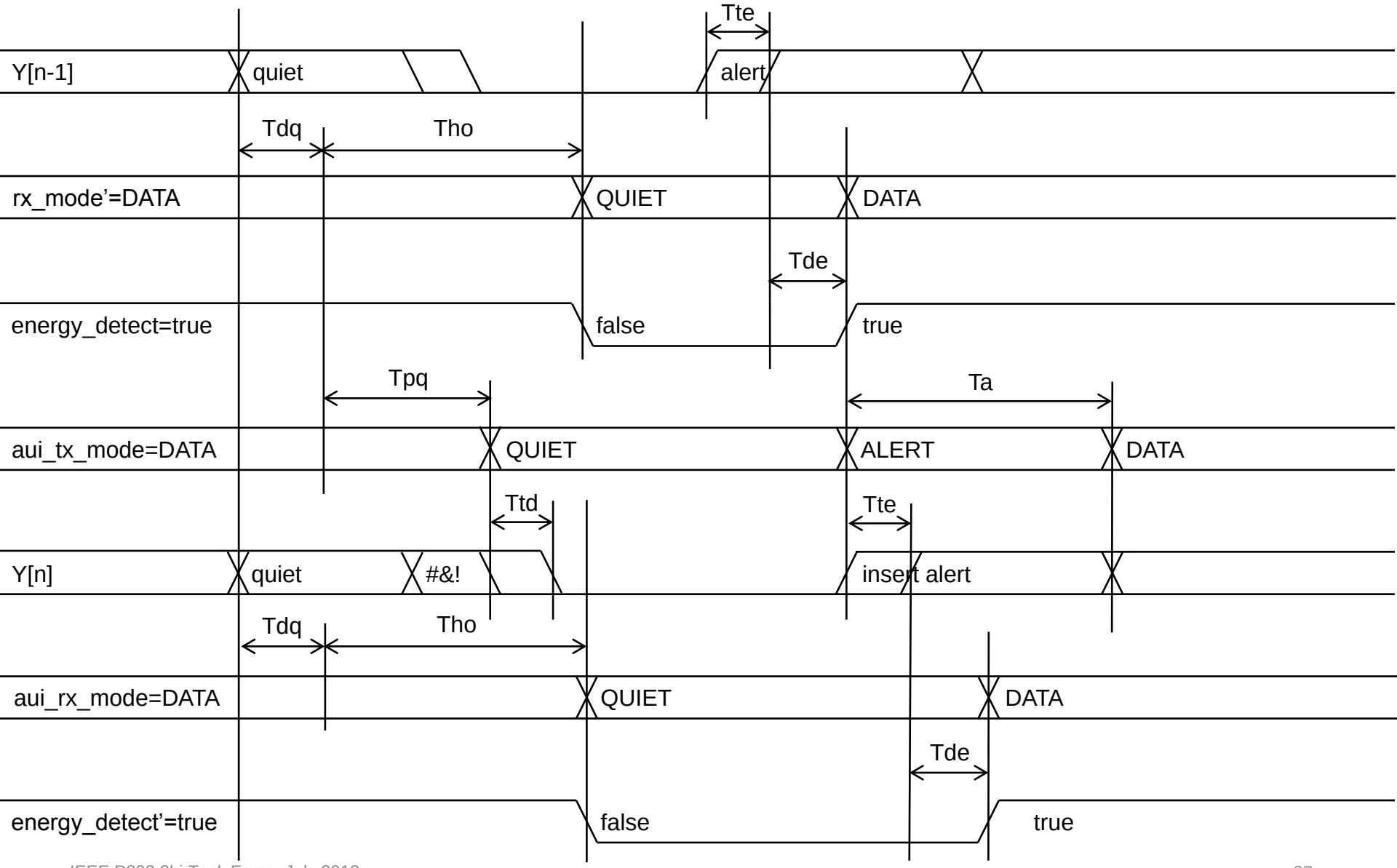
Tx direction timing: Tq = Tpd+



Tx direction timing: Tq < Tpq



Rx direction timing

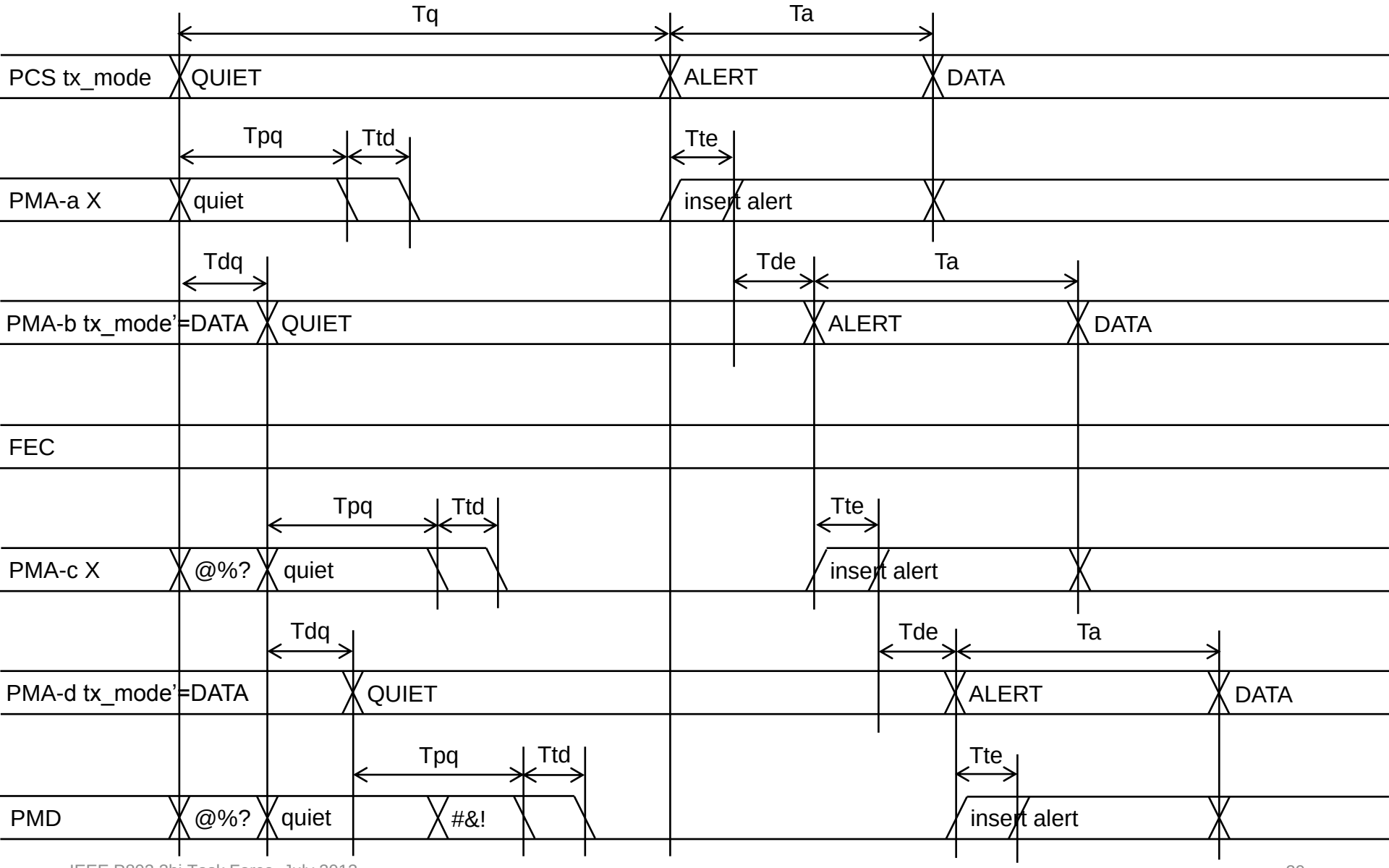


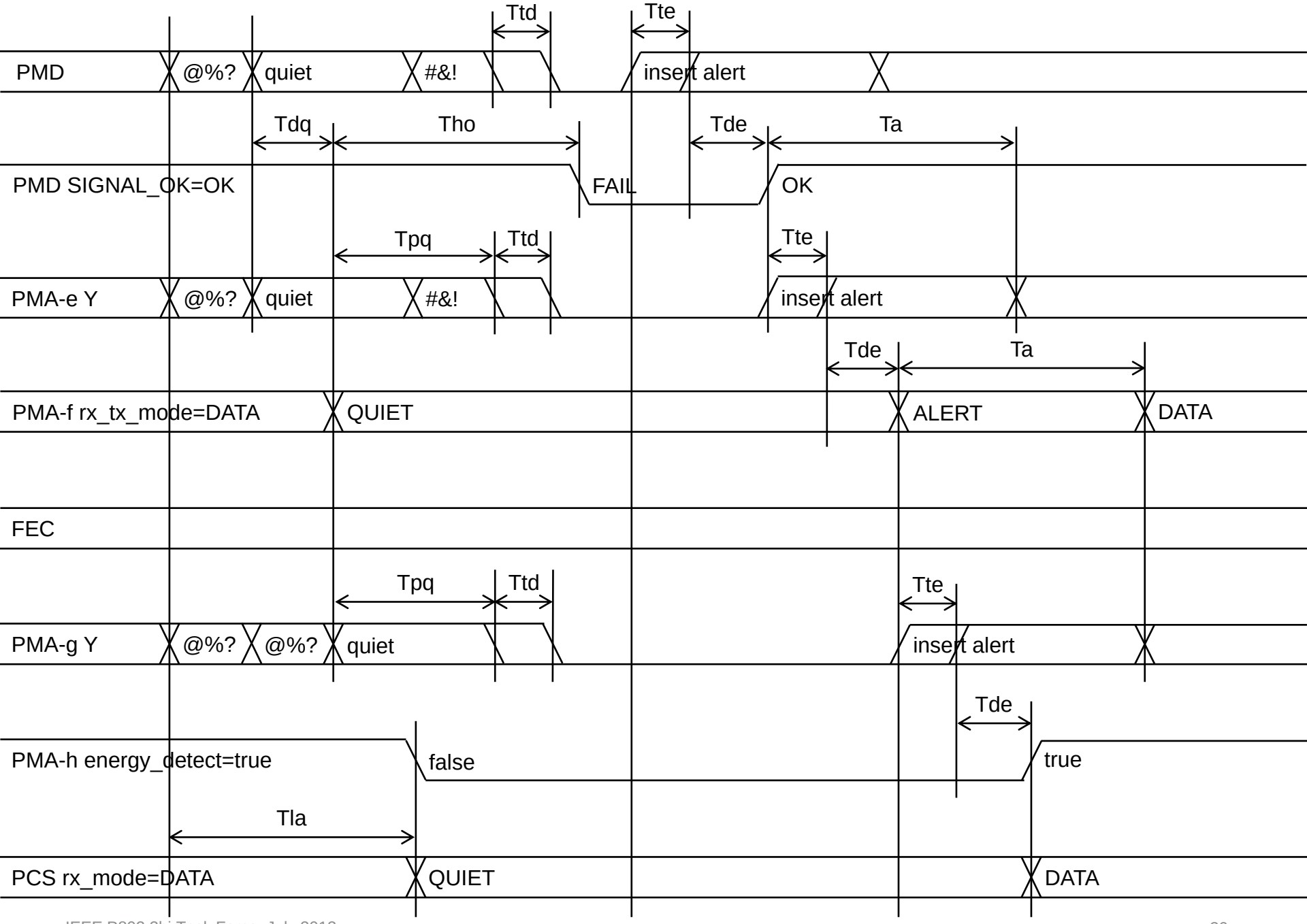
Summary of timing constraints

- Tho(max.) = Tpq(min.) + Ta(min.) – Tdq(max.) – Tde(max.)
- Tho(min.) = Tpq(max.) + Ttd(max.) – Tdq(min.)
- Tho(min.) = Tpq(max.) + Tdn(max.) – Tdq(min.)
- Tho(min.) = Tpq(max.) + Ttd(max.) + Tdq(max.) – Tdq(min.)

Timing parameters have been chosen to satisfy these constraints

End-to-end timing




Ttd
Tte
PMD
@%?
quiet
#&!
insert alert
Tdq
Tho
Tde
Ta
PMD SIGNAL_OK=OK
FAIL
OK
Tpq
Ttd
Tte
PMA-e Y
@%?
quiet
#&!
insert alert
Tde
Ta
PMA-f rx_tx_mode=DATA
QUIET
ALERT
DATA
FEC
Tpq
Ttd
Tte
PMA-g Y
@%?
@%?
quiet
insert alert
Tde
PMA-h energy_detect=true
false
true
Tla
PCS rx_mode=DATA
QUIET
DATA

Implications to wake time

- For each nAUI instance that is allowed to shut down, the wake time must be increased by Tte(max.) + Tde(max.) which is 1000 ns

- For each nAUI instance that is allowed to shut down, the apparent time from the start of wake to scrambler bypass is shortened; the additional wake time should precede scrambler bypass for PHYs that include BASE-R FEC

- If the quiet time (Tq) is very short, nAUI instances (or the PMD-to-PMD link) may not power down due to PMA quiet signal transmission

Closing remarks

- Proposal addresses a major shortcoming in the EEE deep sleep capability defined in IEEE P802.3bj/D2.1

- Definition of rx_tx_mode for the 100GBASE-KP4 PMA is for further study