

Proposed Modifications to 1000BASE-T1 State Diagrams and Variables

Pittsburgh, PA, USA

May 18, 2015

Peiqing Wang and Mike Tu

peiqingw@broadcom.com

tum@broadcom.com

Contributors

- Jim Graba
- Ahmad Chini
- Steven Chen
- Jay Cordaro
- Mehmet Tazebay

Motivations

- Fix corner cases in D1.4 state diagrams
- Improve robustness of startup sequence
- Provide a way to minimize packet losses under varying noise conditions

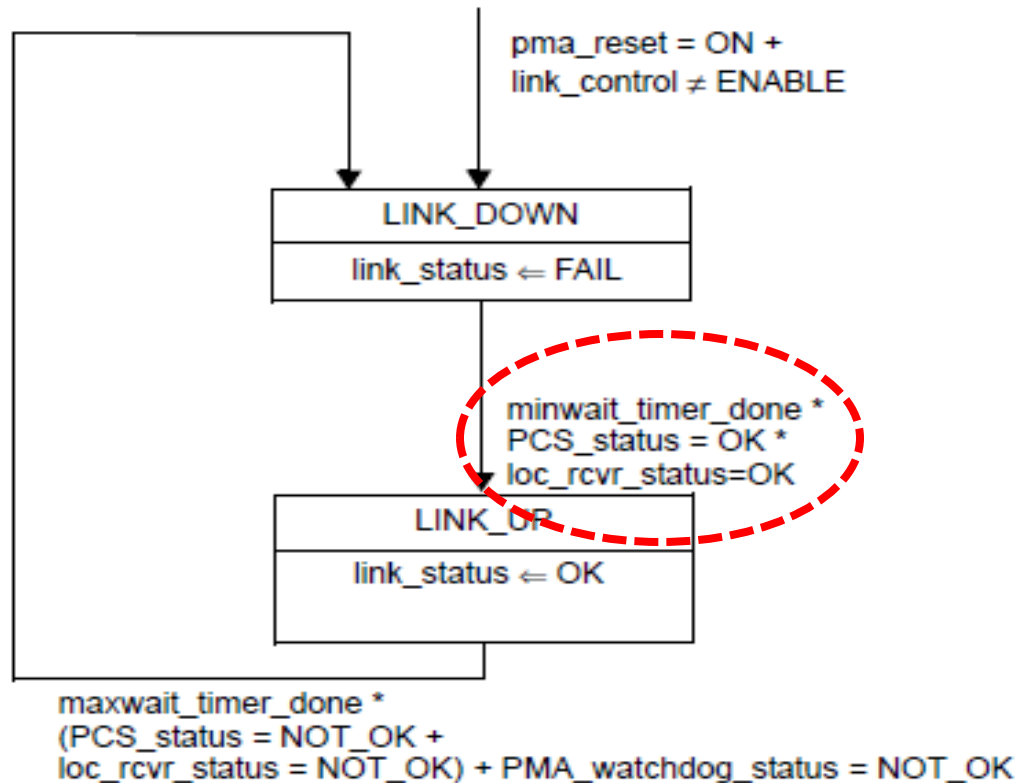
Terminology to avoid confusion

- Relink
 - The two PHY's failed to reach data mode when the specified time limit (100msec) expires, and have to start over either from Auto-Negotiation or the PHY Link Synchronization.
- Retrain
 - The two PHY's failed to reach data mode while still within the specified time limit, or a recovery attempt has failed. They will retry starting from the half-duplex training stage.
- Recover
 - The two PHY's have already reached data mode. Due to temporary disturbances they may attempt recovery while staying in 3B2T mode.

Topic #1:

Issues with Figure 97-23 Link Monitor state diagram

Original Figure 97-23



NOTE 1—maxwait_timer is started in PHY Control state diagram (see Figure 97–22).
NOTE 2—The variables link_control and link_status are designated as link_control_1GigT1 and link_status_1GigT1, respectively, by the Auto-Negotiation Arbitration state diagram (Figure 98–14) if the optional Auto-Negotiation function is implemented.

Figure 97–23—Link Monitor state diagram

D1.4 97.3.7.1

- PCS_status:
 - Indicates whether the PCS is in a fully operational state. **It is only true if block_lock is true and hi_rfer is false.** This status is reflected in MDIO register 3.32.12. A latch low view of this status is reflected in MDIO register 3.1.2 and a latch high of the inverse of this status, Receive fault, is reflected in MDIO register 3.8.10.
- block_lock:
 - Indicates the state of the block_lock variable. This status is reflected in MDIO register 3.32.0. A latch low view of this status is reflected in MDIO register 3.33.15.
- hi_rfer:
 - Indicates the state of the hi_rfer variable. This status is reflected in MDIO register 3.32.1. A latch high view of this status is reflected in MDIO register 3.33.14.
- ➔ PCS_status=OK if (block_lock=true) and (hi_rfer=false)

block_lock definition

- On Page 74, line 23 to 30:
- When the receive channel is **in training mode**, the PCS Synchronization process It attains frame and block synchronization based on the PMA training frames and conveys received blocks to the PCS Receive process. The PMA training sequence includes 1 bit pattern every 180 PAM2 symbols, which is aligned with the PCS Partial RS frame boundary, as well as an InfoField which is inserted in the 15th PCS Partial RS frame. **When the PCS Synchronization process is synchronized to the RS frame boundary using this pattern, block_lock is asserted.**
- ➔ Block_lock will be set to true in PAM2 training after RS FEC frame alignment is done.

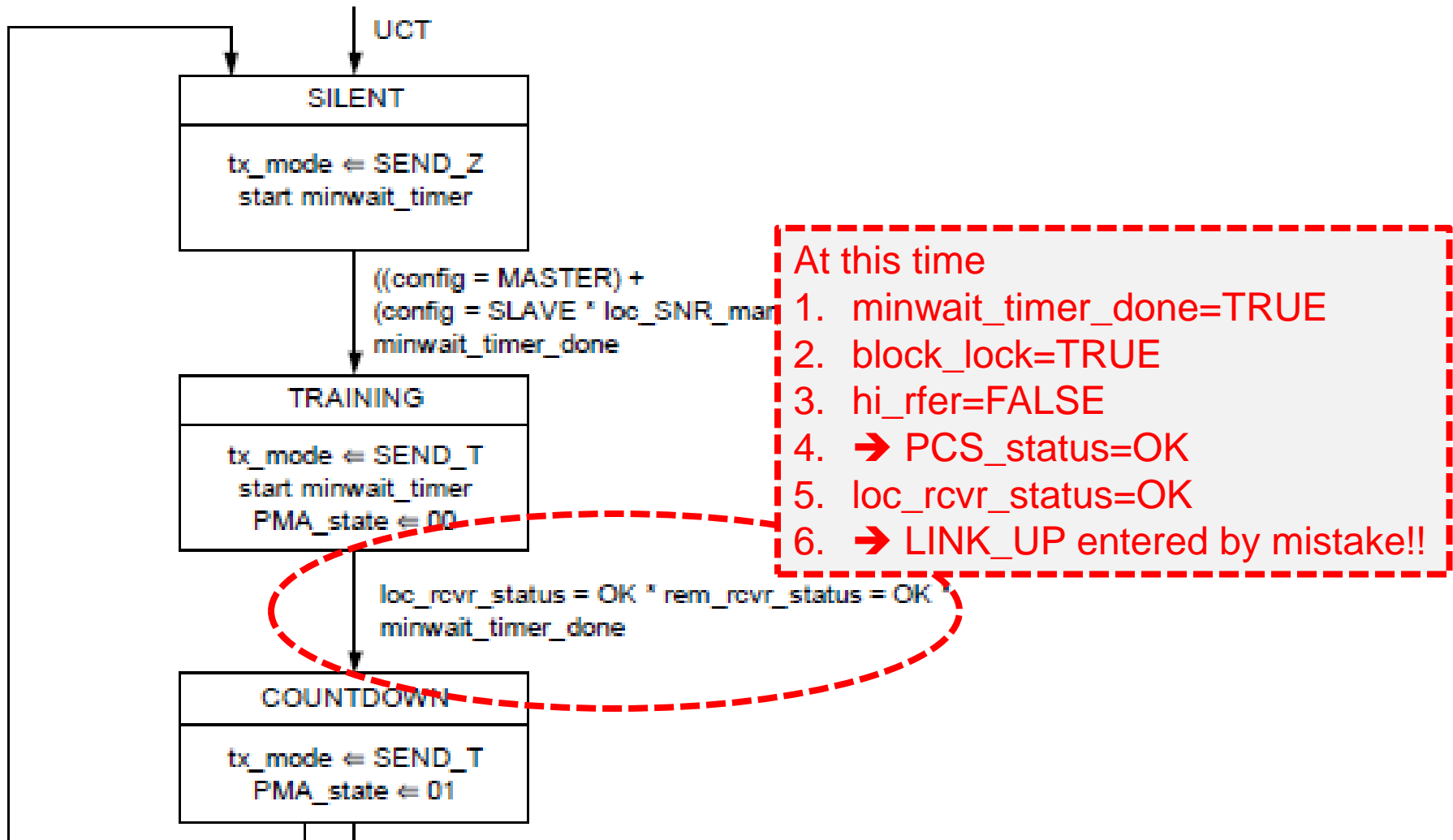
hi_rfer definition

- On page 74, line 16 to 18:
 - When the PCS Synchronization process has obtained synchronization, the RS frame error rate (RFER) monitor process monitors the signal quality asserting hi_rfer if excessive RS frame errors are detected (RS parity error).
- On page 79, line 43 to 45:
- hi_rfer
 - Boolean variable which is asserted true when the rfer_cnt exceeds RFER_CNT_LIMIT indicating a bit error ratio $> 4 \times 10^{-4}$.
- ➔ In training mode, there is no FEC errors, and hence hi_rfer=false.

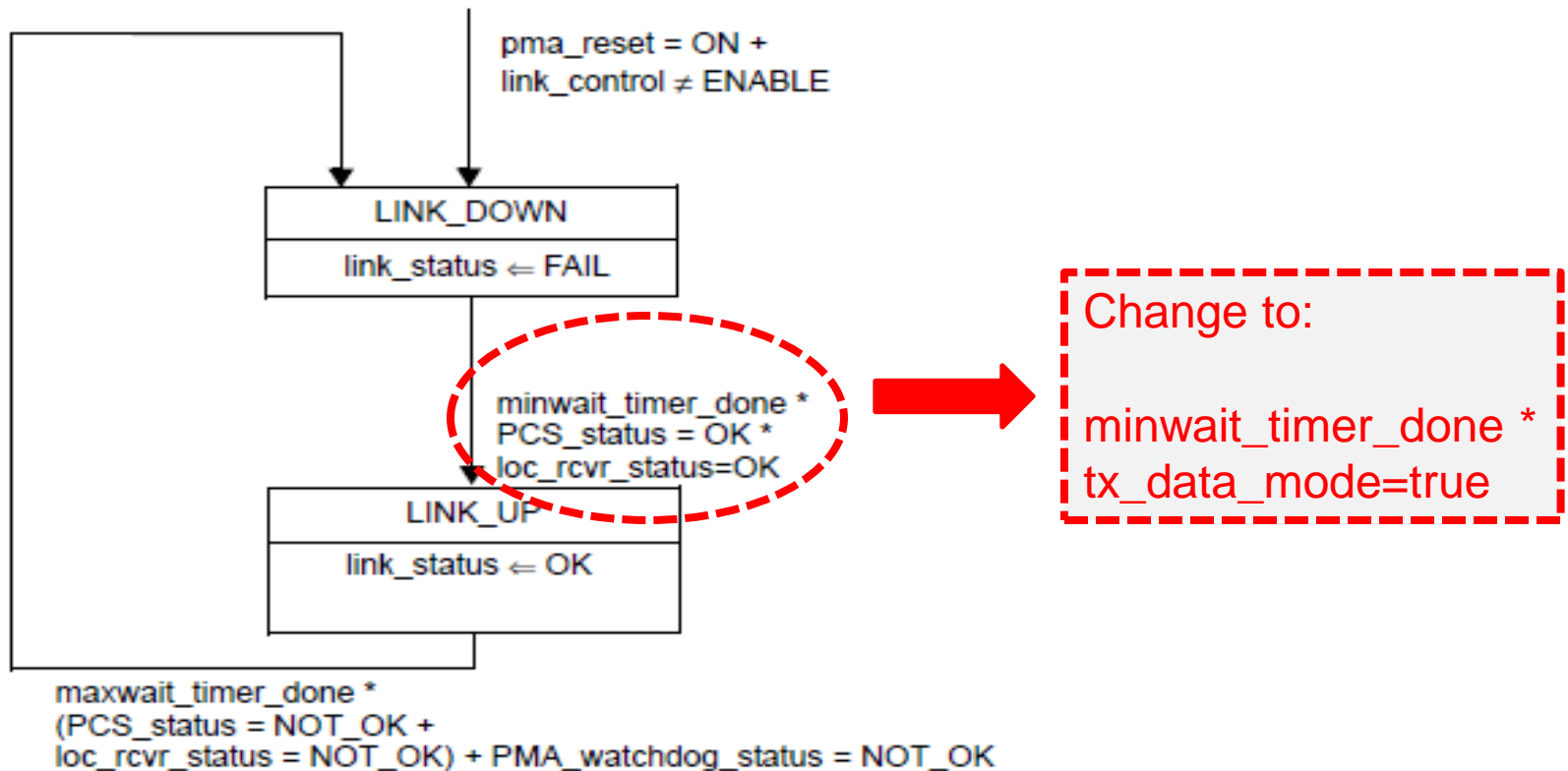
Issue with Figure 97-23 Link Monitor

- The condition from “LINK_DOWN” state to “LINK_UP” state is shown as:
 - `minwait_timer_done * PCS_status=OK *
loc_rcvr_status=OK.`
- However, this condition will become true during the PAM2 training, before entering the data mode, resulting in “LINK_UP” declared too early.

Part of PAM2 Training Sequence



Remedy for Figure 97-23



NOTE 1—`maxwait_timer` is started in PHY Control state diagram (see Figure 97–22).
NOTE 2—The variables `link_control` and `link_status` are designated as `link_control_1GigT1` and `link_status_1GigT1`, respectively, by the Auto-Negotiation Arbitration state diagram (Figure 98–14) if the optional Auto-Negotiation function is implemented.

Figure 97–23—Link Monitor state diagram

tx_data_mode

- On page 80, line 26:
- tx_data_mode
 - Set true when tx_mode = SEND_N, otherwise false.
- This variable is only set to TRUE after tx_mode=SEND_N → no ambiguity.

Topic #2:

A proposal for remote receiver status
in data mode

PCS and PMA Interface

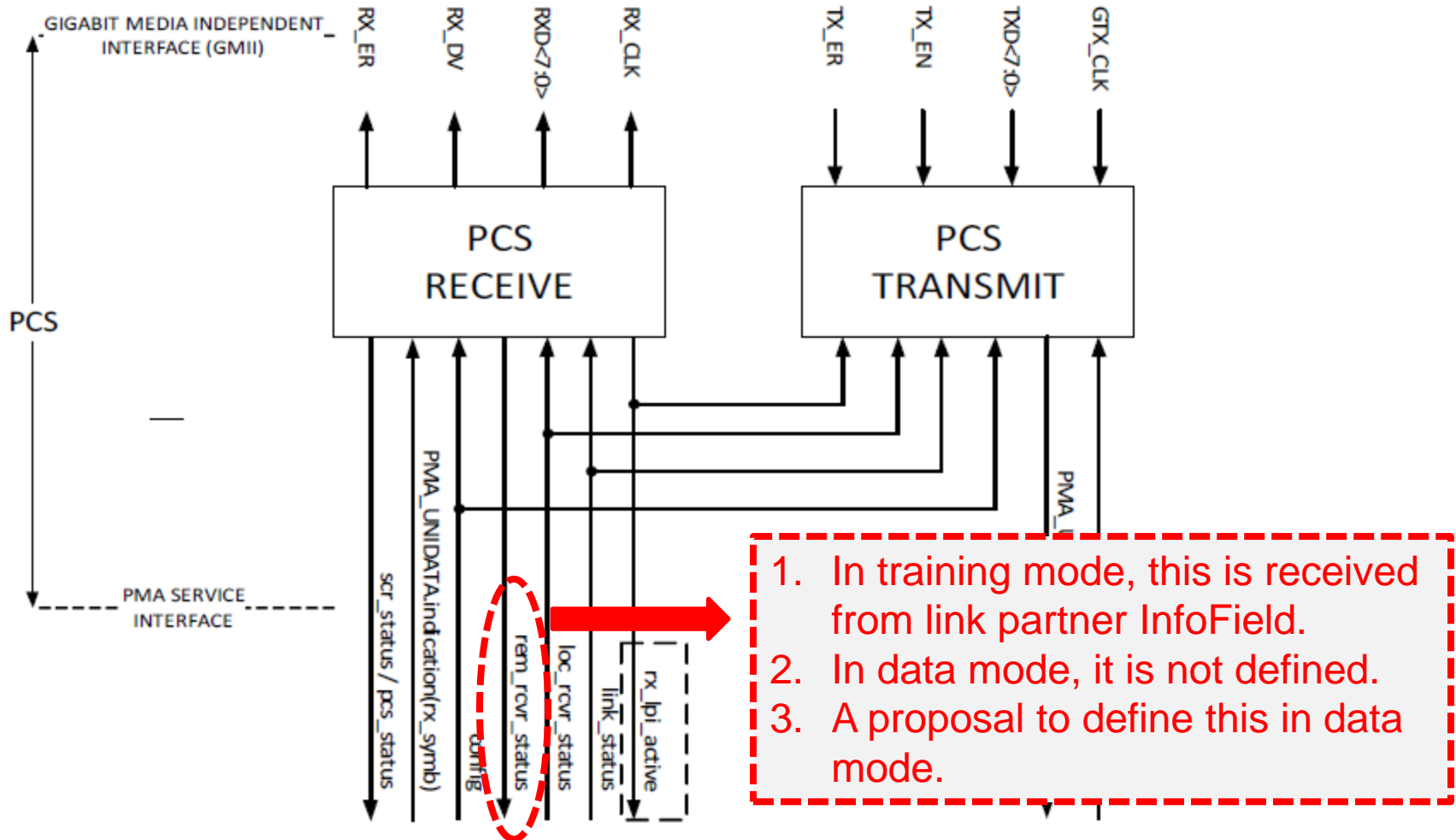


Figure 97-4—PCS reference diagram

Data Mode Variables

- In 1000BASE-T1, rem_rcvr_status is only defined in PAM2 training mode.
- Define new variable “loc_data_ready”:
 - loc_data_ready=OK, if (loc_rcvr_status=OK) and (pcs_status=OK)
 - loc_data_ready=NOT_OK, otherwise
- The “loc_data_ready” variable value will be sent to link partner via PCS 80/81B IPG control words (see next slide).
- The “rem_data_ready” variable is based on the received link partner “loc_data_ready” variable (see next slide).

Modified Table 97-1

NEW

Control Code[0:2]	GMII Transmit	GMII Receive
000	Normal Inter-Frame with <code>loc_data_ready=NOT_OK</code>	Normal Inter-Frame with <code>rem_data_ready=NOT_OK</code>
001	Transmit Error Propagation	Data Reception Error
010	Normal Inter-Frame with <code>loc_data_ready=OK</code>	Normal Inter-Frame with <code>rem_data_ready=OK</code>
101	Assert Low Power Idle	Assert Low Power Idle
other	Reserved	Reserved

**Additional
condition**

- Set data mode variable “`rem_data_ready`”:
 - OK if the control code “010” is received
 - NOT_OK if the control code “000” is received

Text and Figure Changes

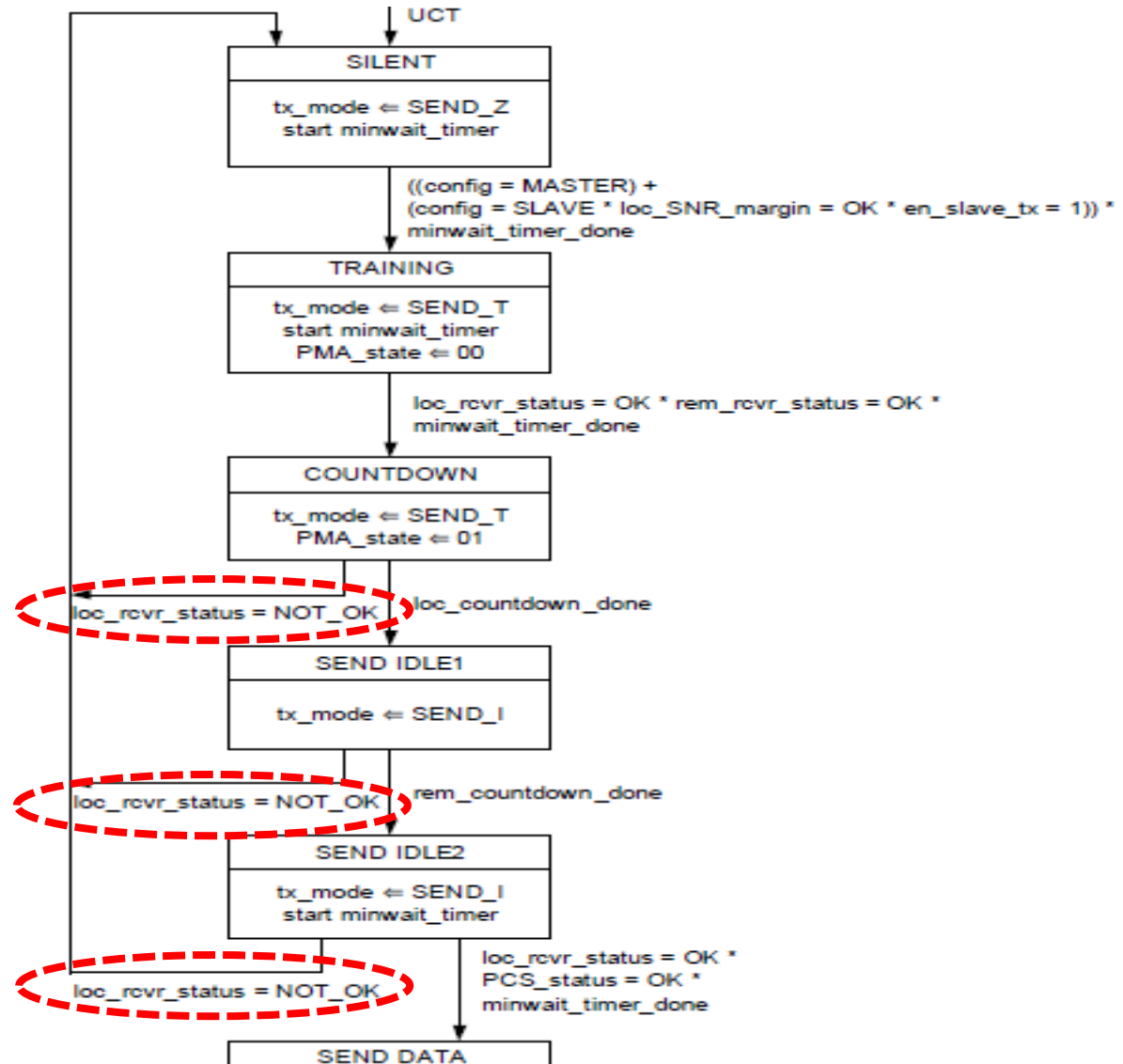
- Will provide a separate presentation to show the text and Figure changes.
- Basically:
 - Define `loc_data_ready` and `rem_data_ready`
 - Change “`rem_rx_status`” to “`rem_rx_status/rem_data_ready`” when appropriate

Topic #3:

Concerns with Figure 97-22 PHY
Control state diagram

Potential Concerns

What happens when link partner drops back to SILENT state during training??



Some Analysis

- Link partner drops back to SILENT state
 - Link partner tx_mode <= SEND_Z
 - Last for at least minwait_timer ≈ 1msec
- The local receiver will receive all 0.
- We want the local PHY to jump to SILENT state quickly in order to attempt retrain.
- Based on D1.4, the local receiver must set loc_rcvr_status=NOT_OK under this condition.
- However ...

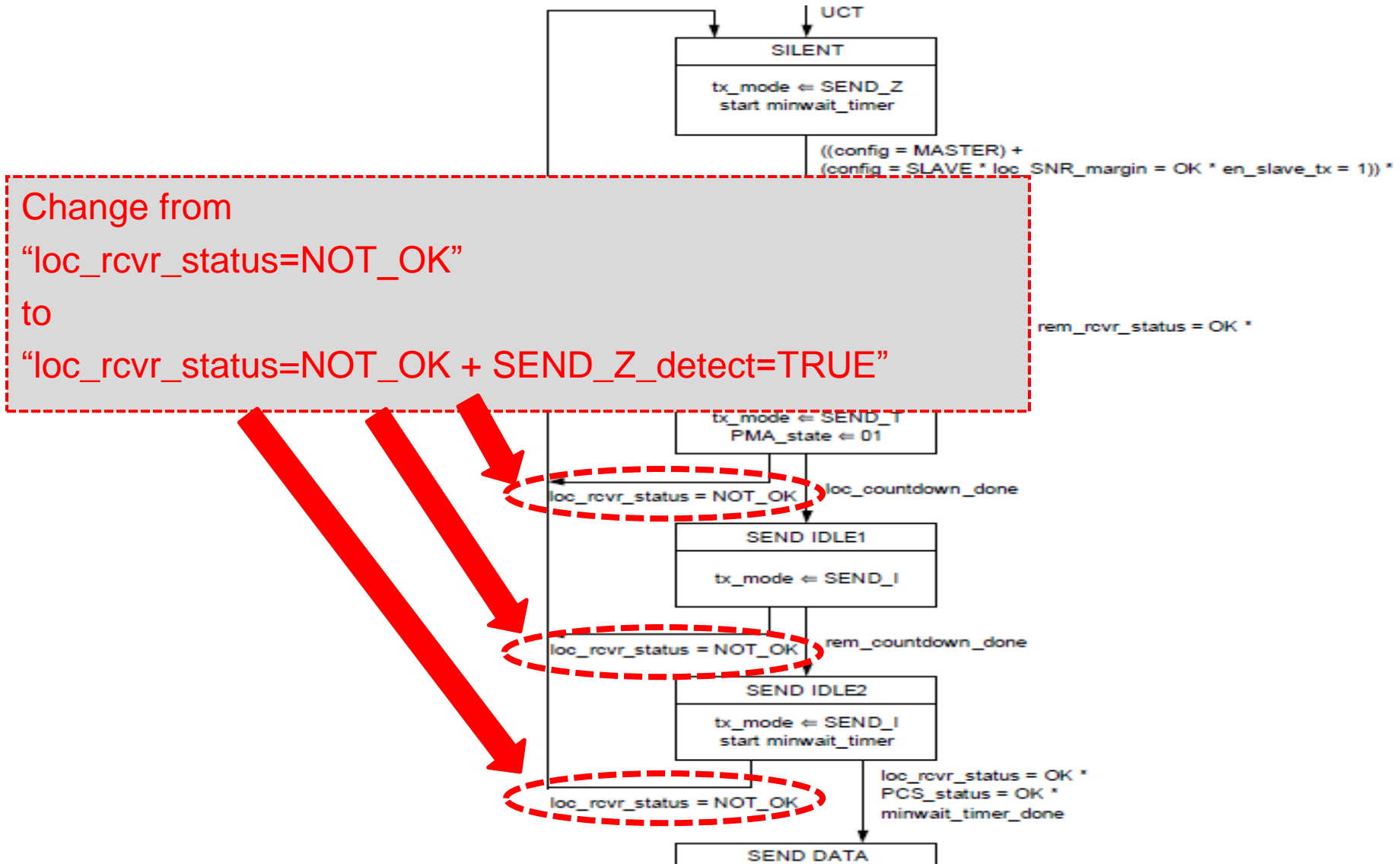
loc_rcvr_status

- Definition for loc_rcvr_status is intentionally vague in current D1.4.
- Based on 97.2.2.7, page 61, line 10 to 13
 - The criterion for setting the parameter loc_rcvr_status is left to the implementor. It can be based, for example, on observing the mean-square error at the decision point of the receiver and detecting errors during reception of symbol stream.
- In order to improve interoperability, a more specific condition is desirable.
- → Add “SEND_Z_detect” in the branch conditions
 - Note: The variable PMA_watchdog_status also tries to detect link partner Tx silence, but they might be subject to different requirements.

Proposed SEND_Z_detect definition

- New state variable for 97.4.4.1
- SEND_Z_detect
 - This variable is used during the COUNTDOWN, SEND_IDLE1, and SEND_IDLE2 states. It detects when the remote link partner is transmitting SEND_Z.
- Values:
 - TRUE: Remote link partner is transmitting SEND_Z.
 - FALSE: Remote link partner is not transmitting SEND_Z.

Proposed Improvements

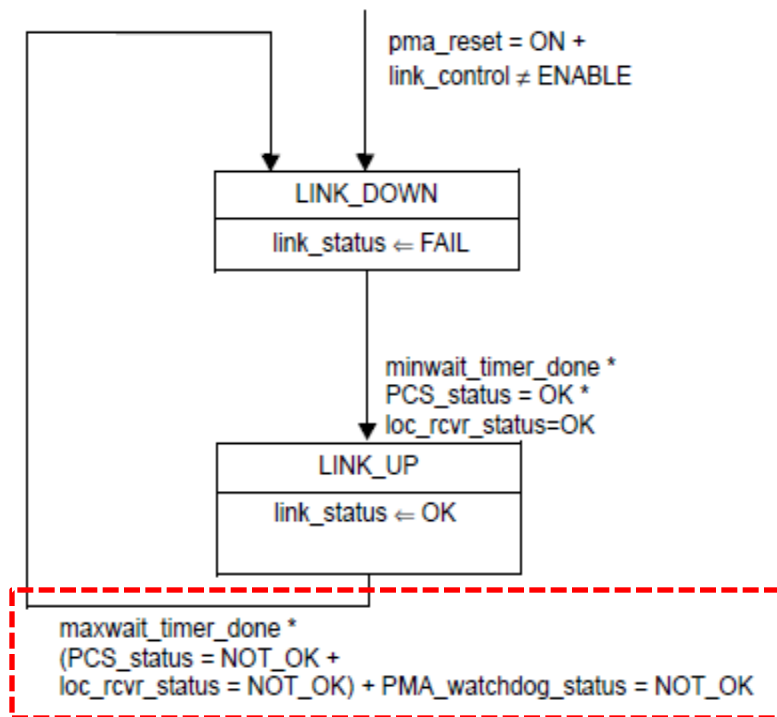


Topic #4:

Add LPI REFRESH Failure to Link
Monitor state diagram

What is the issue?

- Figure supersedes text descriptions.



- 97.4.2.7 Refresh Monitor function

- A 1000BASE-T1 PHY supporting the EEE capability shall implement the Refresh monitor function. The Refresh monitor operates when the PHY is in the LPI receive mode. **The receiver shall force a retrain if Refresh is unreliably detected within a moving window of 50 Q/R cycles (4.32 ms).**

NOTE 1—maxwait_timer is started in PHY Control state diagram (see Figure 97–22).
NOTE 2—The variables link_control and link_status are designated as link_control_1GigT1 and link_status_1GigT1, respectively, by the Auto-Negotiation Arbitration state diagram (Figure 98–14) if the optional Auto-Negotiation function is implemented.

Add this condition into Figure 97-23.

Figure 97–23—Link Monitor state diagram

Background

Three ways to leave LPI mode when receiving unreliable refreshes:

1. Detects PMA_watchdog_status during a refresh
 - Go to LINK_DOWN state within 90 us.

2. Refresh monitor fails
 - Restart auto-negotiation (if auto-negotiation is enabled) or synchronization (if auto-negotiation is disabled).
 - Criteria for unreliable Refresh is not specified but detection window is specified to be 4.32 ms.

3. Send SNR<1:0> = 01
 - Local and remote LP exit LPI within 10.8 us
 - If Link Monitor also has a problem after exiting LPI mode restart auto-negotiation (if auto-negotiation is enabled) or synchronization (if auto-negotiation is disabled).

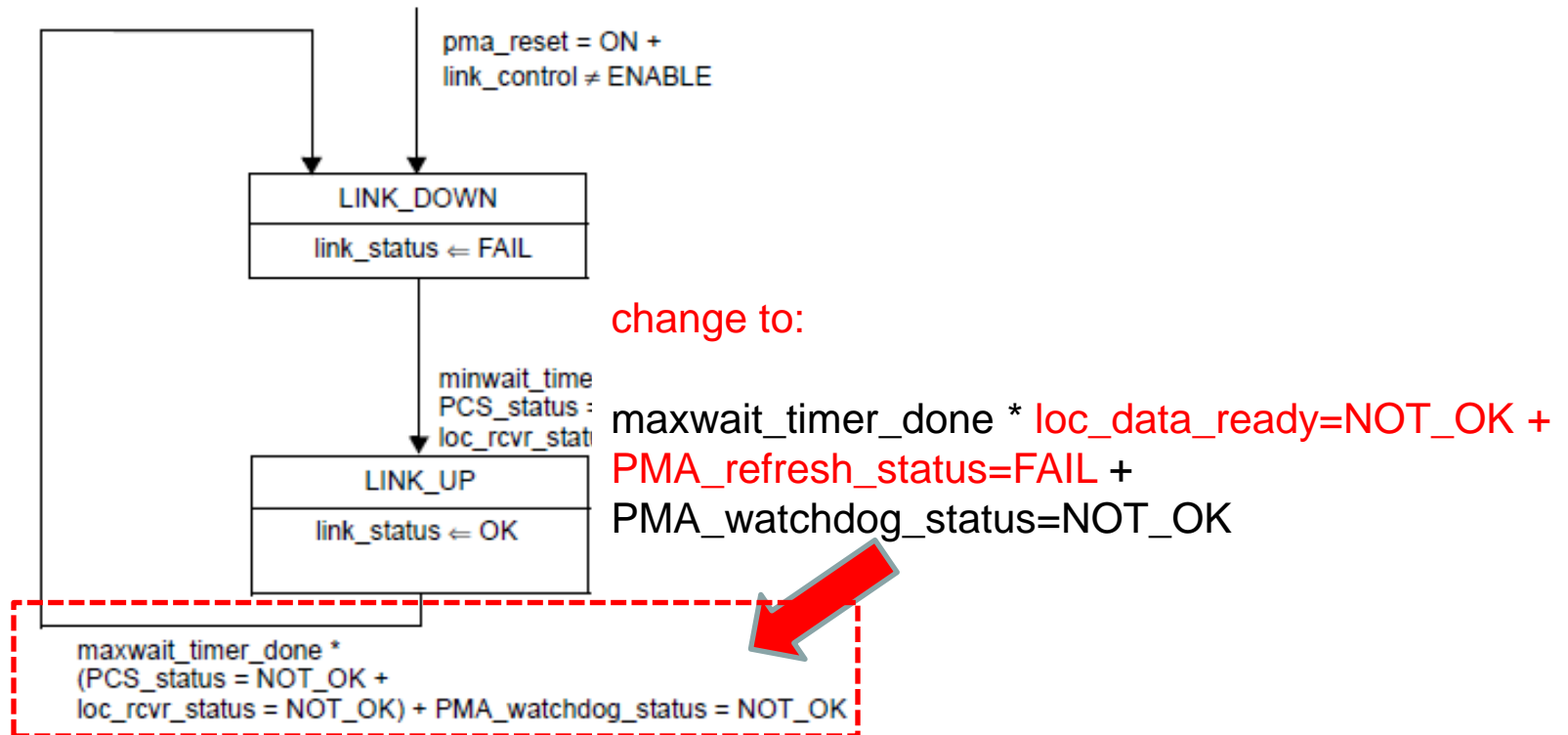
Link failure modification proposal

- Change 97.4.2.7. Page 94, line 4-5, from
 - "The receiver shall force a retrain if Refresh is unreliably detected within a moving window of 50 Q/R cycles (4.32 ms)."
- To
 - "If Refresh is not reliably detected within a moving window of 50 Q/R cycles (4.32 ms), the refresh monitor shall set PMA_refresh_status to FAIL. This will set link_status to FAIL. Subsequently the PHY will restart auto-negotiation (if auto-negotiation is enabled) or synchronization (if auto-negotiation is disabled)."

New state variable in 97.4.4.1

- PMA_refresh_status
 - Variable indicating the status of the Refresh Monitor.
- Values:
 - OK: Refresh is detected reliably.
 - FAIL: Refresh is not detected reliably.

Changes to Figure 97-23



NOTE 1—maxwait_timer is started in PHY Control state diagram (see Figure 97-22).
 NOTE 2—The variables link_control and link_status are designated as link_control_1GigT1 and link_status_1GigT1, respectively, by the Auto-Negotiation Arbitration state diagram (Figure 98-14) if the optional Auto-Negotiation function is implemented.

Figure 97-23—Link Monitor state diagram

Topic #5:
PHY Control state diagram –
entering SEND_DATA

Transition into SEND_DATA state

- In Figure 97-22, when entering SEND_DATA state, the decision is purely based on local receiver and PCS status.
- Once in SEND_DATA mode the local PHY will set link_status<=OK and MAC will start sending data out.
- However at this time the link partner receiver may not be ready yet, for example due to noise events. Under this condition the data packets sent to the link partner will be lost.

Proposed Changes

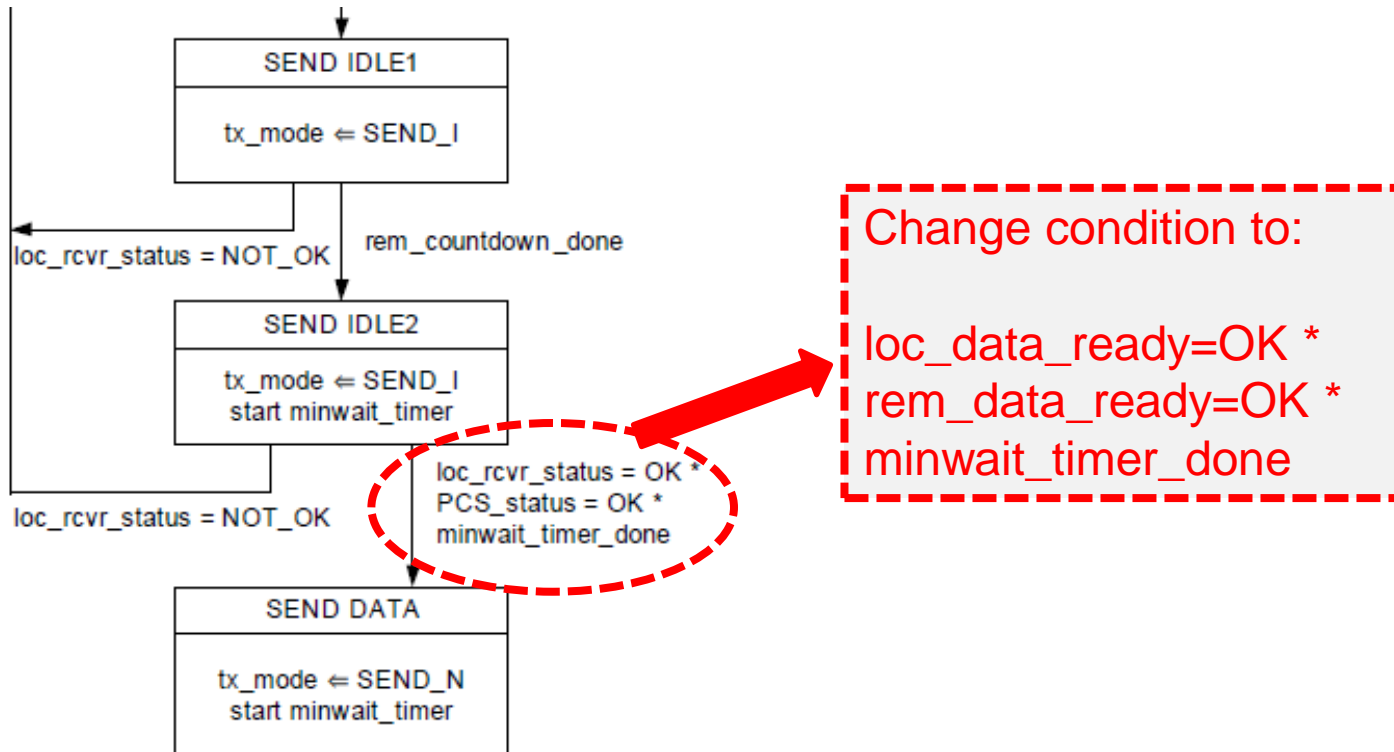


Figure 97–22—PHY Control state diagram

Topic #6:
PHY Control state diagram –
exiting SEND_DATA

Comment #569 from Brett

- CI 97 SC 97.4.5.1 P 97 L 47 # 569
- It is possible for one device to enter the SEND DATA state but the other device to return to the SILENT state. There is no path from SEND DATA to SILENT, so the devices must wait for the link_fail_inhibit_timer to expire and then return to autoneg or SEND_S. If a path is added from SEND DATA to SILENT then the two devices may attempt to retrain with the remaining time of the link_fail_inhibit_timer.
- Suggested Remedy
 - add path from SEND DATA to SILENT with condition "loc_rcvr_status = NOT_OK + PCS_status = NOT_OK"
- → We support this idea

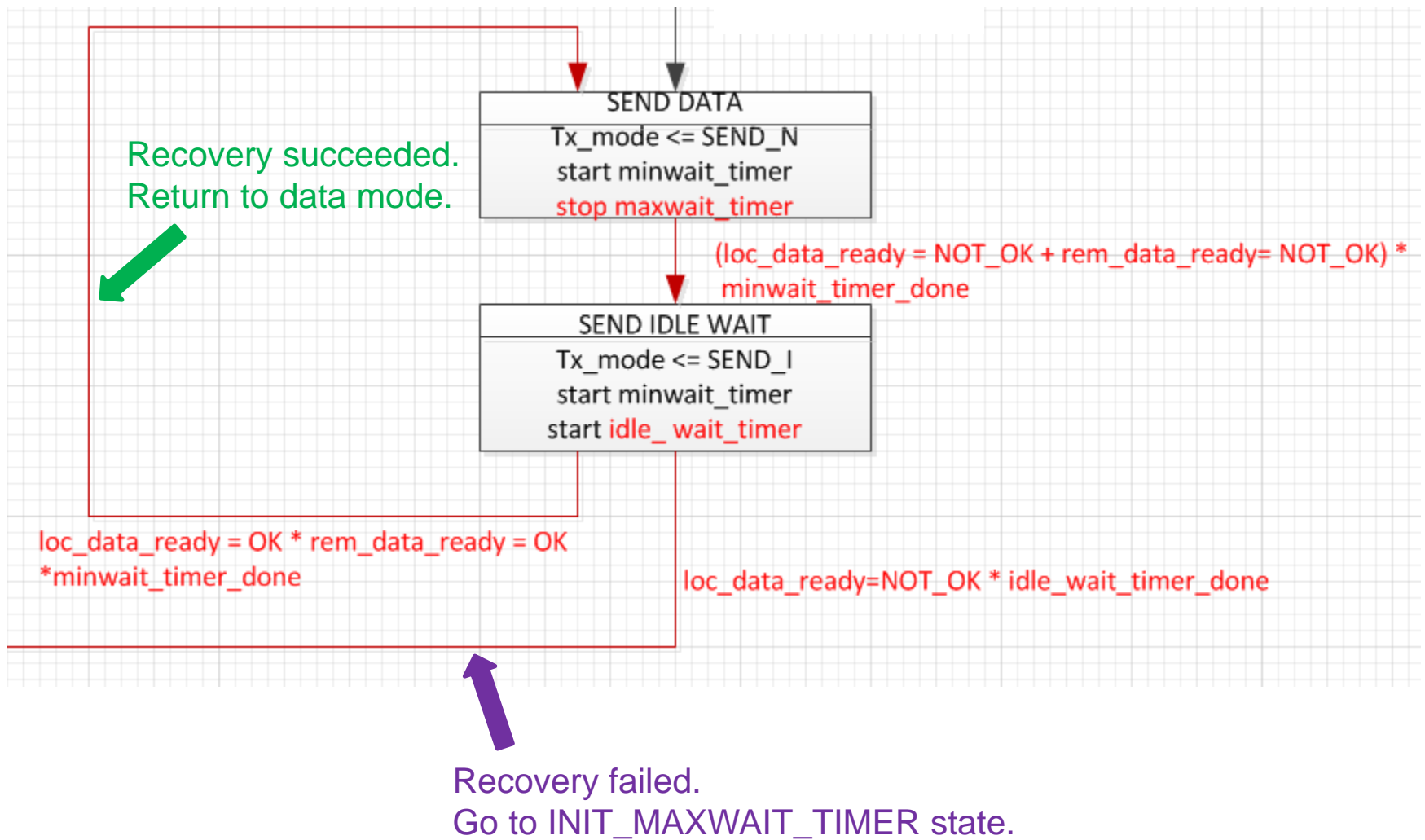
Further Improvements

- What causes PHY link drops?
 - Power or hardware failures
 - Disconnected cable
 - User command/management
 - PHY receiver diverged
 - Variations in noise environment, including EMI, burst noises, etc.
 - Others
- When the link drop is caused by varying noise environment, PHY receiver has a good chance of recovering without going back to training mode.

Further Improvements (cont.)

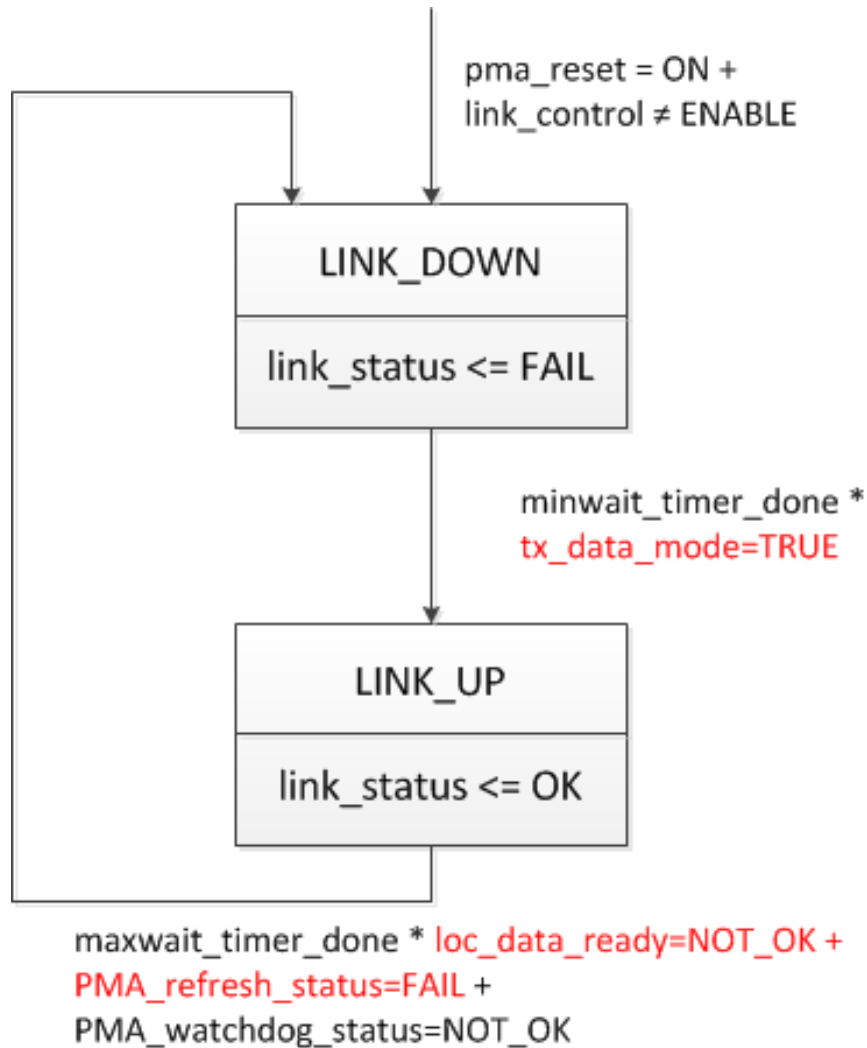
- Before jumping back to half-duplex training mode, stay in 3B2T for “some time”, then check status again.
 - If PHY receiver recovers, then stay in data mode
 - If PHY receiver still can not recover, then do a full retrain
 - During this time, indicate loc_data_ready=NOT_OK to link partner
- Define “idle_wait_timer”
 - About 10msec should be sufficient
- Some changes to Figure 97-22

Changes to Figure 97-22

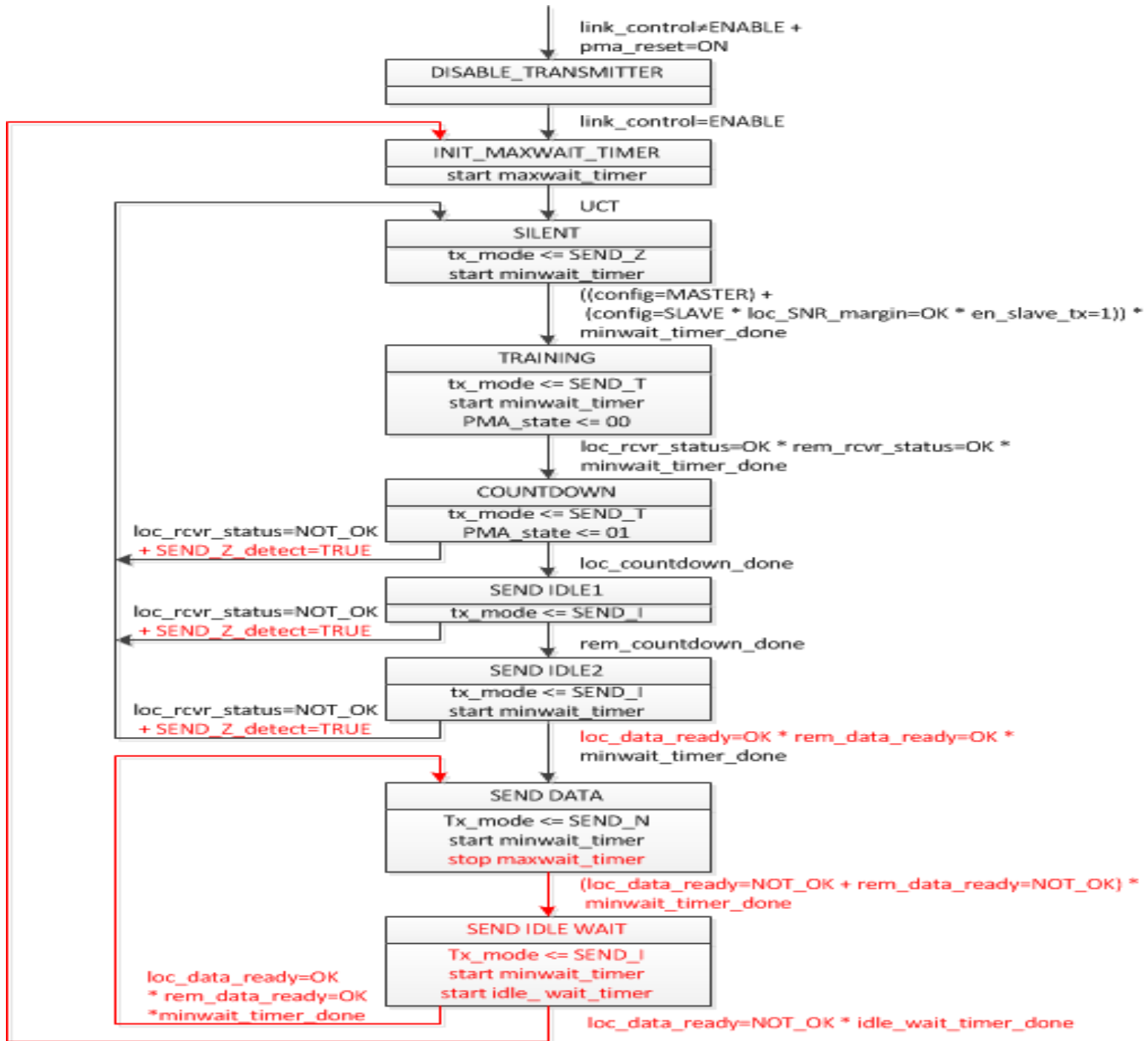


Proposed Figure 97-23 and Figure 97-22 with all changes

New Figure 97-23

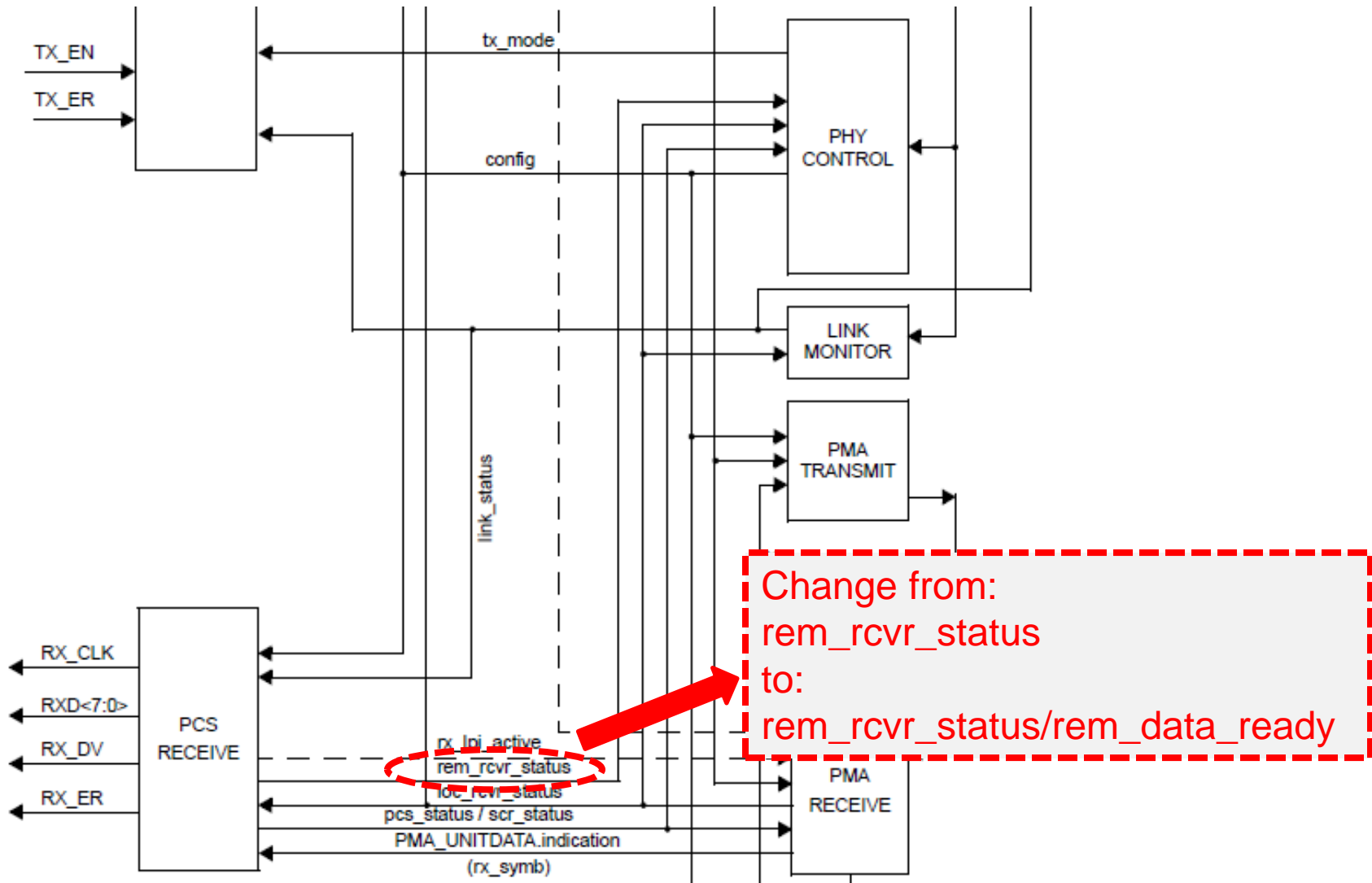


New Figure
97-22



Supplement Materials

Changes to Figure 97-2



Change from:
rem_rcvr_status
to:
rem_rcvr_status/rem_data_ready

Changes to Figure 97-4

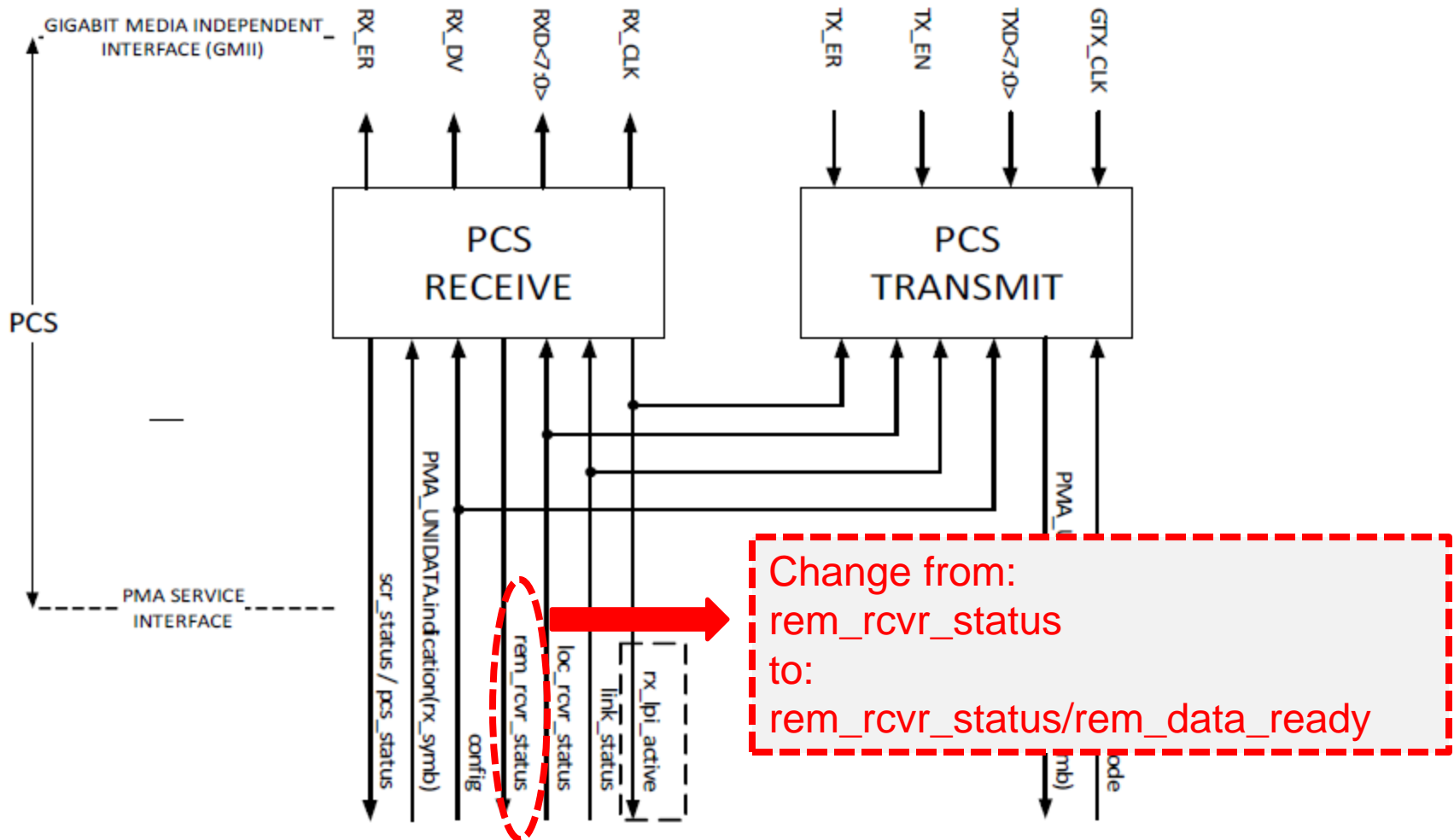


Figure 97-4—PCS reference diagram

Changes to Figure 97-16

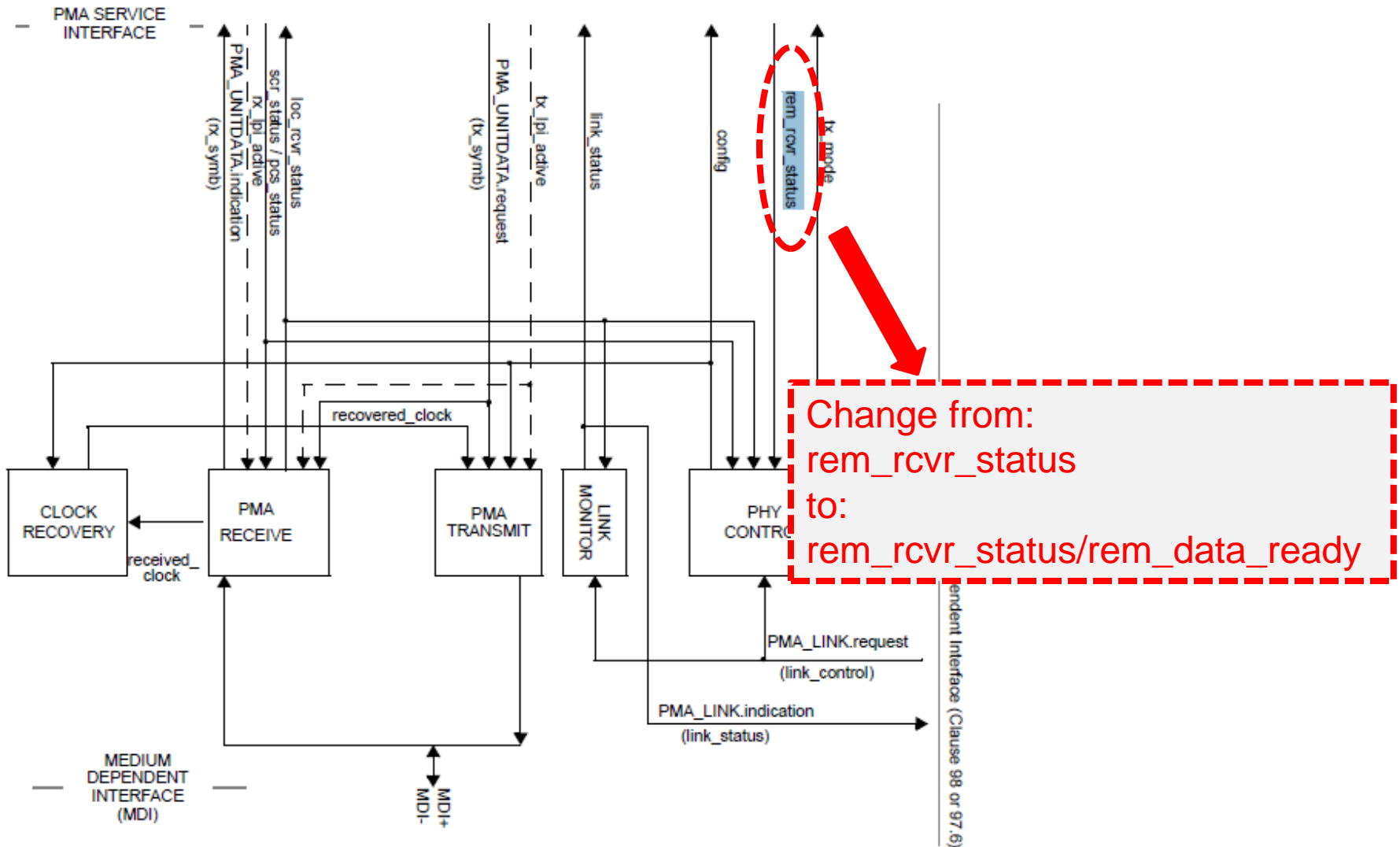


Figure 97-16—PMA reference diagram