PMD Control Timing

Kent Lusted, Intel

P802.3bj Draft 2.1 Comment #94

Cl **92** *SC* **92.7.12** *P* **193** *L* **18** # 94

Ran, Adee Intel

Comment Type **T** *Comment Status* **A**

The response time requirement is dependent on the status of frame_lock_i which may be difficult to verify (e.g. if the MDIO interface is unavailable) and synchronize with a captured waveform. In addition, it is not available to the link partner.

It is relatively easy to make the lane frame lock state available as part of the status report field. This information would be very useful in analyzing link training issues and thus promote interoperability.

Comment applies to clauses 93 and 94 as well.

SuggestedRemedy

In clauses 92 and 93, assign cell 14 of the status report field (currently reserved) to represent the value of the PMD status variable frame_lock_i.

In clause 94, use cell 7 of the status report field instead of cell 14 (14 is already assigned, 7 is currently reserved).

Editorial license granted.

Response *Response Status* **C**

ACCEPT IN PRINCIPLE.

After initial frame lock the the response time should be 2 ms regardless of whether the receiver loses frame lock or not.

Replace:
"In addition to the coefficient update process specified in 72.6.10.2.5, when frame_lock_i is TRUE for lane i
(where i represents the lane number in the range 0 to 3), the period from receiving a new request to responding
to that request shall be less than 2 ms."

With:
In addition to the coefficient update process specified in 72.6.10.2.5, after responding to the first request after training begins, the period from receiving a new request to responding to that request shall be less than 2 ms.

Comment #94 as Implemented in P802.3bj Draft 2.2

92.7.12 PMD control function

Each lane of the 100GBASE-CR4 PMD shall use the same control function as 10GBASE-KR, as defined in 72.6.10.

The training frame structure used by the 100GBASE-CR4 PMD control function shall be as defined in 72.6.10 with the exception that 25.78125 GBd symbols replace 10.3125 GBd symbols and 100GBASE-CR4 UI replace 10GBASE-KR UI.

In addition to the coefficient update process specified in 72.6.10.2.5, after responding to the first request after training begins, the period from receiving a new request to responding to that request shall be less than 2 ms. The start of the period is the frame marker of the training frame with the new request and the end of the period is the frame marker of the training frame with the corresponding response. A new request occurs when the coefficient update field is different from the coefficient field in the preceding frame. The response occurs when the coefficient status report field is updated to indicated that the corresponding action is complete.

- Change also made in Cl 93.7.12 and Cl 94.3.10.7.5

Offline Consensus Discussion Items

- Is the 2 msec response to a coef update from a partner enough time?
  - i.e. if the local RX somehow loses frame_lock in TRAIN_LOCAL or TRAIN_REMOTE states?
- Is it necessary to bound the time from the start of the training process to the first coef update request response?
  - There is no limit today.

Training State Diagram Reference

- Frame_lock could be deasserted by a local RX in the TRAIN_LOCAL and TRAIN_REMOTE states
- Two timers in Fig 72-5:
  - Wait_timer governs exit of PMD control to data mode
  - Max_wait_timer is a watchdog timer to completion of PMD control

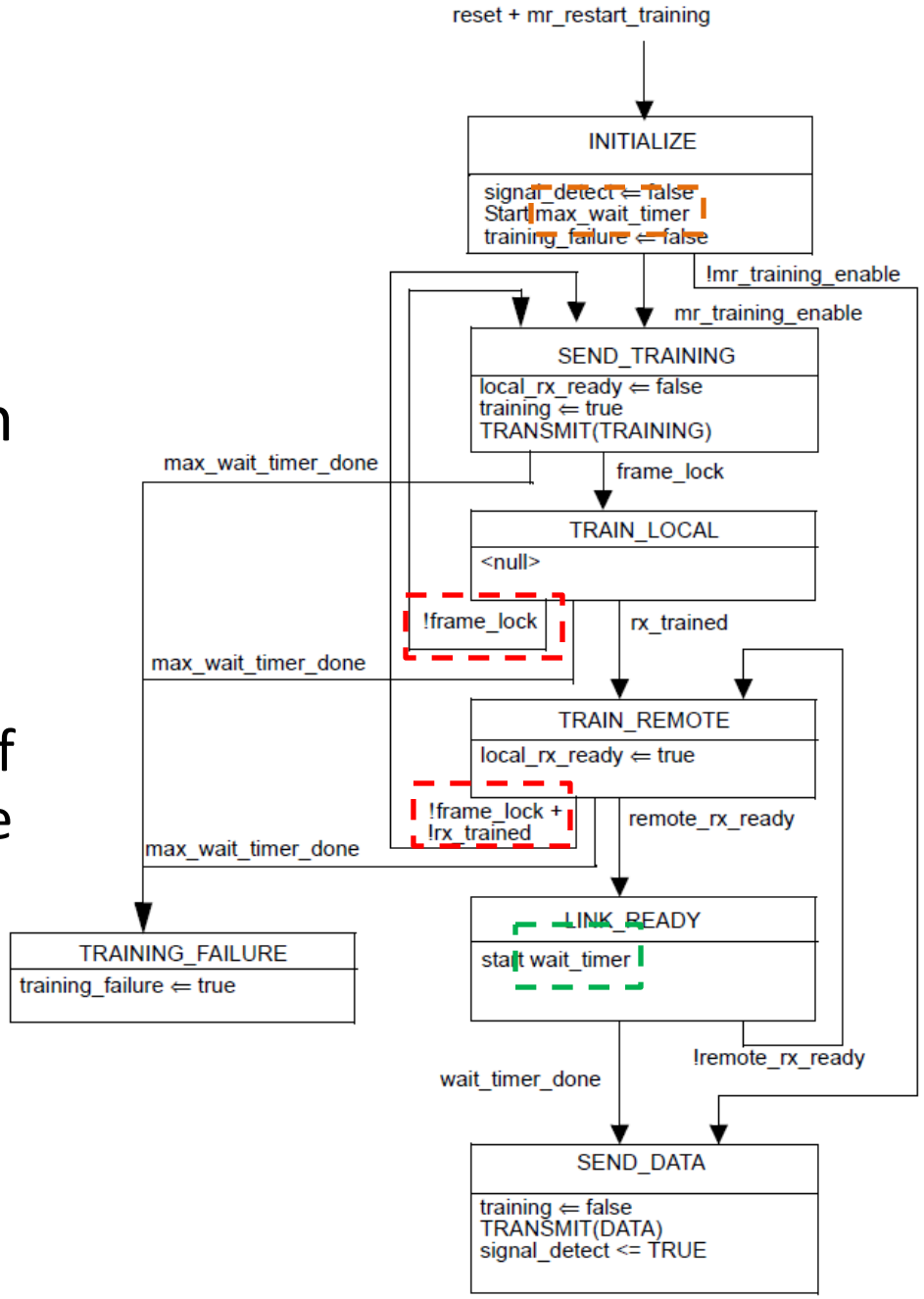

Figure 72–5—Training state diagram

2 ms Response

- In D2.2, what happens if the 2 ms response is not possible due to something?
  - i.e. Remote TX is broken in a way that a response is not possible
- D2.2 has the 2 ms response as normative with PICS statement
- What should the spec require to fix it?
- Possible solutions:
  - ~~Do nothing (don't change the spec in any way)~~
  - ~~Remove 2 ms normative requirement~~
  - Revert to draft 2.1 text
    - Does not prevent parts from being compliant
    - No easy way to verify response time against the normative specification
  - Revert to draft 2.1 text but add frame_lock indication in the control channel
    - Original comment #94
  - ~~Create a bailout mechanism for the aforementioned case~~
- Consensus to revert to draft 2.1 text

2 ms = ~12k 100GBASE-CR4/100GBASE-KR4 training frames or ~1.8k 100GBASE-KP4 frames
* 1x 100GBASE-CR4/KR4 frame = ~162 nsec
* 1x 100GBASE-KP4 frame = ~1.1 usec

Bounding the Start

- Problem statement: It is possible that the max_wait_timer budget is consumed before the local RX gets a response to PMD control requests.
  - i.e., a partner RX takes 499 ms of the max_wait_timer budget (500 ms +/-1%) to respond to the first request
- Possible solutions:
  - ~~Don't change spec at all~~
  - Implement 2 ms response with indication (original comment #94 solution) and don't fix this
  - Implement 2m response without indication (revert to draft 2.1 text) and don't fix this
  - ~~Require a partner to respond within X ms to the first request after training begins using a new solution~~
- Consensus not to fix this issue