800GBASE-LR1 Coherent APSU Continued

Matt Brown, Qualcomm
Jeff Slavick, Broadcom
Gary Nicholl, Cisco
Adee Ran, Cisco
Marco Mascitto, Nokia
Leon Bruckman, Nvidia
Eric Maniloff, Ciena

Introduction

- Task force discussions on Monday led to apparent preference for the encoded PRBS31 sequence to indicate RTS state
- We also need to think about how we would map the signaling using PRBS31 and inverse into the APSU state diagrams.
- But we had not thought to consider squelch as an option.
  - Perhaps we should give this a little thought.
- These slides will explore both.
- This slide package is not a recommendation.

Straw polls

SP2: For the LR1 PHY, the APSU approach, outlined in mascitto_3dj_01a_2603, that I would prefer to adopt is:

A) Option A: PRBS31(Pages 6-7)
B) Option B: O2 Training Frame (Pages 8-9)
C) Option C: Pilot Sequences (Page 10)

(Chicago Rules) Results: A: 39, B: 18, C: 32

Straw Poll #4 – LR1 APSU

For the LR1 PHY, the APSU approach, outlined in mascitto_3dj_01a_2603, that I would prefer to adopt is:

A> PRBS31(Pages 6-7)

B> Pilot Sequences (Page 10)

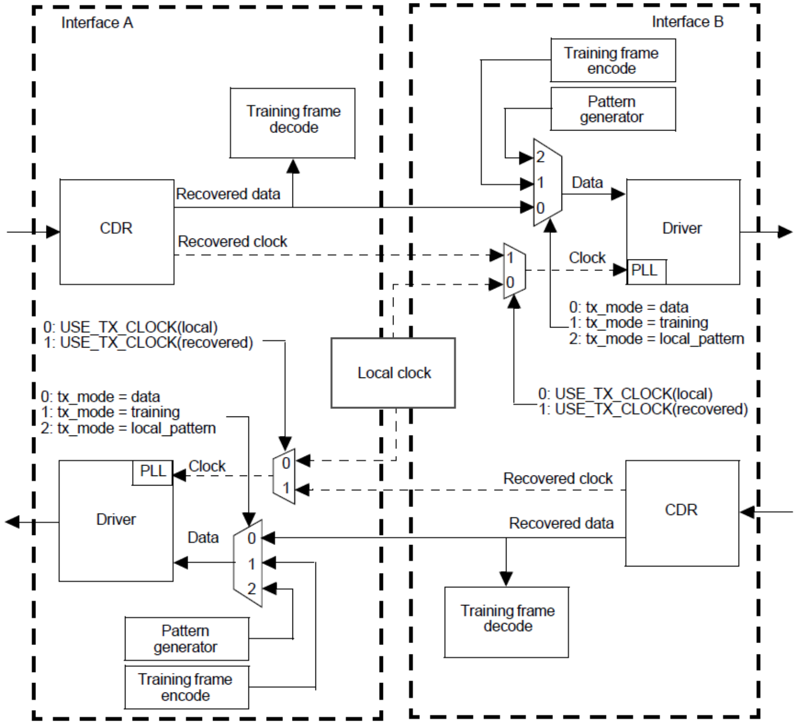
(Choose one) Results: A: 37, B: 28

Based on these two straw polls for 800GBASE-RL1 the room seems to be most supportive of using PRBS31 and inverse for signaling the RTS state.

Retimer model (clocking perspective)

Regardless, transition to/from mission clock and mission data is based on the retimer model in Figure 178B-13. This transition requires careful coordination accommodated by RTS signaling and state diagrams.

AUI ISL

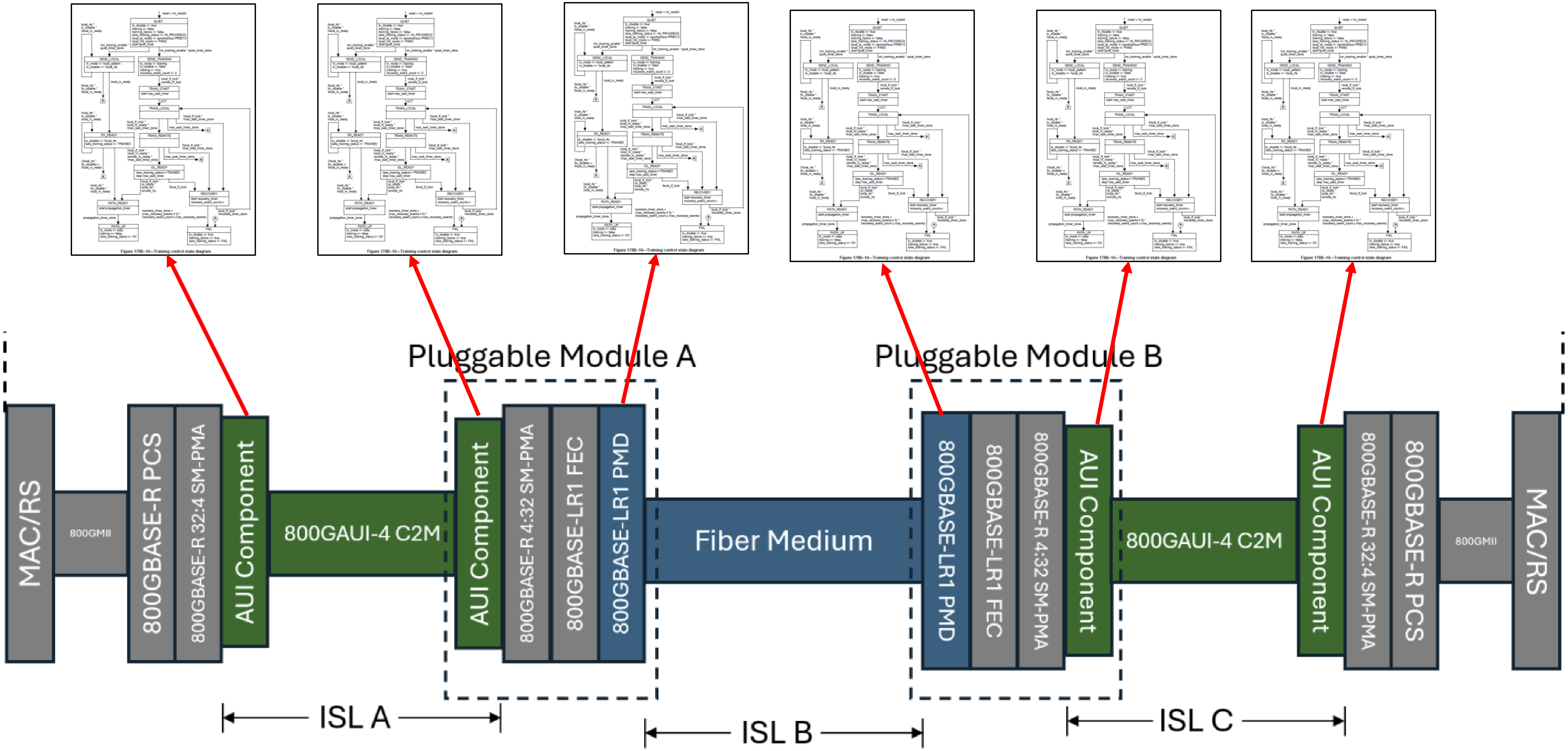


Coherent ISL

NOTE—The dashed lines represent clocking connections that are needed within a retimer for APSU.

Figure 178B–13—Retimer reference model

APSU state machines required at each of the 6 interfaces



800GBASE-LR1

Using PRBS31 and PRBS31 inverted (no squelch)

General

- Use the local_pattern path on the state diagram.
- Some state machine variables need to be redefined.
- Note state diagram is required for careful switchover of clocks from local clock to adjacent clock, then switch to data.
- PRBS31 encoded by Inner FEC used to convey not RTS (or equivalently tx_disable = true).
  - Permits currently defined pattern checker to be used to detect.
- Inverted PRBS31 encoded by Inner FEC used to convey RTS.
  - This is an intermediate pattern sent while waiting for reliable data from the adjacent interface.

local_rx_ready

local_rx_ready

Boolean variable that is set to true when the receiver on a lane of the interface has determined that the peer interface transmitter is transmitting a PAM4 signal, that the remote transmit has been optimized if mr_training_enable is true, the local receive equalizers have been optimized, and that no further adjustments are required for normal data transmission.

For 800GBASE-LR1, it is not waiting for a PAM4 signal.

Need to reword to be universal to both PAM4 and the coherent DP-16QAM.

So change "PAM4 signal" to "valid signal with appropriate modulation format" or similar.

This change is also needed for the squelch as signaling method.

State machine variable tx_disable

tx_disable

Boolean variable that controls the transmitter output on the lane. It is set by the training control state diagram (Figure 178B–10). When it is true, the transmitter output on the lane is disabled. When it is false, tx_mode controls the content of the transmitter output on the lane.

For 800GBASE-LR1, the behavior in response to tx_disable state is as follows:
tx_disable is true: transmit PRBS31 pattern encoded by the Inner FEC
tx_disable is false and tx_mode is local_pattern: transmit inverted PRBS31 pattern encoded by the Inner FEC
tx_disable is false and tx_mode is data: transmit data from the adjacent encoded by the Inner FEC

Potential implementation of this is covered on later slides.

State machine variable remote_rts

remote_rts

Boolean variable that indicates the value of local_rts variable in the peer interface. If mr_training_enable is true and the "continue training" bit of the control field of received training frames on all lanes of the interface is zero then remote_rts is set to true, otherwise it is set to false. If mr_training_enable is false then remote_rts is always true.

We will need an exception for remote_rts along the following lines. This variable is set to true if not detecting PRBS31 encoded by the Inner FEC and local_rx ready is set to true. It is otherwise set to false.

A later slide provide potential implementation of this as an exception in Clause 185.

The peer interface might be sending mission data rather than PRBS31 before remote_rts is asserted. Other information from the PMA might be helpful but requires further definition of these service interfaces to convey the information.

Waiting for adjacent interface then send inverted PRBS31 encoded by Inner FEC.
Waiting for peer to transmit then Rx to adapt.
Could happen in either order.
Note that until Rx has adapted, we cannot identify which PRBS31 state is being sent.

Transmitter is inverted PRBS31 encoded by Inner FEC.
Tx clock is derived from adjacent interface.
Rx is receiving a good signal and has completed adaptation.

Inner FEC now encodes the data from the adjacent interface (mission data).

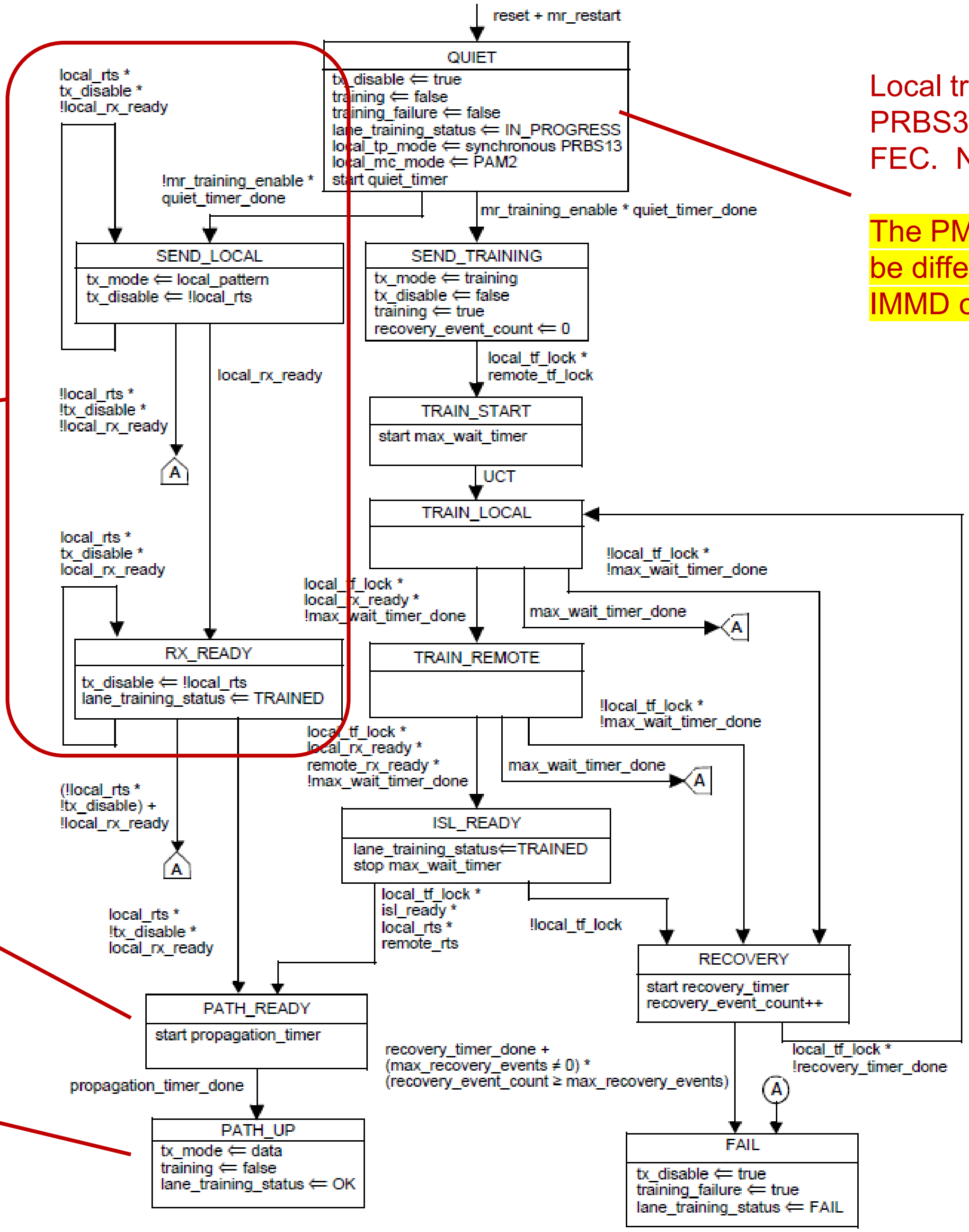


Figure 178B–10—Training control state diagram

Local transmitter is sending PRBS31 encoded by the Inner FEC. Not squelched.

The PMD behavior here would be different from copper and IMMD optical PMDs.

Implementation: New APSU subclause

Add a new subclause in Clause 185 as follows:

185.x.x Autonomous path startup functions (APSU) functions

A PMD shall provide the ready-to-send (RTS) function and the ILT function as specified in Annex 178B.

The variable mr_training_enable is always set to false.

When tx_disable is true, the PMD transmits PRBS31 encoded by the 800GBASE-LR1 Inner FEC, as defined for the test pattern generator in 184.6.1.

When tx_disable is false and tx_mode is local_pattern (see 178B.8.3.1) the PMD transmits inverted PRBS31 encoded by the 800GBASE-LR1 Inner FEC, as defined for the test pattern generator in 184.6.1, except that all PRBS31 bits are inverted.

When tx_disable is false and tx_mode is data transmit data from the adjacent interface encoded by the Inner FEC. [this sentence may not be required as a similar sentence is not provided in other PMD clauses]

The remote_rts variable is redefined as follows. The remote_rts is set to false in the QUIET state. Remote_rts is set to true after failing to identify PRBS31 encoded by Inner FEC.

Note that we might need to be prescriptive about how PRBS31 is identified as being present or not.

Detecting PRBS31

- Detection of PRBS31 might be accomplished by monitoring errors over an appropriate window using the PRBS31 test pattern checker defined in 184.6.2.
- Note however that the checker provides only a set of bin counters that need to be summed together. An alternate method, perhaps at the discretion of the implementer, may be needed.

184.6.2 Test pattern checker

The Inner FEC shall include a PRBS31 test pattern checker over the inverse permutation output PCS lanes (see 184.4.3). To recover the PRBS31 test pattern, 10-bit blocks are round-robin collected from the 32 inverse permutation function output flows.

The test pattern checker is enabled by the rx_checker_enable control variable.

The PRBS31 test pattern checker shall include block error detection and 17 related counters. Block error detection and behavior of the counters is defined in 174A.9.

The following counters shall be implemented:

test_block_error_bin_0_*k*
A set of 16 48-bit counters where counter *k* counts once for each test block received with exactly *k* errored test symbols, *k* = 0 to 15.

test_block_error_bin_0_16p
A 48-bit counter that counts once for each test block received with 16 or more errored test symbols.

Generating PRBS31

PRBS31 test pattern input to the Inner FEC is defined in 184.6.1.

Note that the (not inverted) PRBS31 generator, shown in Figure 49-9, includes an inversion of the feedback shift register.

Inverted PRBS31 defined as an alternate pattern requires a second inversion.

184.6.1 Test pattern generator

The Inner FEC shall include a PRBS31 test pattern generator (Figure 184–2). In test mode, 10-bit blocks from the PRBS31 generator are round-robin distributed to the 32 pcsla Inner FEC flows.

The PRBS31 generator implements the PRBS31 pattern defined by Equation (49–2) and shown in Figure 49–9.

The test pattern generator is enabled by the tx_generator_enable control variable.

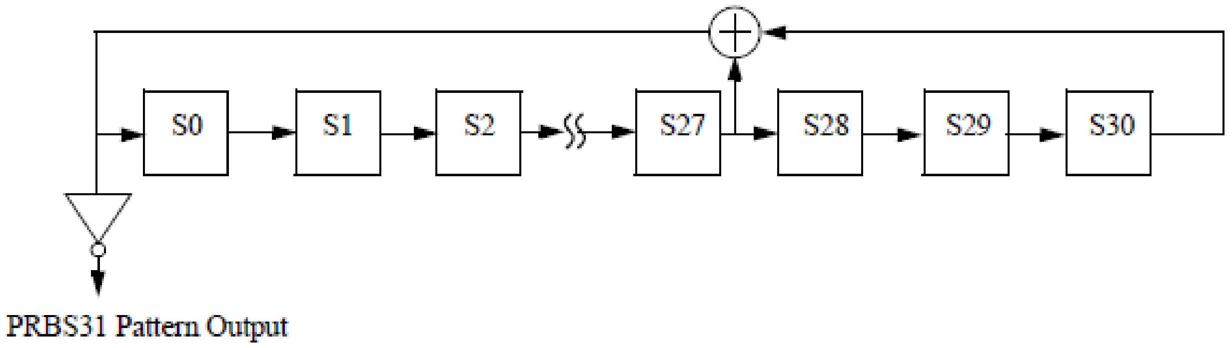


Figure 49–9—PRBS31 pattern generator

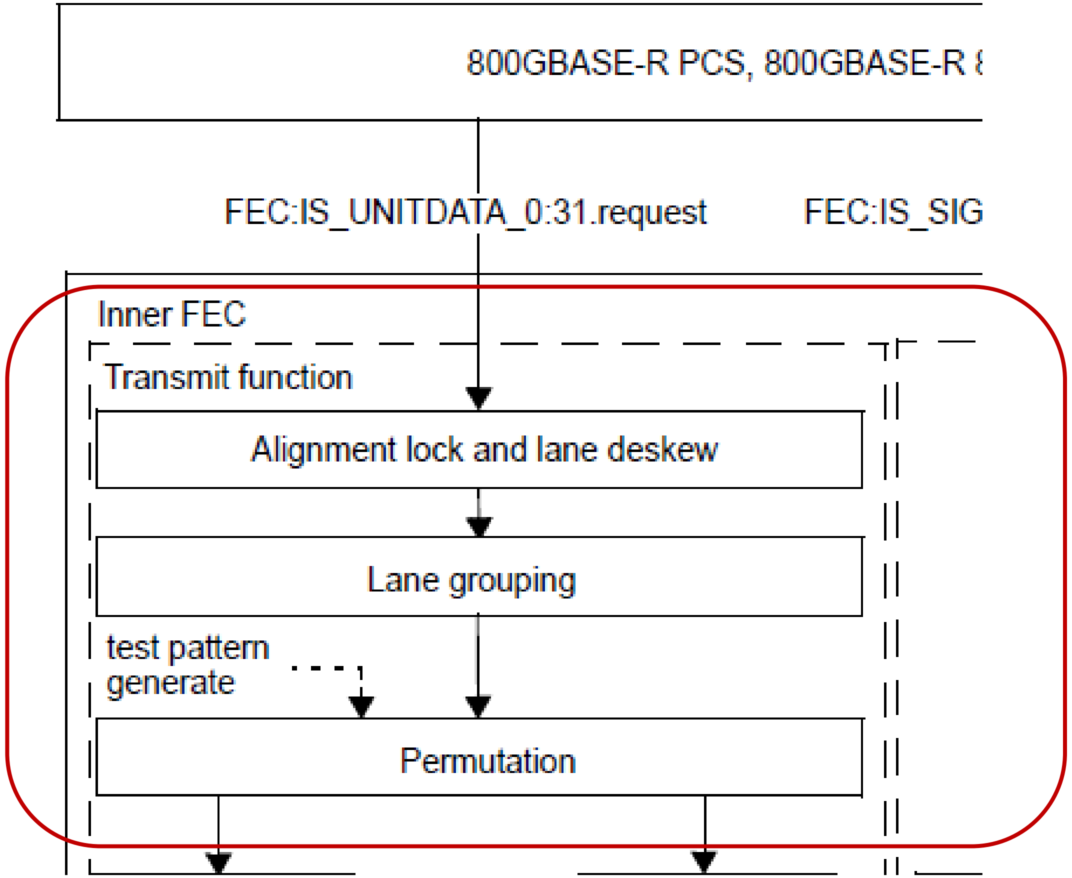
Service interface

- Need to update the 800GBASE-LR1 sublayer service interfaces to include signal_ok in the downward direction.
  - FEC:SIGNAL.request(SIGNAL_OK) [above]
  - PMD:SIGNAL.request(SIGNAL_OK) [below]

- This is also need for the squelch signaling method.

Switching from PRBS31 to data

- The PRBS31 test pattern generator is replacing aligned/deskewed PCS lanes as provided in the Inner FEC.
- Switching from PRBS31 pattern to data requires both signal_ok from above to be OK but also the local alignment lock and deskew to be complete.
- Need to work this into the process.
- Note this consideration is relevant to the squelch method as well.



Observations

- RTS signaling using PRBS31 and inverse is viable.
- Previous slides provided some suggestions in defining.
- There are a number of issues to resolve to have complete proposal.

800GBASE-LR1 APSU using squelch

APSU with squelch + PRBS31

- For 800GBASE-LR1, squelch does not result in a system performance issue since it is not used in a multiple wavelength, amplified system.
- However, receiver adaption will likely take much longer than for an IMDD receiver.
  - Is this an issue?
- While waiting for mission data fill the Inner FEC with PRBS31.
- Use tx_disable as we do for IMDD to squelch the PMD output.
- The local_pattern path on the training control state diagram (Figure 178B-10) can otherwise be used with only minor rewording.

Transmit disable

185.5.7 PMD global transmit disable function (optional)

The PMD global transmit disable function is optional and allows the optical transmitter to be disabled.

When the PMD_global_transmit_disable variable is set to one, this function shall turn off the optical transmitter so that it meets the requirements of the average launch power of the OFF transmitter in Table 185–5.

If a PMD fault is detected (see 185.5.8), then the PMD may set the PMD_global_transmit_disable variable to one, turning off the optical transmitter.

Transmit disable is already defined for 800GBASE-LR1. The definition is essentially the same as for the IMDD PMDs.

It is currently optional, but would have to be mandatory if we need it for APSU.

Waiting for adjacent interface then enable TX and send PRBS31 within Inner FEC.
Waiting for peer to transmit, then Rx to adapt.
Could happen in either order.

Transmitter is not squelched.
Sending PRBS31 encoded by Inner FEC.
Tx clock is derived from adjacent interface.
Rx is receiving a good signal and has completed adaptation.

Inner FEC now encodes the data from the adjacent interface (mission data).

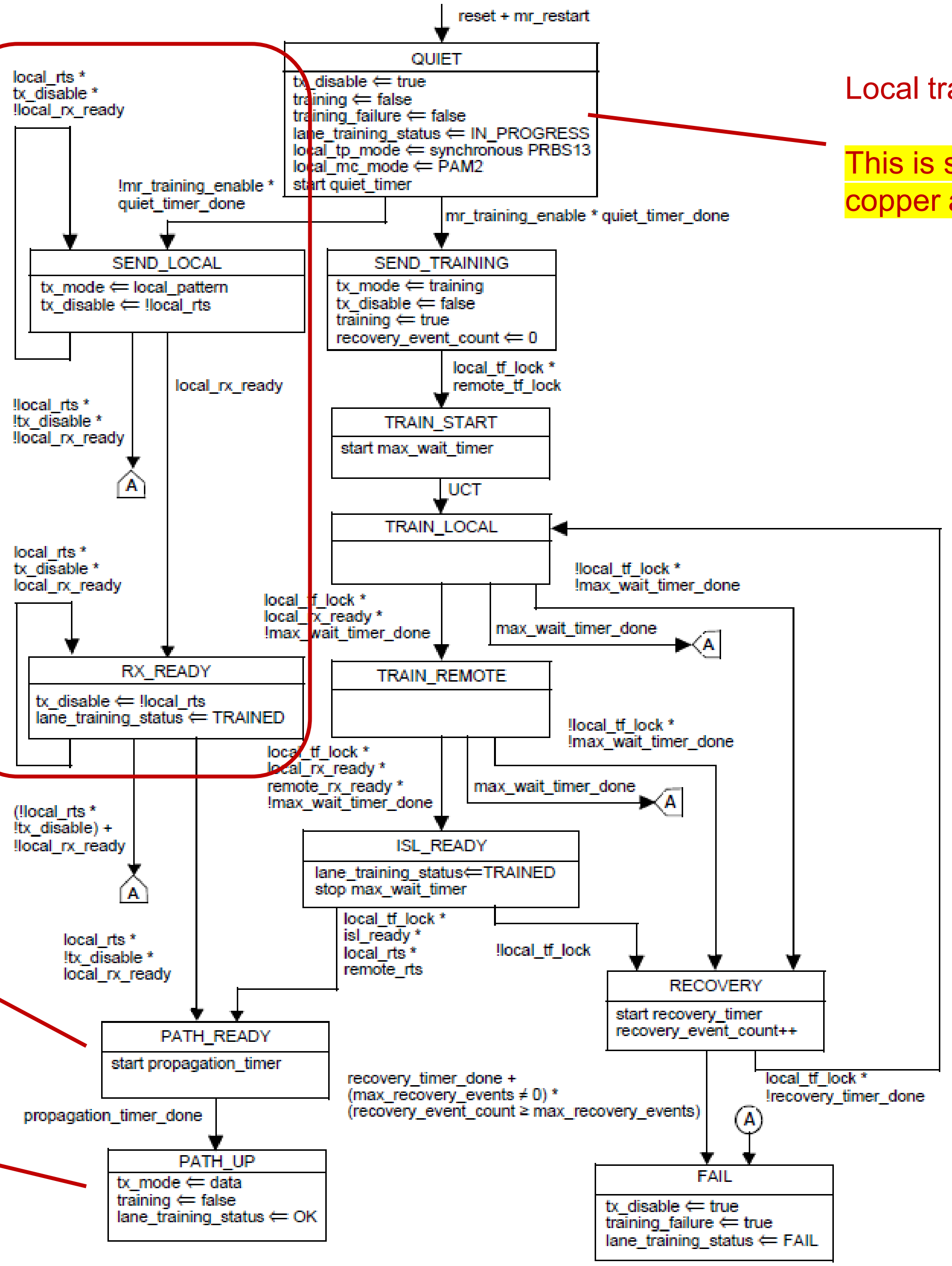


Figure 178B–10—Training control state diagram

Local transmitter is squelched.

This is same behavior as copper and IMDD PMDs.

New APSU subclause

Add a new subclause in Clause 185 as follows:

185.x.x Autonomous path startup functions (APSU) functions

A PMD shall provide the ready-to-send (RTS) function and the ILT function as specified in Annex 178B.

The variable mr_training_enable is always set to false.

When tx_mode is local_pattern (see 178B.8.3.1) the PMD transmits PRBS31 encoded by the 800GBASE-LR1 Inner FEC as defined for the test pattern generator in 184.6.1.

Observations

- RTS signaling using squelch is viable.
- The mapping is simpler in some ways compared with the PRBS31 method.
- Previous slides provided some suggestions in defining.
- There are a number of issues to resolve to have complete proposal; but fewer than for the PRBS31 method.