



# Fast link recovery proposal v2.0

12/7/2009

Gavin Parnaby

# Capability overview

- Use autoneg bit to advertise support for fast retrain capability
  - **Both sides must have enabled fast retrain**
  - **Fast retrain can be disabled using LLDP**
- Either side can detect a link failure
  - **refresh monitor state diagram and vendor specific criteria**
    - snr degradation or missing refreshes
- Signal to the other side that the link is bad by signaling link failure, using an easily detected symbol sequence
  - **The data link goes down on both sides immediately after the alert**
- Both sides re-enter training in coefficient exchange state
  - **Use the normal startup protocol, but reduce timing allowances**
  - **Use a max\_wait\_timer with reduced period**
  - **Leverage existing protocol**

# Add overview text in 55.4.2.5.14a

## 55.4.2.5.14a Fast retrain capability

PHYs with the EEE capability also support a fast retrain mechanism. This allows PHYs on links with degraded performance to return to the normal operational mode much more rapidly than through a normal retrain.

The fast retrain mechanism is controlled by the fast retrain state diagram (Fig 55-??). When the PHY detects a link failure, it sets the variable `loc_fr_req` to TRUE. This causes the transmission of an easily-detected link failure signal. Following the link failure signal, the two link partners transition back to the `pma_coeff_exch` state and follow the training procedure described in 55.4.2.5.14, with the exception that the initial infocfield countdown values are reduced as shown in Figures 55-25 and 55-26.

To ensure interoperability the training times in Table 55-?? should be observed.

<insert table from slide 18 as new Table 55-??>

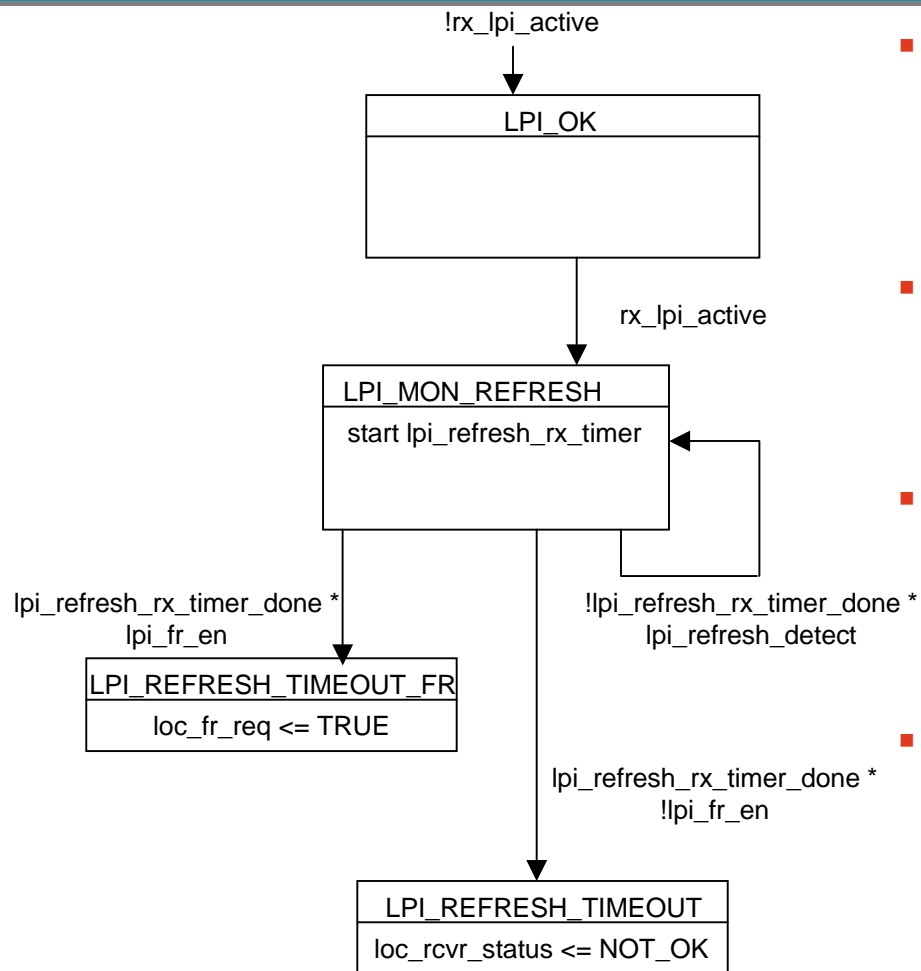
# Add overview text for refresh monitor in 55.4.2.6a

## 55.4.2.6a Refresh monitor

EEE capable PHYs shall implement the Refresh Monitor state diagram shown in Figure 55-??.

This function ensures that PHYs detect silent link partners during the receive low power mode within 330us (two complete quiet-refresh cycles). When a silent link partner is detected, the PHY forces a link retrain.

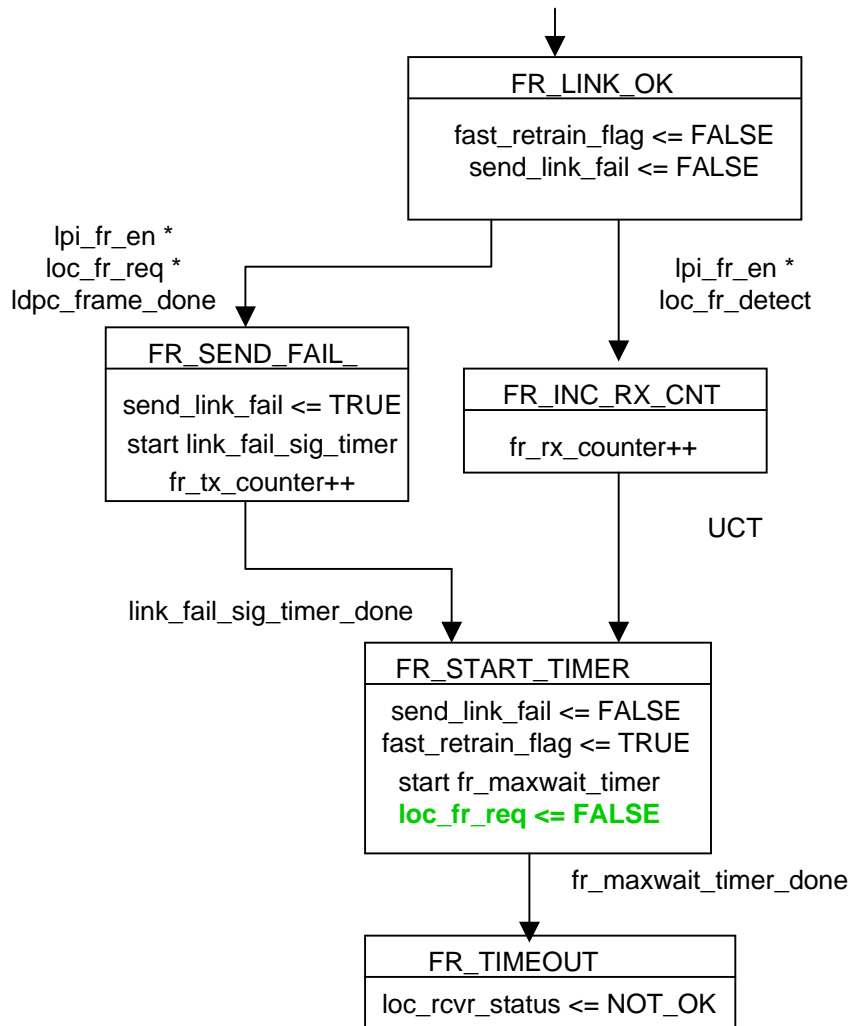
# Add refresh monitor state diagram to 55.4.6



- Add variable **lpi\_refresh\_detect**
  - Set TRUE when the receiver has reliably detected a refresh signal. The exact criteria left to the implementer.
- Add timer **lpi\_refresh\_rx\_timer**
  - Period of 330us
  - 330us <-> 2 complete qr cycles
- Add variable **lpi\_fr\_en**
  - When TRUE the fast retrain capability is enabled. The variable is set through a management register.
- **LPI\_REFRESH\_TIMEOUT\_FR** and **LPI\_REFRESH\_TIMEOUT** both force a link retrain (fast / normal respectively)

# Add fast retrain state diagram to 55.4.6

$tx\_mode = (SEND\_N * !loc\_fr\_req * !loc\_fr\_detect) + !lpi\_fr\_en$



- Add a variable **loc\_fr\_req**
  - when set indicates the local PHY has detected a link failure
  - set through the refresh monitor state diagram and optionally through other vendor-specific means. It causes a transition to FR\_SEND\_FAIL, sending the link failure signal to the link partner.
- Add a variable **loc\_fr\_detect**
  - set true when the link failure signal is reliably detected in the PMA
- In FR\_START\_TIMER both sides of the link set fast\_retrain\_flag<=TRUE
  - to send the PHY control state machines back to PMA\_Coef\_Exch
- fr\_maxwait\_timer is fallback in case fast retrain fails
  - loc\_rcvr\_status <= NOT\_OK forces full retrain
  - PCS\_Status=OK sends state machine back to FR\_LINK\_OK

# New variables: add to 55.4.5.1

- **lpi\_fr\_en**
  - Set TRUE through a management register. Advertised/resolved during autoneg.
- **loc\_fr\_req**
  - Set TRUE when the receiver has detected a link failure condition and is requesting a fast retrain, set FALSE otherwise
- **loc\_fr\_detect**
  - Set TRUE when the receiver has reliably detected the link failure signal. It is highly recommended that loc\_fr\_detect is qualified with the reception of errored blocks at the LDPC decoder output. Set FALSE when the link failure signal is not detected.
- **send\_link\_fail**
  - When TRUE indicates that the PMA should send the link failure signal. When FALSE the variable has no effect.
  - needs to be added to the PCS/PMA interface
- **lpi\_refresh\_detect**
  - Set TRUE when the receiver has reliably detected a refresh signal and FALSE otherwise. The exact criteria left to the implementer.

## New timers: add to 55.4.5.2

- **link\_fail\_sig\_timer**
  - Determines the length of time the PHY sends the link failure signal. Has a period equal to 4 LDPC frame periods.
- **fr\_maxwait\_timer**
  - Determines the period of time the PHY has to set PCS\_Status = OKAY following a fast retrain before the fast retrain is aborted and a full retrain performed. Has a period of 30 ms



## New counters: new section 55.4.5.4

- **fr\_tx\_counter**
  - Counts number of transmit link failure signals
- **fr\_rx\_counter**
  - Counts number of receive link failure signals
- Both counters need management registers (see slides 18-19)

## Link failure signaling – add to PMA Alert subclause 55.4.2.2.1, add text to 55.4.2.4

- **Link retrain request signaling is generated when send\_link\_fail is TRUE**
  - Has priority over LPI signaling (replaces LPI alert/refresh/quiet)
- **Inverted alert sequence used as ‘link failure signal’**
  - Four frames of inverted (multiplied by  $-1$ ) LPI alert signaling on the alert pair
    - No new hardware needed, can use existing alert generator/detector
  - **Recommend that detection is qualified with repeated frames of receive errors at LDPC decoder output**
- Reliably detecting this alert sequence sets loc\_fr\_detect TRUE

# Fast retrain : changes to Fig 55-24

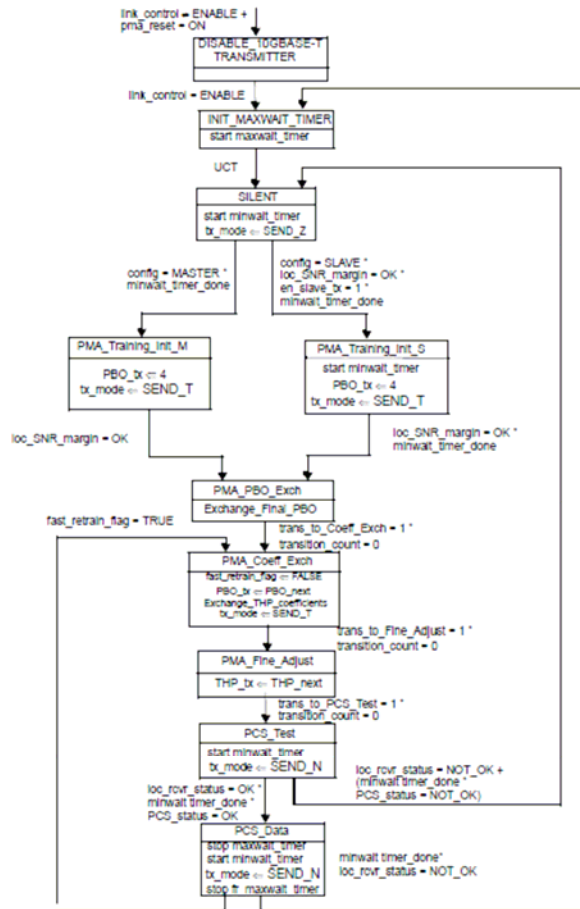
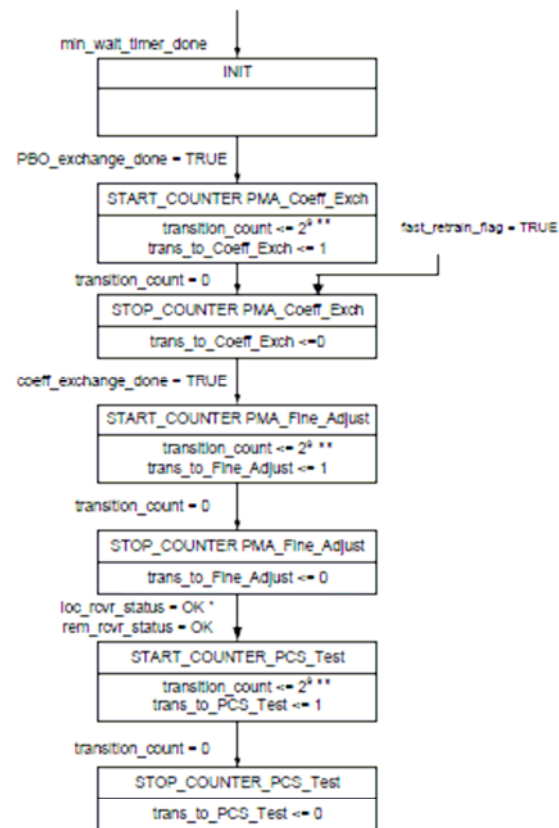


Figure 55-24—PHY Control state diagram

- Re-enter at PMA\_Coeff\_Exch
- PAM2 signaling eliminates slicer errors
- Robust training
- Reuse existing states and transition protocols
- Minimize new text

# Figure 55-25 changes



\*\* - devices supporting fast retrain use  $2^5$

Figure 55-25—MASTER transition counter state diagram

# Figure 55-26 changes

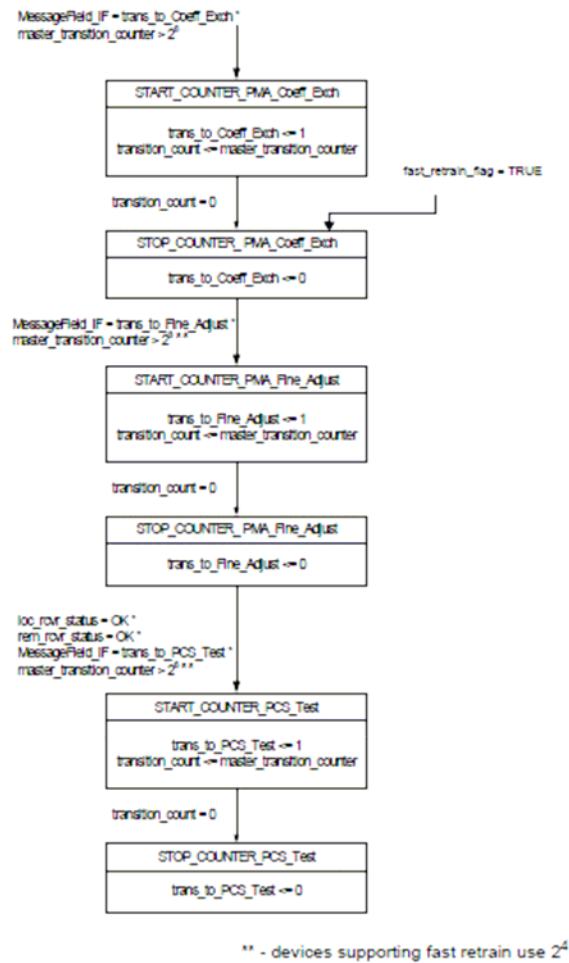


Figure 55-26—SLAVE transition counter state diagram

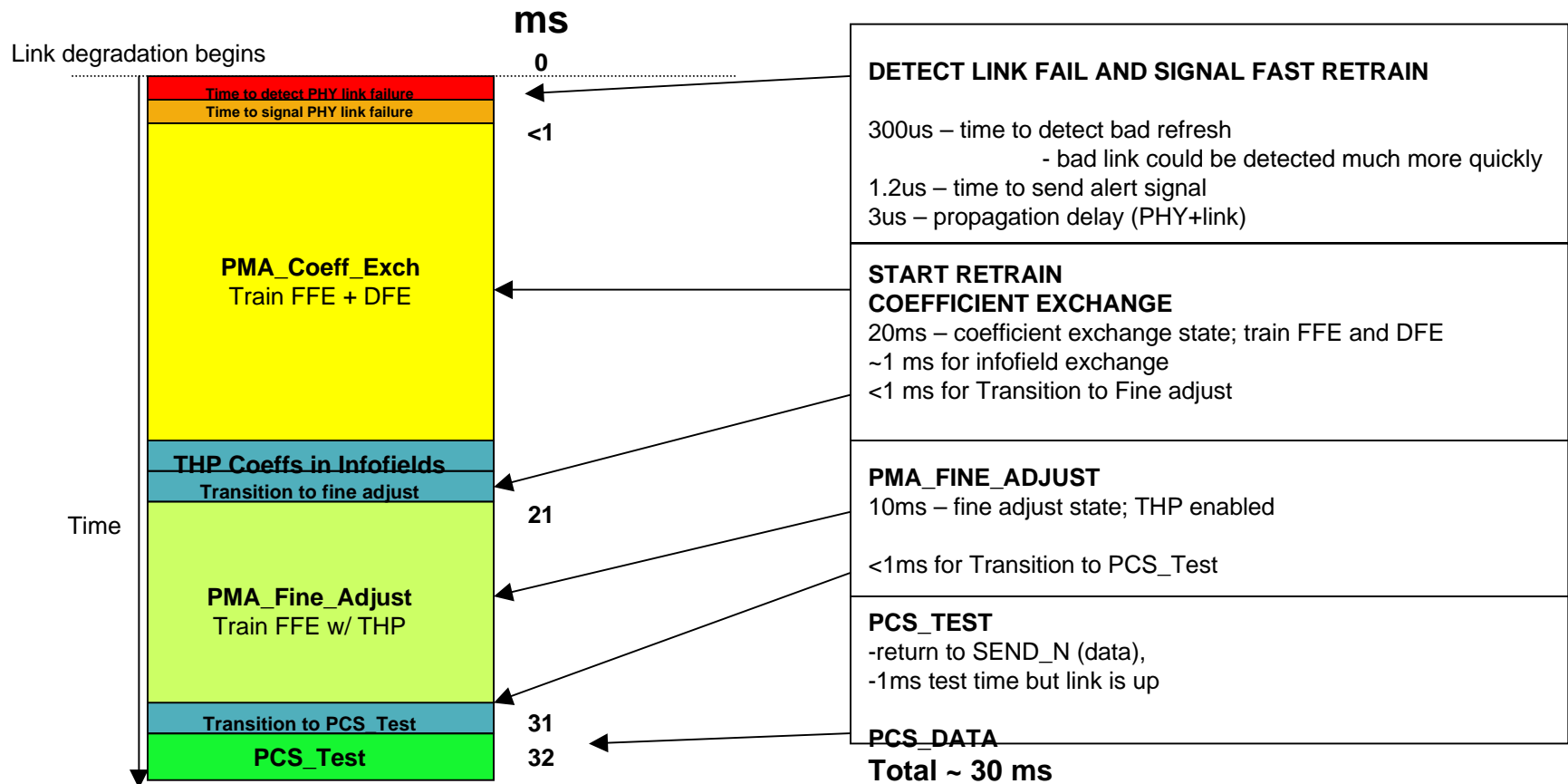
# Timing analysis

- Timing recovery does not need to be reacquired from scratch
  - **The slave should still be locked to the master**
  - **SNR has degraded somewhat below operating threshold**
- Therefore recommend that Master/Slave have symmetric training times
- Coefficient exchange needs to train FFE + DFE together
- Training update uses THP with fixed coefficients
  - **fewer filters to train, FFE should already be converged**

# Justification for 30ms

- Training time required depends on how much degradation there has been on the link
  - **difficult to quantify level of degradation**
  - **clearly not starting from empty state**
- The assumption is that the receiver is still close to the operating condition but SNR has degraded below the PAM16 operating SNR
- Short burst of PAM2 training used to get back to operating condition
- 30ms is
  - **below the 50ms threshold stated as required to avoid system link failover**
  - **high enough that many links can be recovered**
  - **small enough that if the fast retrain fails it does not materially affect the normal retrain time**
- If you have a different number please propose it so that it can be discussed

# Fast retrain time budget



**NOTE: Infocoeff exchange calculation**  
 [4 coeffs per infocoeff, 16 coeffs per pair, 64 coeffs to exchange -> 16 infocoeffs;  
 An infocoeff every 16384 symbols (20us)  
 Therefore minimum 0.32ms for all coeffs; pad to 1ms]



# Shortened state timing : include in 55.4.2.5.14a

- Reduce infocfield countdown from 10ms to <1ms
  - **Minimize transition sync. overhead**

<b>State</b>	<b>Recommended Maximum time (ms)</b>
PMA_Coeff_Exch state	20
PMA_Fine_Adjust state	10

# Changes to clause 45 (Management/Autoneg)

## **45.2.1.75a** NEW: Fast retrain status and Control register 1.147

1.147.0 Fast Retrain Enable Type: R/W

1 = Enable fast retrain

0 = disable fast retrain

(Note: disabling this bit while a link is up will cause the PHY to stop supporting fast retrain and the link will drop if the link partner initiates a fast retrain.)

1.147.10:6 LD Fast Retrain Count – counts the number of fast retrains requested by the local device Type: RO/NR

The BER counter is a 5 bit count of the number of fast retrains requested by the local device for 10GBASE-T. These bits shall be reset to all zeros when

read or upon execution of the PCS reset. These bits shall be held at all ones in the case of overflow.

1.147.15:11 LP Fast Retrain Count – counts the number of fast retrains requested by the link partner Type: RO/NR

The BER counter is a 5 bit count of the number of fast retrains requested by the link partner for 10GBASE-T. These bits shall be reset to all zeros when

read or upon execution of the PCS reset. These bits shall be held at all ones in the case of overflow.

## **45.2.7.10** 10GBASE-T AN control register

7.32.1 Fast Retrain ability Type: R/W

1 = Advertise PHY as 10GBASE-T fast retrain capable.

0 = Do not advertise the PHY as 10GBASE-T fast retrain capable

## **45.2.7.11** 10GBASE-T AN status register Type: RO

7.33.1 Link Partner Fast Retrain capability

1 = Link partner is able to perform fast retrain

0 = Link partner is not able to perform fast retrain

# Changes to clause 55

55.6.1.2 10GBASE-T Auto-Negotiation page use

Type: RW

Table 55-11

U19

Fast Retrain ability

(1 = support of Fast Retrain and 0 = no support)

## Next steps

- **Generate Framemaker file from this proposal**
  - **Adds two new state diagrams, small modifications to three existing PHY control state diagrams, 1 new autoneg bit in Clause 55**
  - **Add variables / counters / two new subclauses**
  - **Add management registers/bits to Clause 45**