10GBASE-T ad hoc report

Fast retrain and link monitoring

Gavin Parnaby, Solarflare Communications

January 26th 2010

New Orleans

IEEE 802.3az

Energy Efficient Ethernet

10GBASE-T ad hoc

- The 10GBASE-T ad-hoc held two meetings in December to create a detailed proposal in response to comment #186 on draft 2.1
- Two ad hoc meetings were held
  - Initial phone conference December 3rd 2009
  - Face-to-face meeting December 10th 2009 held in Irvine, CA
- Thank you to those who attended

Review of contributions

- Phone conference
  - grimwood_01_1209.pdf, Mike Grimwood, Broadcom
    - pointed out issue with LPI during training (not fast-retrain specific) – comment made on draft 2.1
    - requests justification for retrain timer values
- Face-to-face meeting
  - parnaby_01_1209.pdf, Gavin Parnaby, Solarflare
    - adds details on retrain timer values
    - shows suggested text to be used in the draft
    - adds details to fast retrain proposal in parnaby_03_1109.pdf
    - replaces all TBDs with recommended values
    - fixes one state machine bug

Consensus output - draft text

- Final text is in
  - http://www.ieee802.org/3/az/public/adhoc/10gbt/dec1009/8023az-55_fast_retrain_text.pdf
- Minor changes to original proposal
  - changed link monitor timing (using 8.2ms window)
  - removed LLDP requirement
    - since fast retrain disable can be done unilaterally
  - text was edited to state that fast retrain is optional and not a requirement for EEE PHYs
  - added requirement that fr-compliant PHYs shall transition to PAM2 within 9 LDPC frames of receiving a fast retrain request
  - changed transition condition into FR_LINK_OK to fix a state machine reset issue pointed out by Mike Grimwood
  - added details on autoneg and management requirements

Clause 45

45. Management Data Input/Output (MDIO) Interface

Insert a new subclause 45.2.1.75a after the existing 45.2.1.75 as shown below:

45.2.1.75a 10GBASE-T fast retrain status and control register (Register 1.147)

Table 45–49a—10GBASE-T fast retrain status and control register bit definitions

Bit(s)	Name	Description	R/W[a]
1.147.15:11	LP fast retrain count	Counts the number of fast retrains requested by the link partner	RO/NR
1.147.10:6	LD fast retrain count	Counts the number of fast retrains requested by the local device	RO/NR
1.147.5:1	Reserved	Value always 0, writes ignored	RO
1.147.0	Fast retrain enable	1 = Fast retrain capability is enabled 0 = Fast retrain capability is disabled	R/W

[a]R/O = Read only, R/W = Read/Write, NR = Non Roll-over

Autoneg changes

45.2.7.10 10GBASE-T AN control register (Register 7.32)

Table 45–140—10GBASE-T AN control register

Bit(s)	Name	Description	R/W[a]
7.32.1	Fast retrain ability	1 = Advertise PHY as 10GBASE-T fast retrain capable 0 = Do not advertise the PHY as 10GBASE-T fast retrain capable	R/W

[a]R/W = Read/Write, R/O = Read only

45.2.7.11 10GBASE-T AN status register (Register 7.33)

Table 45–141—10GBASE-T AN control register

Bit(s)	Name	Description	R/W[a]
7.33.1	Fast retrain ability	1 = Link partner is capable of 10GBASE-T fast retrain 0 = Link partner is not capable of 10GBASE-T fast retrain	R/O

[a]R/O = Read only, SC = Self-clearing, LH = Latching high

55.4.2.6a Refresh monitor function

55.4.2.6a Refresh Monitor function

Insert the following subclause after subclause 55.4.2.6 in draft 2.2

The Refresh monitor is required for PHYs which support the EEE capability. The Refresh monitor operates when the PHY is the low power receive mode. The Refresh monitor shall comply with the state diagram of Figure 55–27a. The function forces a link retrain if a refresh signal is not reliably detected within a moving time window equivalent to 50 complete quiet-refresh cycles (nominally equal to 8.192 ms), when the PHY is in the lower power receive mode.

Link monitoring

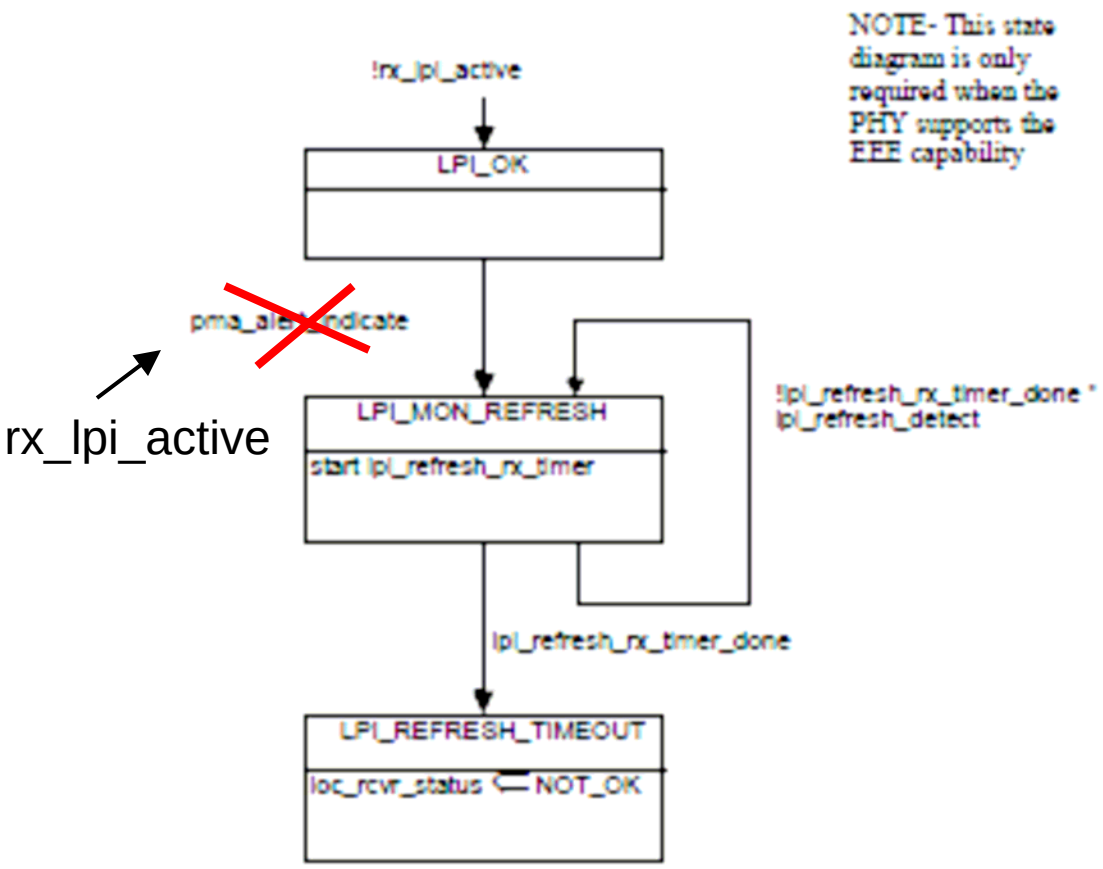

Figure 55–27a—EEE Refresh monitor state diagram

- Forces PHY to retrain if refresh signaling is NOT detected within 8.2ms window
- Avoids zombie state during LPI
- Note cut/paste error in draft

55.4.2.5.15 Fast retrain function

55.4.2.5.15 Fast retrain function

Insert the following subclause after subclause 55.4.2.5.15 in draft 2.2

PHYs that support the fast retrain capability shall implement the fast retrain state diagram shown in Figure 55–27b. PHYs may request a fast retrain by setting the variable loc_fr_req to TRUE. This causes the transmission of an easily-detected link failure signal. After transmitting the link failure signal the PHY shall transition to the PMA_Coeff_Exch state and send PAM2 signaling within a time period equivalent to 9 LDPC frame periods after completing the link failure signal.

After the detection of the link failure signal, a PHY shall transition to the PMA_Coeff_Exch state and respond with PAM2 signaling within a time period equivalent to 9 LDPC frame periods after receiving the link failure signal.

Note that reliable traffic on the transmitter may be interrupted when the local receiver requests a fast retrain.

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 infofield countdown values are reduced as indicated in Figures 55-25 and 55-26.

To ensure interoperability the training times in Table 55–6a should be observed during the fast retrain.

Table 55–6a—Recommended fast retrain sequence timing

State	Recommended maximum time (ms)
PMA_Coeff_Exch state	20
PMA_Fine_Adjust state	10

Fast retrain

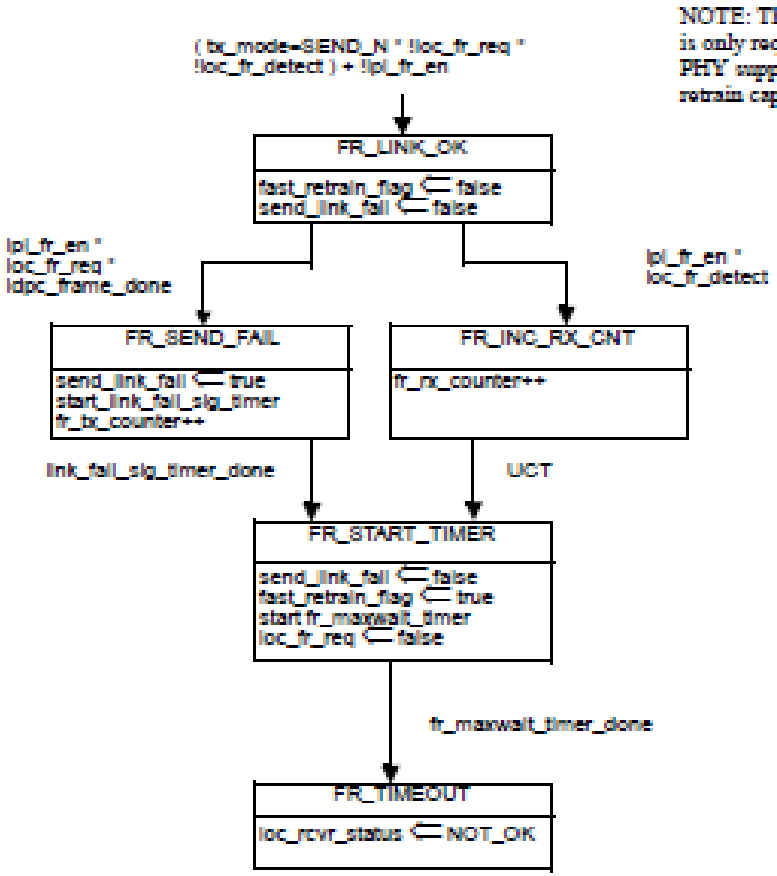

Figure 55–27b—Fast retrain control state diagram

- PHY may request fast retrain using a special link failure signal
- The link has 30ms after the fast retrain signal to return to pcs data mode
  - If the fast retrain fails the link performs a full retrain

PHY Control State diagram changes

55.4.6.1 PHY Control state diagram

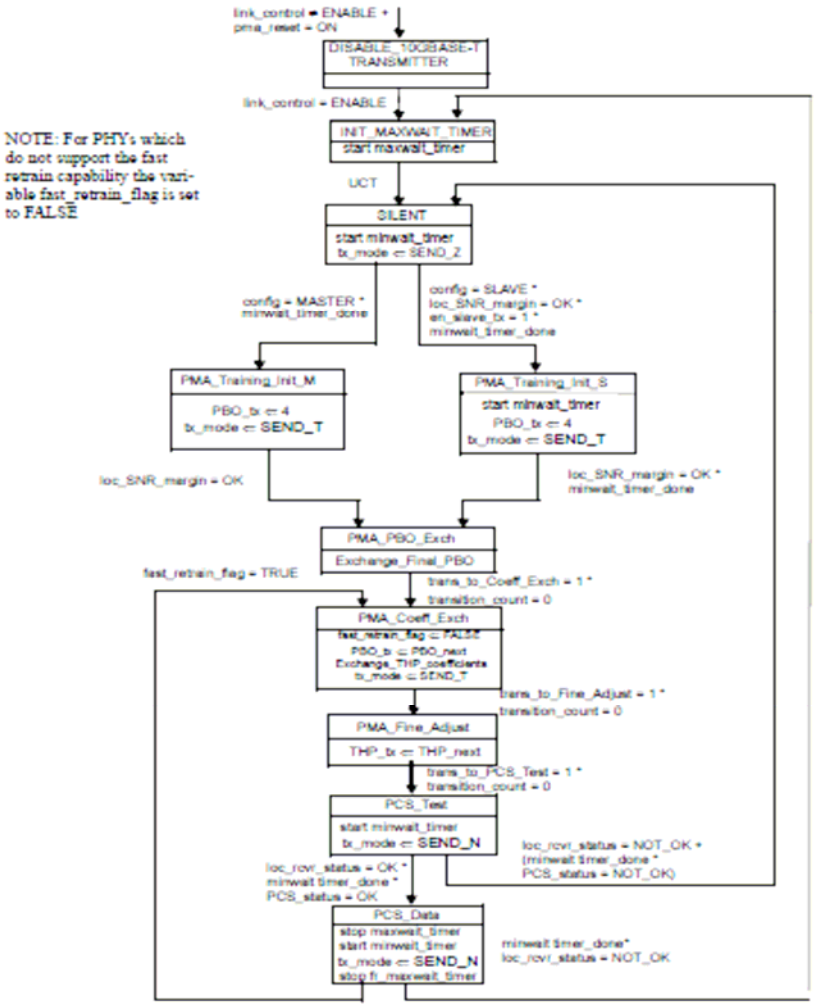

Figure 55–24—PHY Control state diagram

Master/slave transition counter

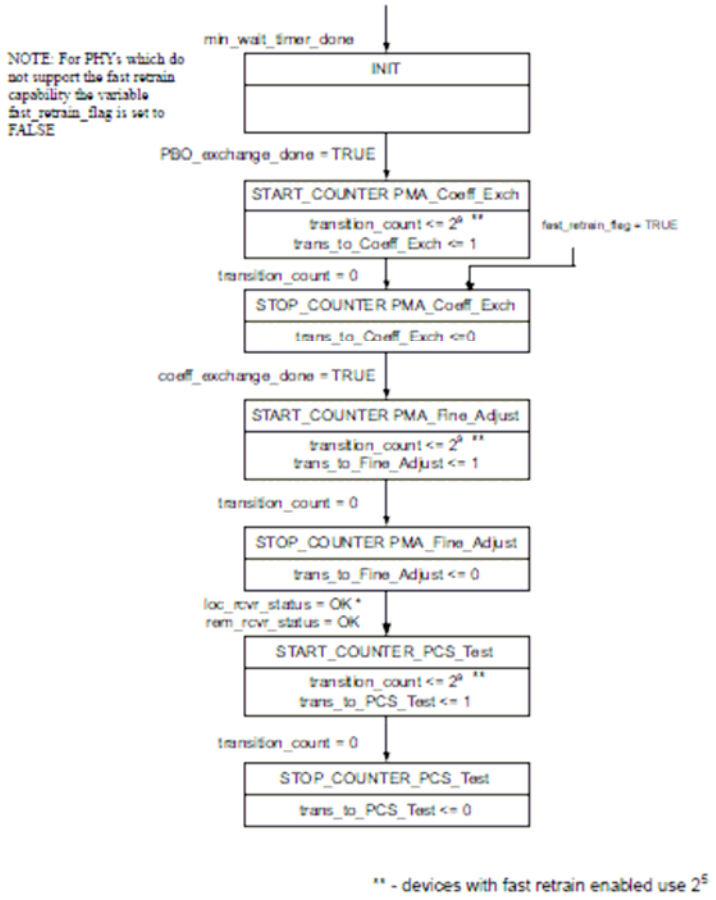

Figure 55–25—MASTER transition counter state diagram

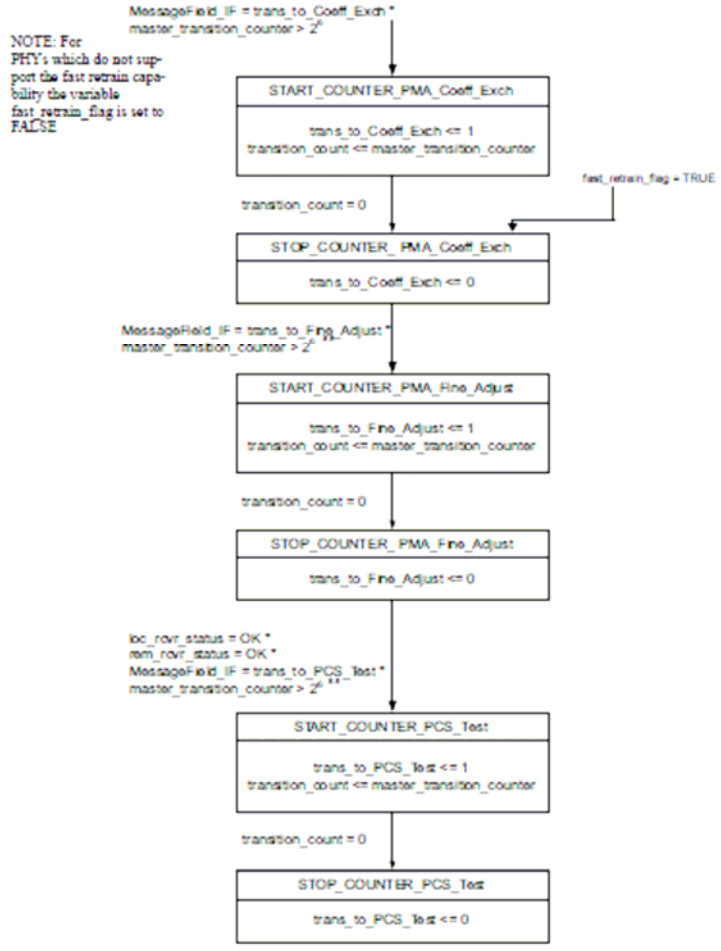

Figure 55–26—SLAVE transition counter state diagram

Discussion

- Consensus reached on fast retrain text
  - mandatory refresh monitor
  - optional fast retrain capability
- Points of discussion
  - ‘May’ vs ‘Shall’ for fast retrain
  - LLDP support

Summary and motion

The 10GBASE-T ad hoc reached consensus on the changes to the text to support fast retrain and link monitoring.

Motion:

The IEEE P802.3az Task Force adopts the fast retrain and link monitoring function as described in 8023az-55_fast_retrain_text.pdf with the edit shown on slide 8

- **Move:**
- **Second:**

Attendance

- 3rd December
  - Mike Grimwood, Bill Bliss, Tom Souvignier, Pat Thaler (Broadcom)
  - Mike Bennett (LBNL)
  - Bill Woodruff (Aquantia)
  - Matt Brown, Dariush Dabiri (AMCC)
  - Scott Kip (Brocade)
  - Dimitry Taich, Jose Tellado, Jennifer Wu, Sanjay Kasturia (Teranetics)
  - Dan Dove (HP)
  - Gavin Parnaby, George Zimmerman, Brett McClellan (Solarflare)
- 10th December
  - Mike Grimwood, Tom Souvignier, Jim Graba (Broadcom)
  - Bill Woodruff (Aquantia)
  - Matt Brown (AMCC)
  - Dmitri Taich, Sanjay Kasturia (Teranetics)
  - Gavin Parnaby, Brett McClellan (Solarflare)