EEE Improvements For Highly Sparse Traffic

Saied Benyamin

George Zimmerman

Mission Statement

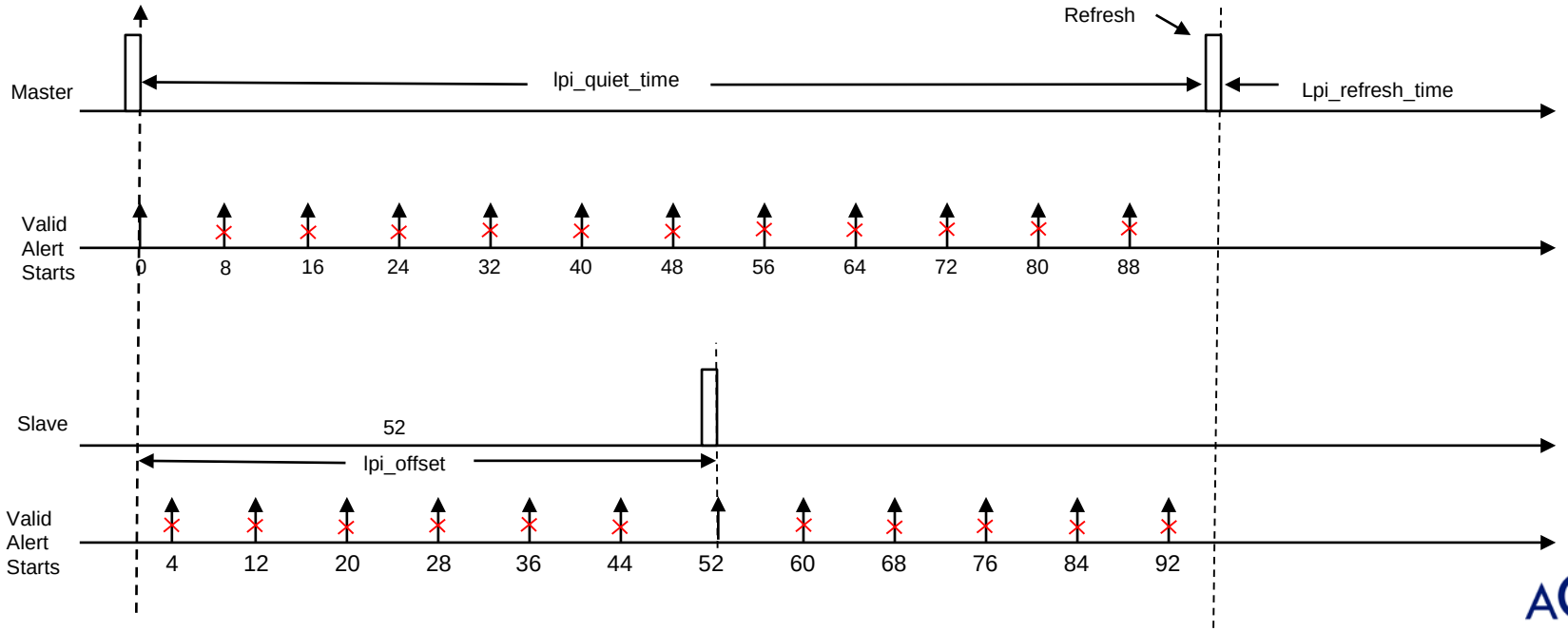
- The automotive infrastructure includes a large number of camera interfaces. These interfaces currently use technologies such as MIPI or LVDS which are simplex connections and do not carry the burden of power consumption for the upstream direction. The upstream commands to the camera use low speed, low power consumption communication such as I2C.
- In order for ethernet to proliferate into this interface, it is highly desirable for us to have a solution with competing power.

EEE power management

- While in LPI, the receiver can power down portions of its circuitry during quiet periods, but it must still be constantly or very frequently listening for Alert. Extra power consumption is resulted because of demands for rapid ‘instant’ return to data mode.
- When traffic is not backing up or otherwise longer wake times are not disadvantageous, we may extend the time to return to data transmission.
- Many applications, such as camera devices, can tolerate a slow return to data traffic in the direction with sparse traffic. Taking advantage of this leads to possible simplifications in ALERT recovery to full speed.
- Propose
  - A method to allow a lower power Alert Detect circuitry at the cost of additional wake up time
  - Allow each port to independently decide if it wants to use this method for its transmission of alert and communicate this to the link partner. The power saving is mainly on the receive side.

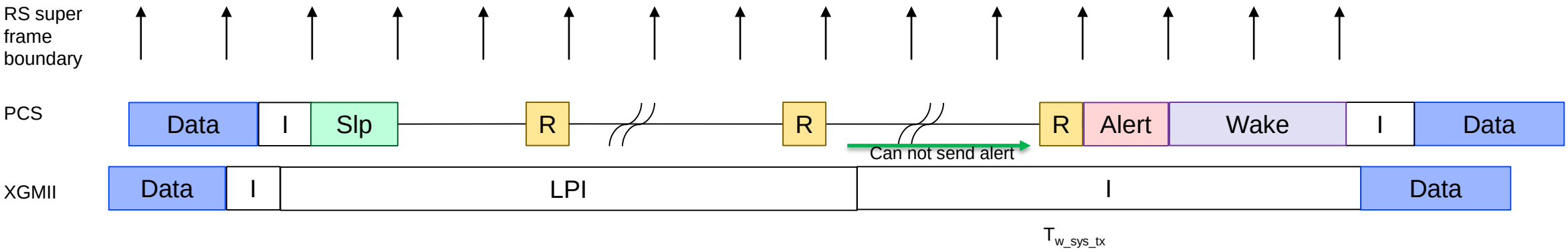
Slow Wake: Less Frequent Alert

- Alert is currently allowed to start on every 8th RS-frame
- When Slow Wake is active, limit the locations in the QR cycle at which alert can be transmitted (similar proposal by Lo_3ch_03a_0319.pdf - Low Datarate EEE)
- When Slow Wake is set, alert can only start right after refresh
- This means that phy may take up to an additional ~30uS to be ready for data transmission. Tw_sys_tx must be increased by the approximately 30uS in 10G.



Slow Wake

- Slow Wake allows the alert detect to remain powered down through most of the QR cycle.
- As you see in picture below, when XGMII idle starts, PCS has to wait for the next refresh to send ALERT to signal end of QR
- Receiver detector can take advantage of data path already turned on for refresh



Slow Wake: Mechanics of enabling

- Configuring Slow Wake is made through the management entity for the transmit direction and communicated to the link partner.
  - When a delay of 92 RS-Frame periods (~30uS/S) in the transmitted sparse data is tolerable, Slow wake should be enabled to allow the link partner to save power.
  - Configuration is done before the training phase and can not be changed unless link drops and training is repeated.
- Slow Wake is set independently in each direction of communication
- When Slow Wake is configured, at the local end, the transmitter:
  - increases its Tw_sys_tx by 30.72uS
  - sets the slow Wake bit in the phy infofield training register bit group
- The transmitter notifies the link partner that it is going to use Slow Wake using infofield exchange in register 1.2311. The receiver detects that its link partner is using Slow Wake by checking register 1.2312.

Conclusion

- Highly asymmetric traffic patterns require a more power efficient EEE, this is a crucial requirement for 802.3ch in order to proliferate into all the interfaces available in an automobile.
- We propose that when a system determines that its transmitter can tolerate a latency in switching from LPI to data mode, it allows the link partner to limit the locations in the QR cycle where it has to detect Alert by setting Slow Wake.
- The Slow Wake is enabled in each direction independently via infofield exchange

Modifications to the standard Slow Wake

- Table 149-10: Bit 6 change to SlowWakeEnable
- Paragraphs below the table (P138 L36) additions in red:

The format of PHY capability bits is Oct9<6> = SlowWakeRequest, Oct9<7> = EEEen, Oct10<0> = OAMen, Oct10<2:1> = InterleaverDepth, and Oct10<4:3> = PrecodeSel. EEEen and OAMen indicate EEE and MultiGBASET1 OAM capability enable, respectively. The PHY shall indicate the support of these two optional capabilities by setting the corresponding capability bits.

The optional EEE capability shall be enabled only if both PHYs set the capability bit EEEen = 1. The optional BASE-T1 OAM capability shall be enabled only if both PHYs set the capability bit OAMen = 1. InterleaverDepth indicates the requested data mode interleaving depth. PrecodeSel indicates the requested precoder. SlowWakeRequest is set to indicate to the link partner that the PHY will transmit alert only immediately following a refresh. The remaining bits shall be reserved and set to 0.

Modifications to the standard Slow Wake

- Change the heading of the following section
  - **45.2.1.194.2 User field (1.2311.10:45)**
- Add the following section
  - 45.2.1.194.2B Slow wake Request (1.2311.4)
  - When Slow Wake is set, the link partner will transmit alert only immediately following refresh

- Change heading of following section
  - **45.2.1.195.2 Link partner user field (1.2312.10:45)**
- Add the following section
  - 45.2.1.195.2B Link Partner Slow Wake Requested (1.2312.4)

    When Slow Wake is enabled, the phy will transmit alerts only immediately after refresh.

Modifications to the Standard for Slow Wake – **Table 45-155C**

Bit(s)	Name	Description	R/W
1.2311.15:13	Reserved	Value always 0	RO
1.2311.12:11	Interleave request	00: L=1, no interleaving, default for 2.5GBASE-T1 01: L=2, RS-FEC interleaving factor two, default for 5GBASE-T1, Reserved for 2.5GBASE-T1 10: L=4, RS-FEC interleaving factor four, default for 10GBASE-T1, Reserved for 2.5GBASE-T1 and 5GBASE-T1 11: Reserved	R/W
1.2311.10:5~~4~~	~~User field~~Reserved	~~7-bit user defined field to send to the link partner~~Value always 0	~~R/W~~RO
1.2311.4	Slow Wake request	1 = Alert window start is only immediate frame after refresh 0 = Alert window start is every 8th RS frame after refresh	R/W
1.2311.3:2	Precoder requested	00 = no precoder requested 01 = 1-D precoder requested 10 = 1+D precoder requested 11 =1-D^2 precoder requested	R/W
1.2311.1	MultiGBASE-T1 OAM advertisement	1 = MultiGBASE-T1 OAM ability advertised to link partner 0 = MultiGBASE-T1 OAM ability not advertised to link partner	R/W
1.2311.0	EEE advertisement	1 = EEE ability advertised to link partner 0 = EEE ability not advertised to link partner	R/W

Note – there is no User field, this should be Reserved, and is addressed in another comment.

Modifications to the Standard for Slow Wake – **Table 45-155d**

Bit(s)	Name	Description	R/W
1.2312.15:13	Reserved	Value always 0	RO
1.2312.12:11	Link Partner Interleave request	00: L=1, no interleaving, default for 2.5GBASE-T1 01: L=2, RS-FEC interleaving factor two, default for 5GBASE-T1, Reserved for 2.5GBASE-T1 10: L=4, RS-FEC interleaving factor four, default for 10GBASE-T1, Reserved for 2.5GBASE-T1 and 5GBASE-T1 11: Reserved	RO
1.2312.10:54	~~Link Partner User field~~ Reserved	~~7-bit user defined field received from the link partner~~ Value always 0	RO
1.2312.4	Link Partner Slow Wake requested	1 = Alert window start is only immediate frame after refresh 0 = Alert window start is every 8th RS frame after refresh	RO
1.2312.3:2	Link Partner Precoder requested	00 = no precoder requested 01 = 1-D precoder requested 10 = 1+D precoder requested 11 =1-D^2 precoder requested	RO
1.2312.1	Link Partner MultiGBASE-T1 OAM advertisement	1 = Link partner has MultiGBASE-T1 OAM ability 0 = Link partner does not have MultiGBASE-T1 OAM ability	RO
1.2312.0	Link Partner EEE advertisement	1 = Link partner has EEE ability 0 = Link partner does not have EEE ability	RO

Note – there is no User field, this should be Reserved, and is addressed in another comment.

Modifications to the Standard for Slow Wake - **149.3.9.3**

Add a new sentence to the end of the 4th paragraph of 149.3.9.3 (P100 L13) as follows:

ALERT, a four RS-FEC frame long sequence, shall start at the beginning of any eighth PHY frame boundary starting at the beginning of the frame following a refresh PHY frame. This offsets the MASTER and SLAVE ALERT start times by alert_period/2 = 4 PHY frames and provides the following two benefits: The MASTER and SLAVE allowable ALERT transmissions do not overlap and ALERT does not overlap the device's own refresh. The MASTER and SLAVE shall derive the tx_refresh_active and tx_alert_start signals from the transmitted PHY frames (tx_rsfc) as shown in Table 149–4 and Table 149–5. When Slow Wake is active, Alert can be transmitted in only a single QR cycle location.

EEE Alert Active, section 149.3.5.1

Table 149–4—Synchronization logic derived from slave signal RS-FEC frame count

Slave-side variable	u=tx_rsfc
tx_alert_start_next = true	mod(u,alert_period) = alert_period/2 - 1 when Slow Wake = 0 mod(u,lpi_qr_time) = lpi_qr_time/2 + alert_period/2 - 1 when Slow Wake = 1

Table 149–5—Synchronization logic derived from master signal RS-FEC frame count

Master-side variable	v=tx_rsfc
tx_alert_start_next = true	mod(v,alert_period) = alert_period - 1 when Slow Wake = 0 mod(v,lpi_qr_time) = lpi_qr_time – 1 when Slow Wake = 1

Modifications to the Standard for Slow Wake

- Modify title of Figure 149-12 (defines Valid Alert start) as follows:

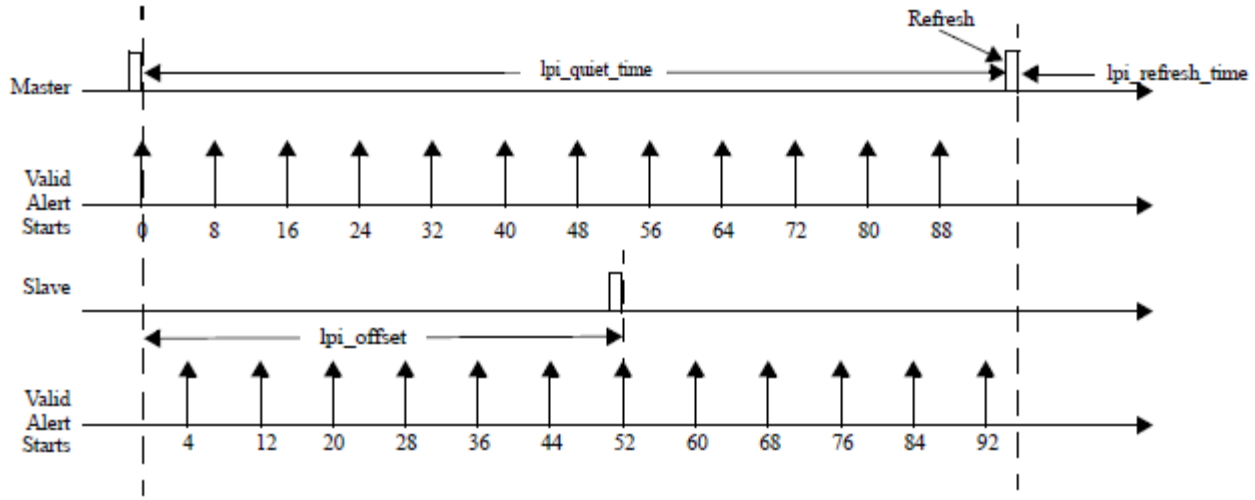


Figure 149–12—Timing periods for LPI signals when Slow Wake not active

Modifications to the Standard for Slow Wake

- Insert new figure after 149-12 for timing when slow wake is active
  - Change text references to LPI timing Figure 149-12 to say “Figure 149-12 and Figure 149-12a” (page 99 lines 11 and 51 – figure will be auto-numbered in the draft)

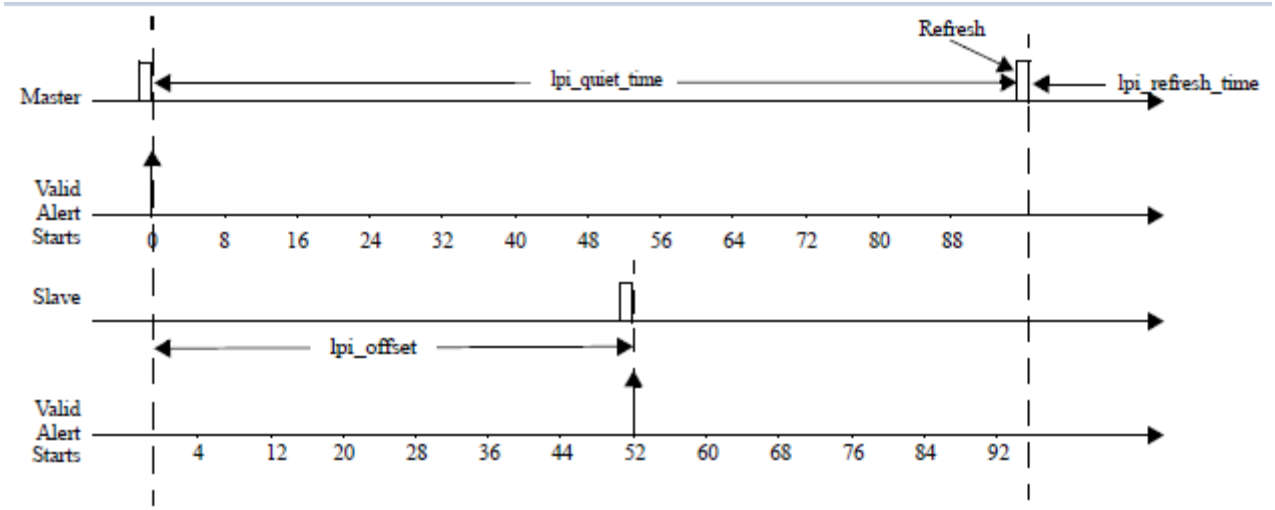


Figure 149-12a – Timing periods for LPI signals when Slow Wake is active

Modifications to the Standard for Slow Wake - **Clause 78**

- Add this paragraph to section 78.5 just above table 78-4
- Case-1 of the PHY in the MultiGBASE-T1 set applies when the PHY is requested to transmit the Alert signal before transmission of the Sleep signal to the Link Partner is complete. Case-2 of the PHY in the MultiGBASE-T1 set applies when the PHY is requested to transmit the Wake signal after transmission of the Sleep signal to the Link Partner is complete and if the PHY has not indicated LOCAL FAULT at any time during the previous 10 ms. Case-3 in MultiGBASE-T1 is the same as case-1 when Slow Wake is active. Case-4 in the MultiGBASE-T1 is the same as Case-2 when Slow Wake is active.

Table 78-4 additions

PHY or interface type	Case	*T*w_sys_tx (min) (uS)	*T*w_phy (min) (uS)	*T*phy_shrink_tx (max) (uS)	*T*phy_shrink_rx (max) (uS)	*T*w_sys_rx (min) (uS)
2.5GBase-T1	Case-1	35.84	35.84			10.24
	Case-2	25.6	25.6			10.24
	Case-3	158.72	35.84			10.24
	Case-4	148.48	25.6			10.24
5GBase-T1	Case-1	17.92	17.92			5.12
	Case-2	12.8	12.8			5.12
	Case-3	79.36	17.92			5.12
	Case-4	74.24	12.8			5.12
10GBase-T1	Case-1	8.96	8.96			2.56
	Case-2	6.4	6.4			2.56
	Case-3	38.4	8.96			2.56
	Case-4	35.84	6.4			2.56

Thank you.