Teranetics

Enhancements to the Low-Power Idle Mode

Contributors:

Dimitry Taich

Sanjay Kasturia

Supporters

- **ADD YOUR NAME HERE**

Overview

Teranetics

- LPI Concept
- Power Saving Analysis
- Synchronous Wake-up procedure
- Asynchronous Wake-up procedure
- Refresh rate re-negotiation method
- Summary
- Conclusions

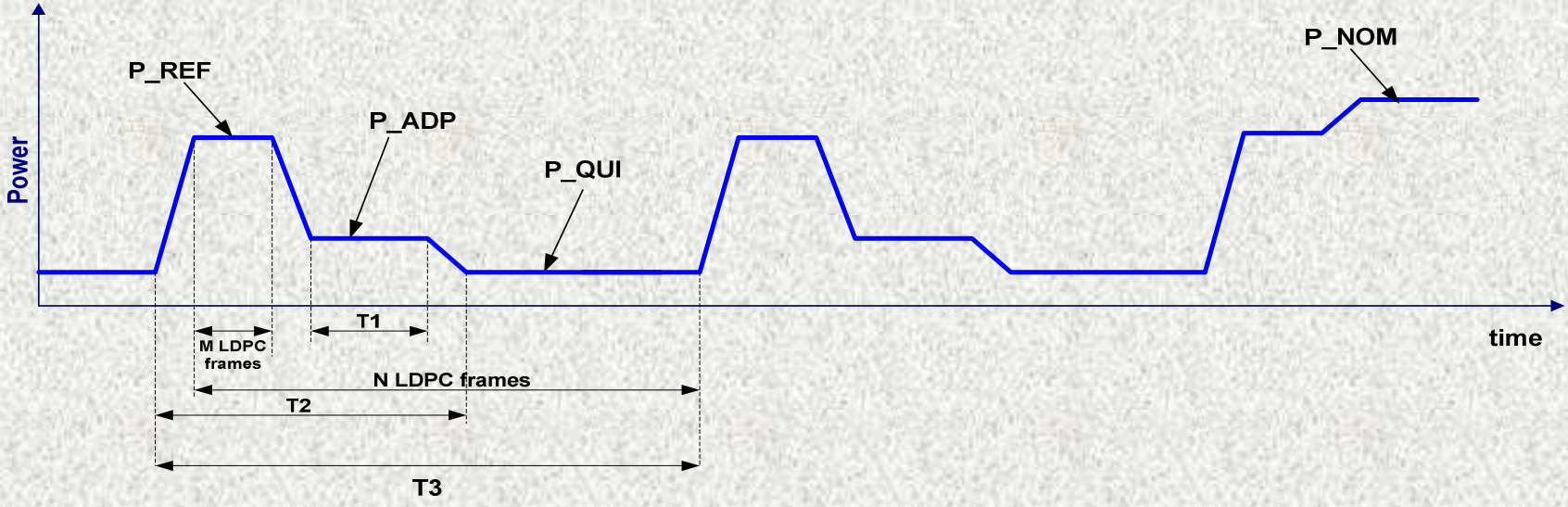
LP-IDLE Concept

Teranetics

- Save power by entering a low-power idle state when there is no data to be transmitted
- While at LP-Idle stage:
  - **All transmit and receive data path circuits can be turned off**
  - **All adaptive coefficients are saved and stored**
  - **Timing circuits free run with acquired frequency**
    - *Only fraction of the nominal power to be consumed*
  - **Periodically refresh local/remote timing so they remain locked**
  - **Periodically refresh all coefficients**
  - **PMA and PCS maintain synchronization**
    - *To enable fast return to full mode of operation*
- Current 10GBASE-T LP-idle proposal allows resumption of normal mode only at specific point in the super-frame
- MAC requests PHY to enter or exit LP_IDLE
  - **If the remote PHY initiates exit, local PHY immediately signals to the MAC**

Power consumption overview

- P_NOM – full 10G Power in normal operational mode
- P_REF - power during refresh idle; P_REF < P_NOM
- P_ADP – additional power required to complete coefficients update after ceasing transmitting and receiving R_IDLE; P_ADP << P_REF
- P_QIU – quiescent power when Tx/Rx off (leakage + timing circuit + mdio/etc); P_QIU << P_REF, P_QIU << P_NOM
- Full refresh cycle T2 takes P_REF + P_ADP + some overhead associated with rapid power switching;
- The target is minimizing power consumption over T3 period; because P_QIU << P_REF and P_ADP < P_REF the key is achieving N >> M.



Power Consumption Estimation - 1

- **P_REF = P_NOM – P_LDPC ≈ P_NOM**
  - **Assumes all interfaces buses are ON**
  - **Higher then in "parnaby_01_0108"; might reflect vendor-specific implementation**
- **P_ADP**:
  - **Majority of the data path circuits can be switched off**
  - **Since channel is very stable full coefficients metrics update can be spread among big number of refresh cycles thus further reducing power consumption per refresh cycle**
  - **Timing circuit might require frequent update but usually consumes very little power**

  P_ADP can be estimated as ~20% of P_NOM

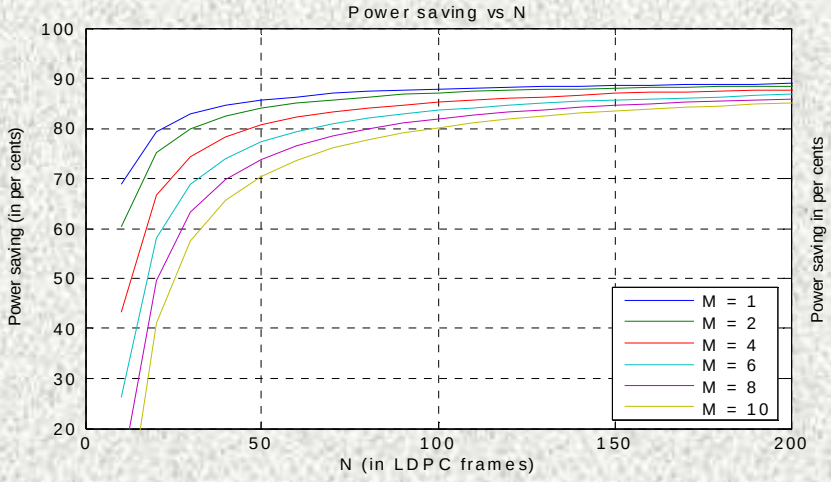
  T1 varies from 1 (N/M >> 100) up to 4 (N/M << 100)
- **P = P_REF*M + 4*P_ADP + Δ**
  - **Δ is overhead of power on/off switching;**
  - **According to our estimation, T2 – T1 – M_frames can be estimated as ≤300nsec ≈ 1 LDPC frame, thus associated power consumption can be approximated by P_REF/2**
- **P_QUI ≈ ~10% of P_NOM**
  - **Slightly lower then in "parnaby_01_0108"**

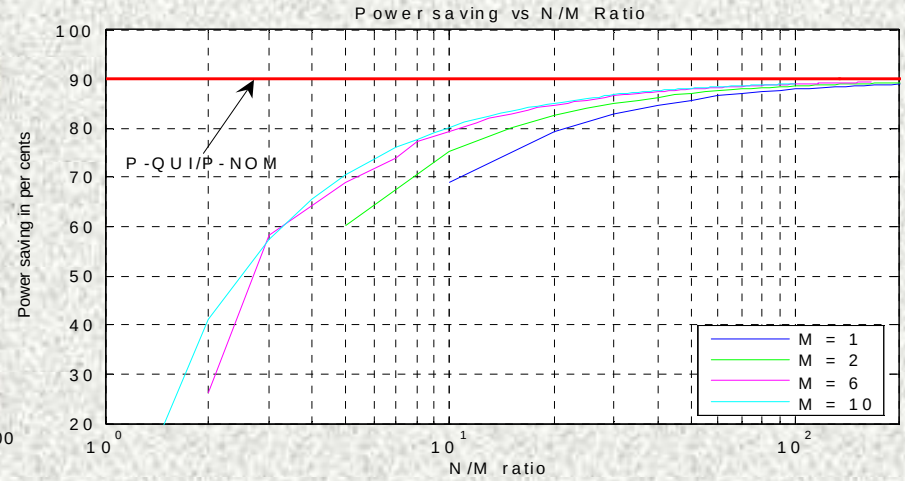
Power Consumption Estimation - 2

- **Thus Power consumption over entire T3 period can be approximated as**

  P_REF x M + P_REF/2 + P_ADP x 4 + P_QUI x (N-M-4-1/2)

  and should be compared to P_NOM x N; Obviously max power saving asymptotically converges to P_QUI/P_NOM (90%)
- **For M << N, overall power consumption is dominated by P_QUI**
  - **N/M = 10 allows better then 75% of possible power saving (68% of absolute power saving)**
  - **N/M = 20 allows better then 87% of possible power saving (79% of absolute power saving)**
  - **N/M = 100 allows better then 97%(!) of possible power saving (88% of absolute power saving); not much energy left to go after…**
- **From PHY perspective M/N ratio larger then 100 brings very small value**





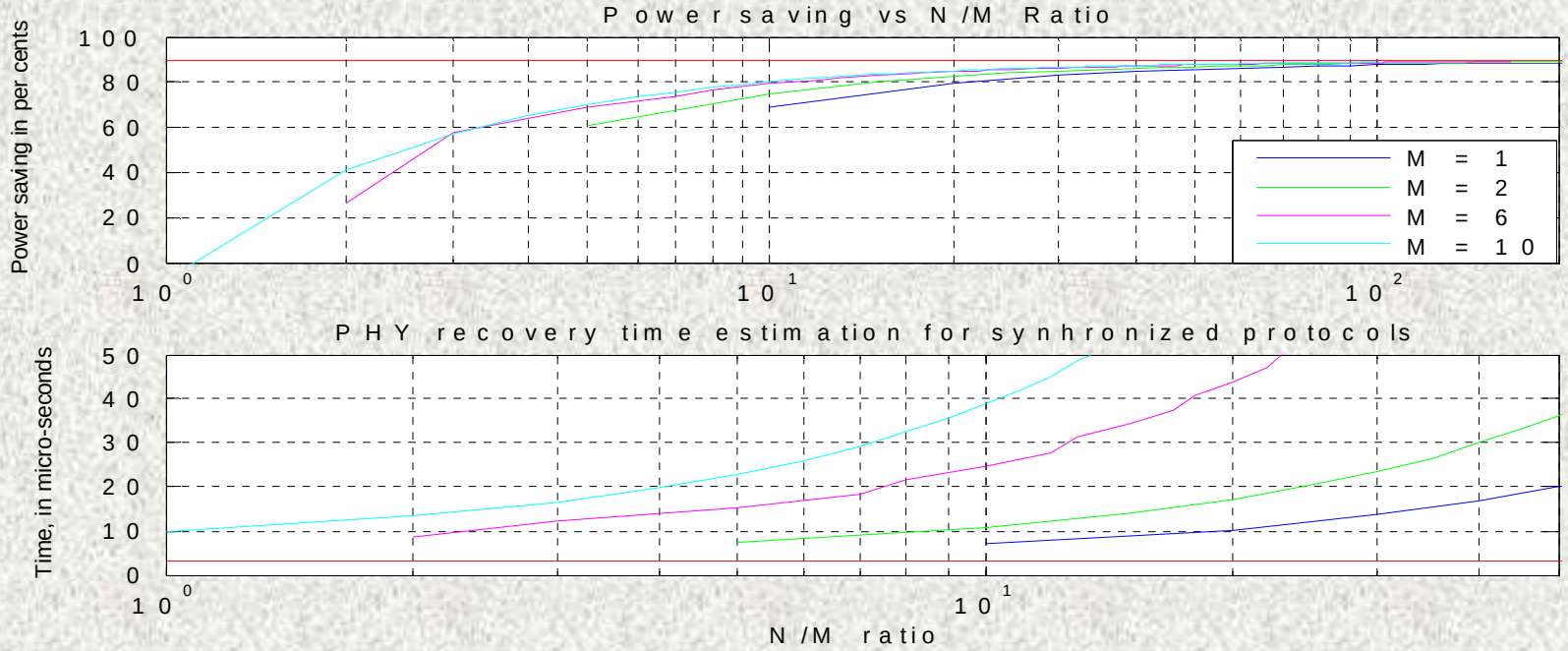
Recovery time - Sync

Teranetics

- Option 1 (synchronized) : By sending pre-defined digital stream
  - Will Require PCS layer to be ON during LP_idle stage
  - Max Next_frame is limited by N LDPC frames
- Option 2 (synchronized): By using auxiliary bit as a communication channel:
  - “0” means stay in LPI mode
  - “1” means wake up;
  - repetition decoder can be applied to prevent false-alarm wake-up
    - Will Require PCS layer to be ON during LPI stage
    - Max Next_frame is limited by N LDPC frames
- Both options allow simple wake-up/power down schedule implementation using sync counters in local and remote PHY’s
- Worst-case latency is dominated by N parameter (time-slot between two refresh frames); thus if latency dominates the choices made, max power saving might not be achieved

Recovery Time Estimation – Sync

- Basically limited by I/O synchronization and PHY latency < 3µsec
- Two main parameter to optimize are worst-case waiting time to next frame (T_NF) and PHY bring-up to normal operational mode.
- For synchronized protocols T_NF is N x 320nsec → suggests that schemes with N less then 20 should be used to meet 10µsec recovery time
  - Feasible but might not provide max power saving

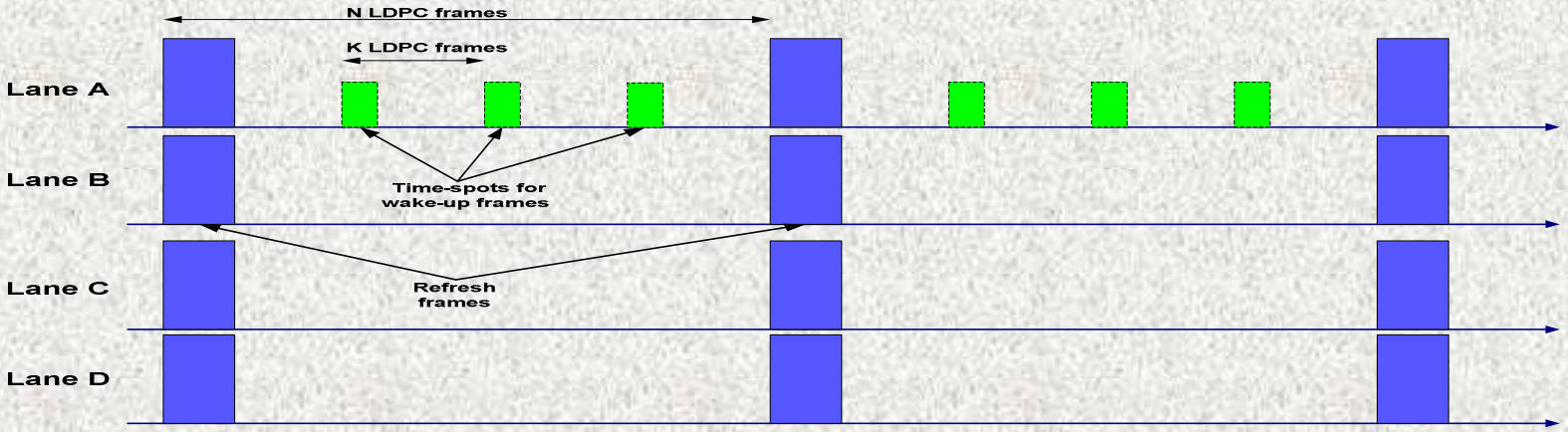


Recovery Time – Sense and Wake

Teranetics

Option 3 (with sense and wake): Defining more "sense" time slots within the super-frame and allow sending wake-up signal in any of these.

- **Simple signal detect circuit can be activated during LPI stage with negligible effect on P_QUI**
  - **Additional power is estimated as 100mW per lane**
- **Time-to-wakeup after the energy is detected may be a few LDPC frames but this will still provide significant reduction in maximum latency**
- **Once wake-up frame is detected local PHY gets ready to receive data frame after L cycles**
  - **Similar as if PHY would be awaken during refresh frame**
- **To prevent collisions half-duplex protocol can be used – say Master on Lane A, Slave on Lane B**
  - **Good for Power Saving too**
- **K parameter can be fixed (say 10 LDPC frames) or negotiated during AN stage**



Real-time refreshing rate adjusting

Teranetics

- **"Portions of MAC, NSE, Memory may be powered down, but will require a finite time (N_IDLE) to resume operation. *This time is not a constant that can be negotiated at link startup as it may be dependent on aggregate system utilization however a minimum may be negotiated at link startup. A means for negotiating N_IDLE is necessary to balance outbound buffering (of source) against inbound N_IDLE requirements (of receiver)"* – see dove_01_0108**
- **Previous analysis showed an advantage of (M,N) pair optimizing from power saving and max latency perspectives**
- **(N,M) can be re-negotiated without new AN process using auxiliary bit as in-band channel information.**
- **Overall channel throughput is ~3MHz (1/320e-9)**
- **Since auxiliary bit is LDPC encoded additional coding is not mandatory**
  - **Requires LDPC decoder to be turned "ON" for refresh cycles**

Information channel using auxiliary bit

Teranetics

- **Range of supported N and M to be established during AN**
- **During AN MAC's select refresh rate setting based on available common denominator and *expected* traffic load**
- **Following traffic change (day/night time, for example) N & M can be optionally re-negotiated**
  - **Can be implemented during Data or Normal Idle Transmission**
  - **Does not require any high-layer changes – can be entirely implemented at PHY layer.**
- **Two-bytes protocol example: 2x8 = 16 consecutive LDPC frames are used - ~5µsec Word-length**
  - **1st nibble: Start of frame delimiter (SFD – for example 0xA)**
  - **2nd nibble: Requested (M,N) pair – selected from AN'ed list of supported values (MN_ID)**
    - K can be added to list of parameters if needed
  - **3rd nibble: Acknowledge of received N and M parameters (ACK)**
  - **4th nibble: End of frame delimiter (EFD – for example 0xF)**
  - **Transmit "0" when no information to be sent**
- **Same 2-bytes Word is sent until Acknowledge is received;**
  - **repetition decoder can be optionally applied**
  - **Other coding options are available but probably overkill**

Summary

Teranetics

- Achievable power saving for LP_idle mode was analyzed and estimated based on currently existing technology
  - Good correlation with “parnaby_01_0108”
- We proposed scheme to define “sense” frames in addition to refresh frames that allows to achieve nearly max power saving without compromising the worst-case latency parameter
- We proposed using auxiliary bit as an information channel allows to adjust refreshing and wake-up rate (N,M, optionally K as as well) without link re-start
- We propose defining following (M,N) pairs – in LDPC frames:
  - (2,40), (2,80), (2,200), (2,500);
  - (4,80), (4,200), (4,500);
  - (8,200), (8,500);
- We propose to support following K (wake-up rate): 10

Conclusions

Teranetics

- LPI approach was presented and investigated from PHY perspective
- Wake-up methods were briefly discussed; “sense and wake” method was presented to improve PHY recovery time
- Method for time on/off parameters re-negotiation without link restart was suggested
- Close to 10x power saving in LPI mode is feasible without PMA/PCS layers changes and under real-life conditions
- 10µsec wake-up time is achievable