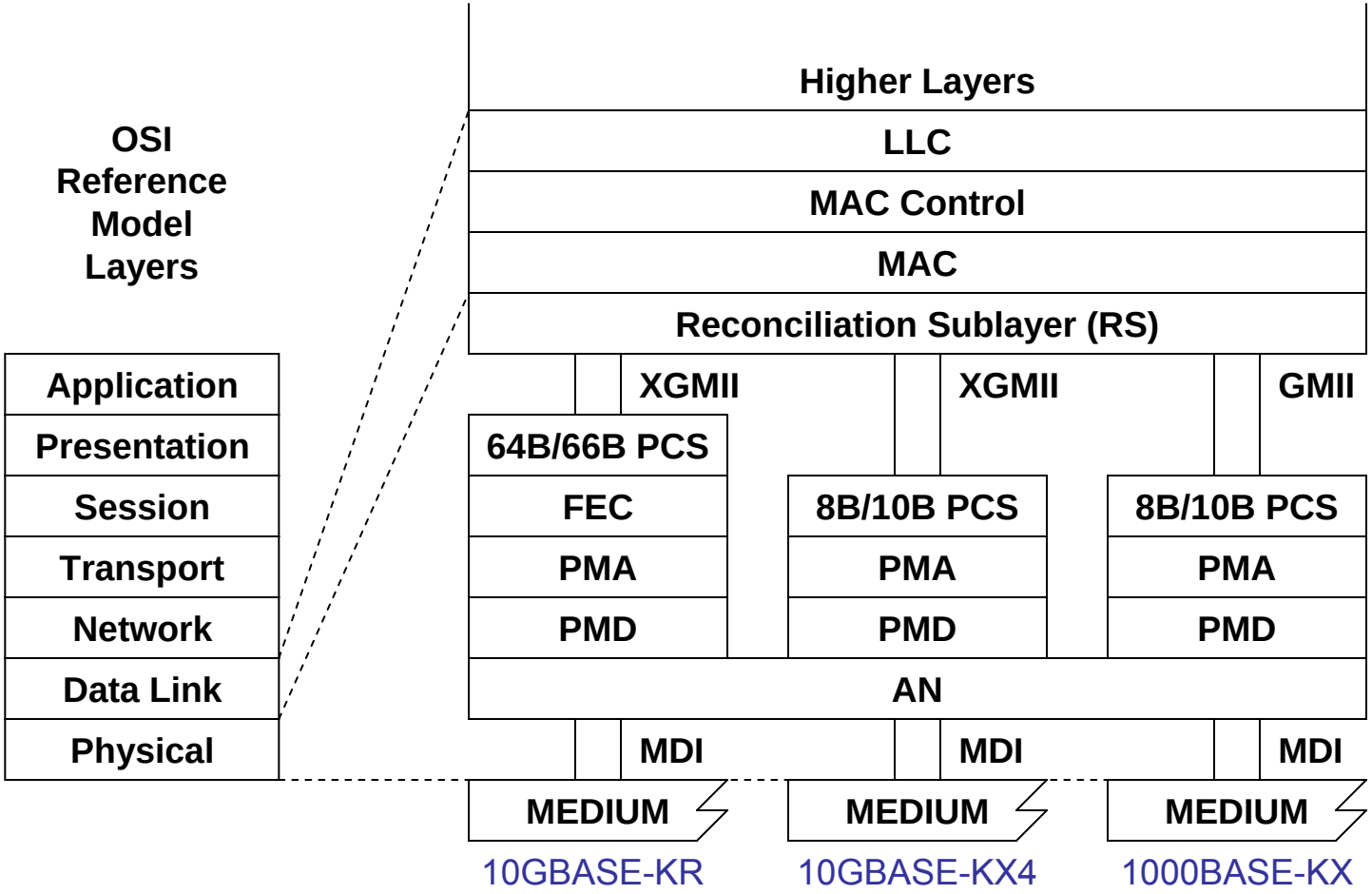
Backplane Ethernet Low-Power Idle Baseline Proposal

David Koenen, HP

July 2008

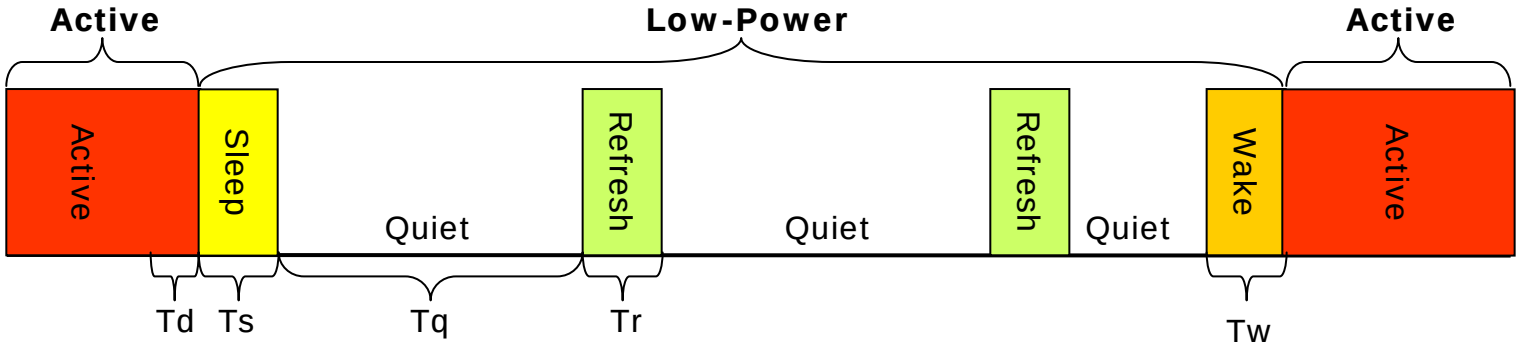
Backplane PHYs



Backplane LPI

- Transmitter goes pre-determined level during quiet state
  - Similar concept to PCIe specification L2,L3 states
- Refresh operation
  - KR to use training mode during refresh cycle
  - KX4 and KX to use set idle pattern
  - KR FEC not refreshed
- Sleep transition
  - Fairly simple after the request is indicated
- Wake transition
  - 10GBASE-KR has greater latency due to coding scheme & FEC
  - KR is multi-staged to decrease synchronization requirements
  - KX4/KX use wake pattern to re-sync the link

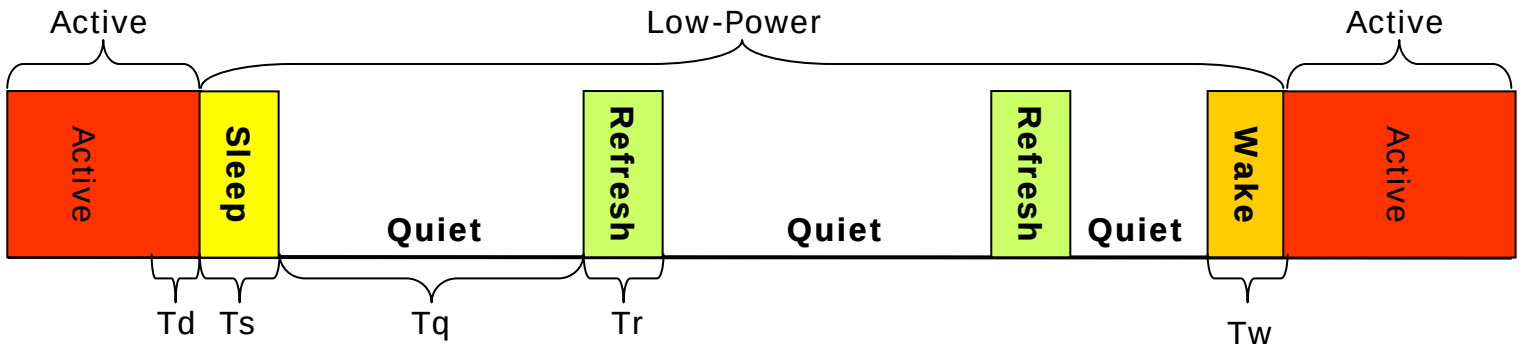
LPI Operating States



Term	Description
Active state	Existing state used for data transmission where either data packets or IPG/Idle symbols are transmitted.
Low-Power state	New state used during periods of no data transmission to allow system power reduction between data packet bursts.

- No change to this concept with backplane LPI

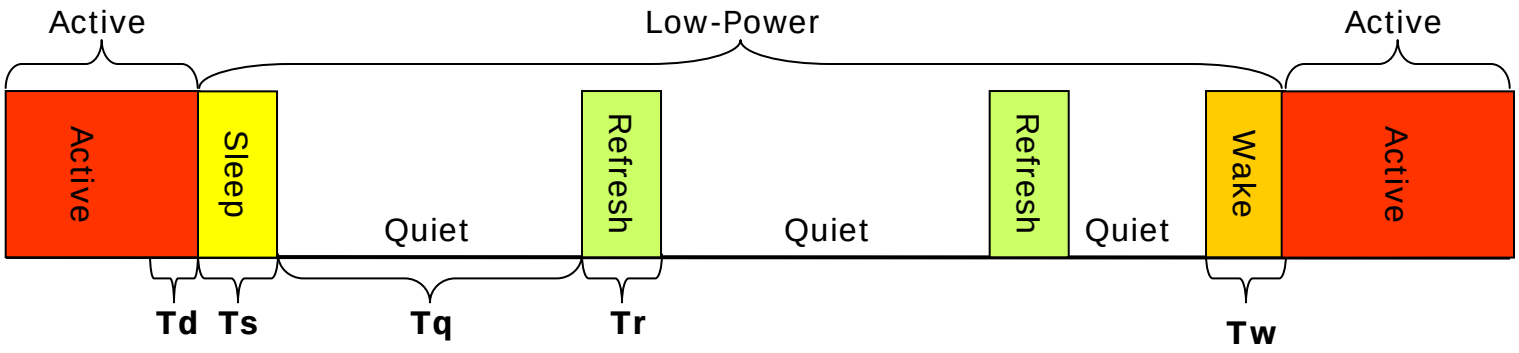
LPI New Line Signals



Term	Description
Sleep	Signal to inform link partner of entry into low power state
Quiet	Steady state, no transitions
Refresh	Periodic signal during low power state for PHY to maintain timing recovery and/or coefficients
Wake	Signal to inform link partner of entry back into active state

- Same concept as with BASE-T technologies

LPI Timing Parameters

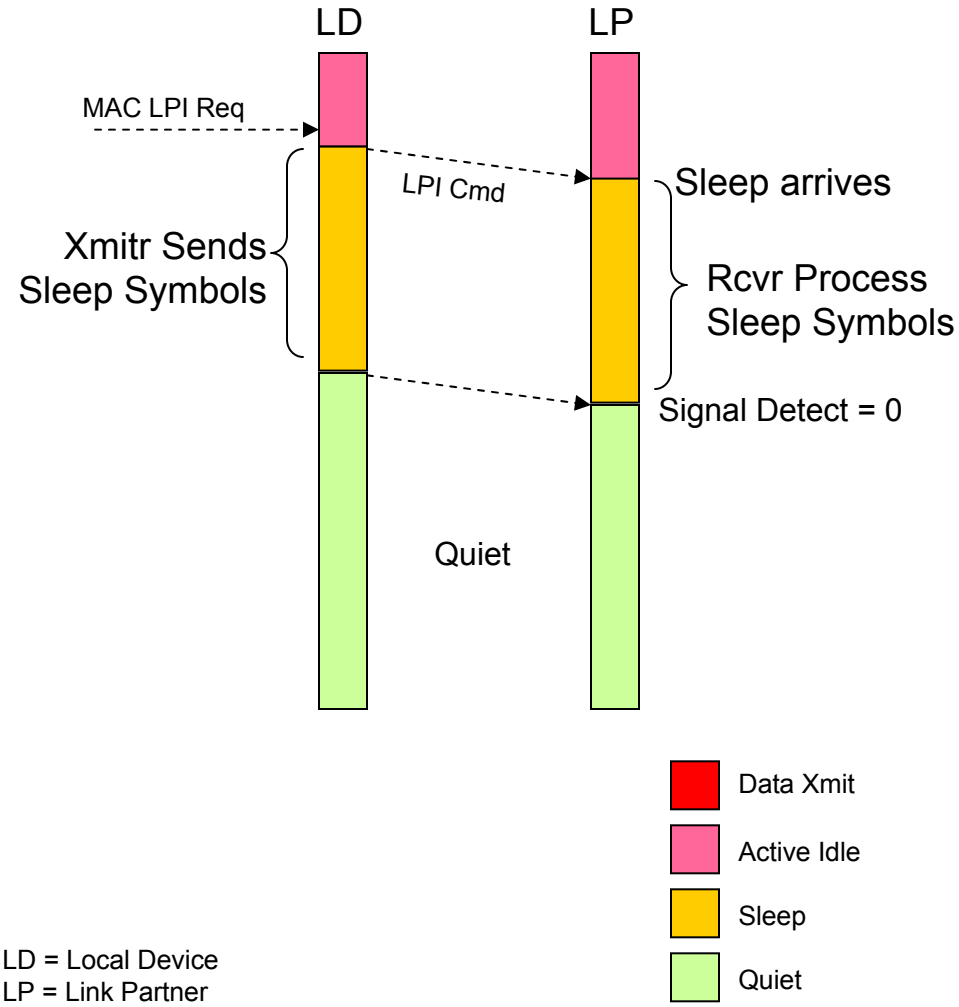


Term	Description
Decision Time (Td)	Higher-layer control policy timing; out of scope
Sleep Time (Ts)	Min. duration Sleep symbols sent before going to Quiet
Quiet Duration (Tq)	Max. duration PHY remains Quiet before Refresh
Refresh Duration (Tr)	Min. duration PHY sends Refresh symbols
Wake Time (Tw)	Max. period to permit the receiving system to wake up
Propagation Delay (Tp)	Max. transmission delay of the media

- No change to this concept with backplane LPI

Entering Quiet Mode

- Pass message to indicate transition to quiet
  - Ordered set or special code with KR/KX4, special code with KX
- Transmitter held at 0mV +/- TBD mV differential voltage



LD = Local Device
LP = Link Partner

Refresh Cycle

- KR
  - Use of training mode
  - Cell 14 of status report field changed from Reserved to be Refresh
- KX4
  - Use of KAR idle pattern only on lane 0
  - Lack of signals on lanes 1-3 indicates a refresh state
- KX
  - Per dove_02_05_08.pdf, slide 13

/LPI/	Low Power Idle		Correcting /LI1/, Preserving /LI2/
/LI1/	Low Power Idle 1	2	/K28.5/D6.5/
/LI2/	Low Power Idle 2	2	/K28.5/D26.4/

Use of KR Training Frame

- Data header of 36 octets (=288 UI) which has low baud rate
  - Does not need phase lock to decode
- Training pattern of 512 octets (=4096 UI) *at line rate*
  - Enables clock recovery and receiver adaptation
- RX PHY detects the training pattern
  - Anticipates its boundary
  - Triggers the framing/FEC synchronization
- RX specifies min # of training frames that TX must send
  - The transmitter can send a final training frame (with a special flag)
  - Switch to active idle.

* Thanks to Adee Ran of Intel

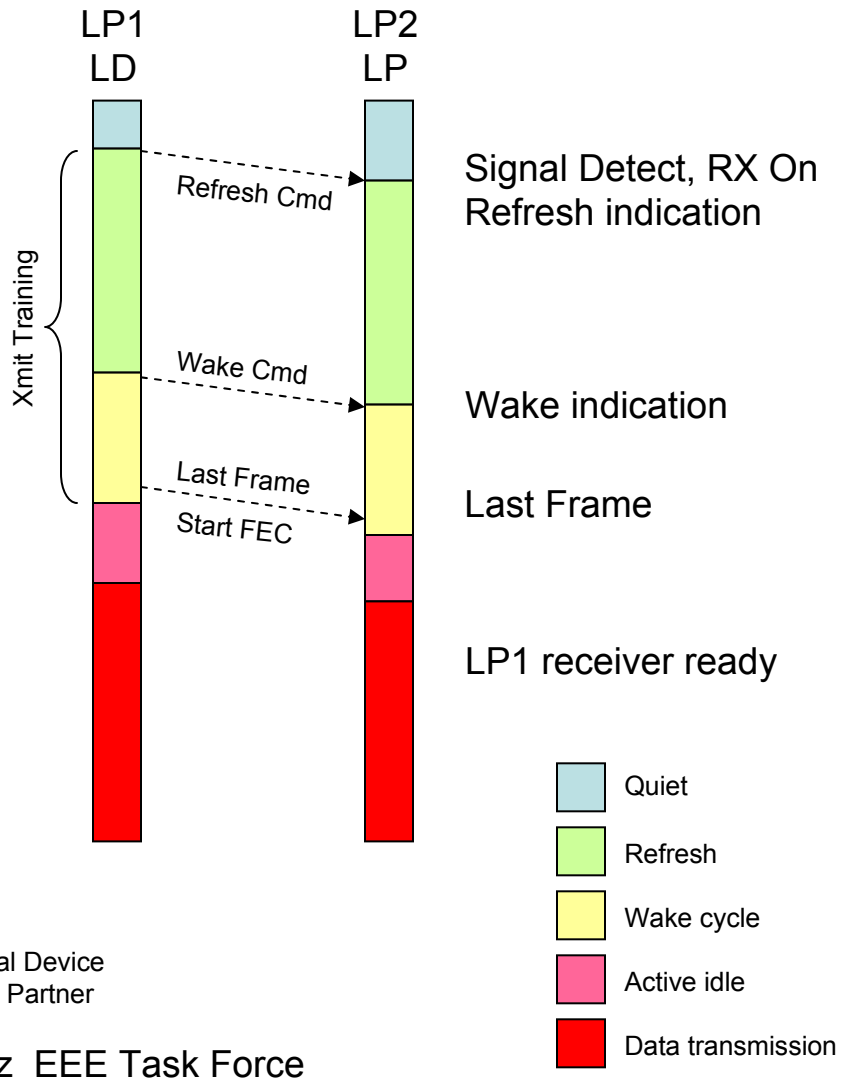
KR Operation

- Training frame length is 4384 UI ≈ 0.5 microsecond
- Wakeup flow should include some time to wake up the receiver circuitry:
  - Initiator turns on TX (and possibly RX) circuits, and starts transmitting (t=0)
  - Training frames activate RX signal detector and indicate wakeup or refresh (TBD us)
  - Additional frames required by the target+initiator for Rcvr Clock Sync for TBD # of Training Frames (# * 0.5us)
  - For local device sends training frame with “last training frame” flag TBD. (0.5 us)
  - Transition to active mode
- Total wakeup time should be in range of 5-10us
- Update w/o wakeup
  - Initiator need not turn on its RX
  - Target need not turn on its TX
  - Initiator sends a fixed number of frames (TBD)

* Thanks to Adee Ran of Intel

KR Wake

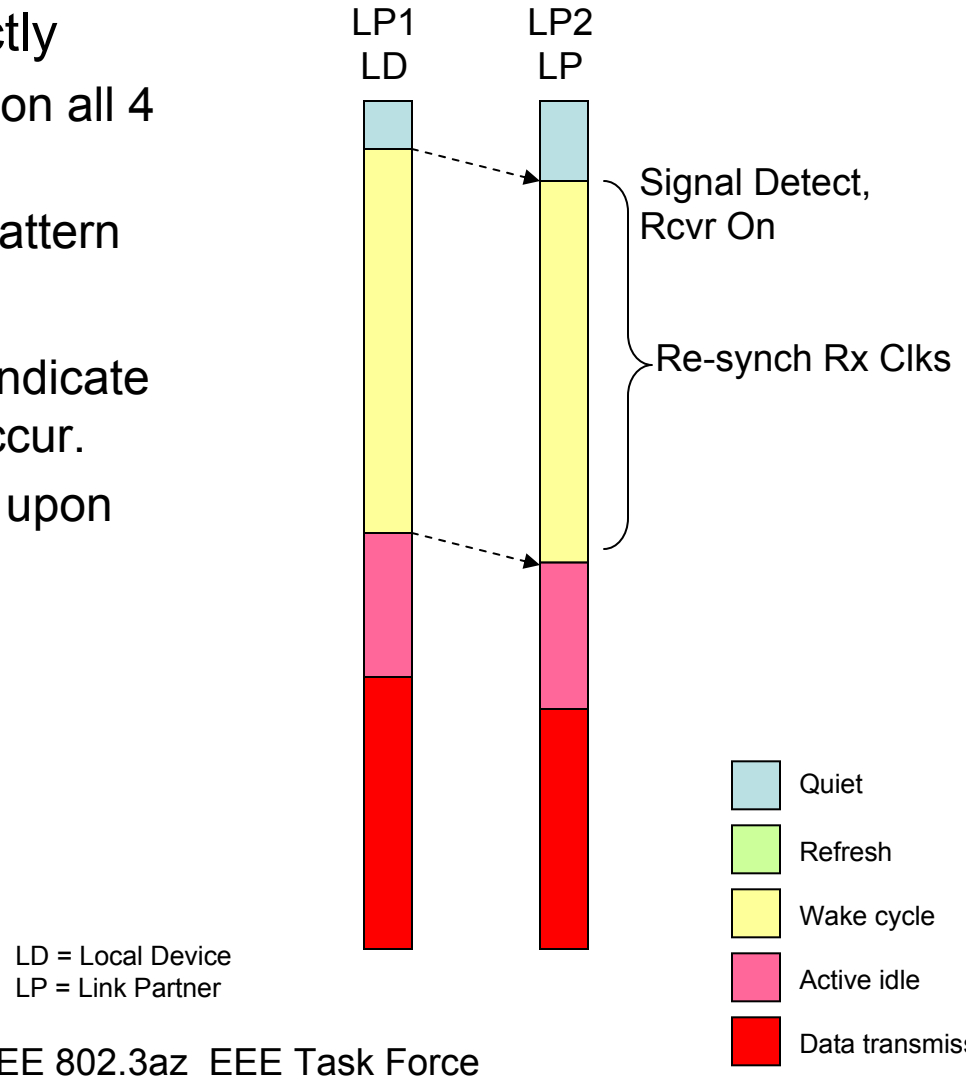
- Enter Refresh Cycle first
  - First transmitted training frame activates LP Signal Detect wakes Rcvr
  - After Refresh period, transition to Wake
- Wake Cycle
  - Add Wake to cell 13 of status report field
  - Change to TX & RX Active after TBD number of frames
- Known transition boundary
  - Simplifies FEC & coding sync



LD = Local Device
LP = Link Partner

KX4/KX Wake Cycle

- Enter Wake Cycle directly
  - KX4 uses idle pattern on all 4 lanes
  - KX uses normal idle pattern
- Ordered sets
  - KX4 could use RF to indicate when synch did not occur.
  - Otherwise, just based upon time
- No FEC complications
  - Yippee!



LD = Local Device
LP = Link Partner

Asymmetric Operation

- The Backplane PHY will support Asymmetric operation
  - The Backplane PHY are all dual simplex.
  - The TX -> RX pair can go in and out of LPI independent from the other direction.
- Refresh
  - LP periodically enters refresh cycle

Summary

- LPI can be applied to backplane Ethernet PHYs
  - Training mode for KR
  - Set idle patterns for other PHYs
- Asymmetric support
  - Use of ordered sets in 10G (KX4 and KR)
  - Use of special idle codes for 1G (KX)