IEEE P802.3az
Energy Efficient Ethernet and Time Synchroniztion

Mike Bennett
mjbennett@lbl.gov

Presented at the TSSG
May 1, 2009
Quebec City, Quebec, CA

Acknolwedgements

- I “borrowed” liberally from slides created by
  - David Law
  - Wael Diab
  - Rob Hays

- Thanks to David, Wael and Rob for their contributions to the 802.3az task force!

Discussion

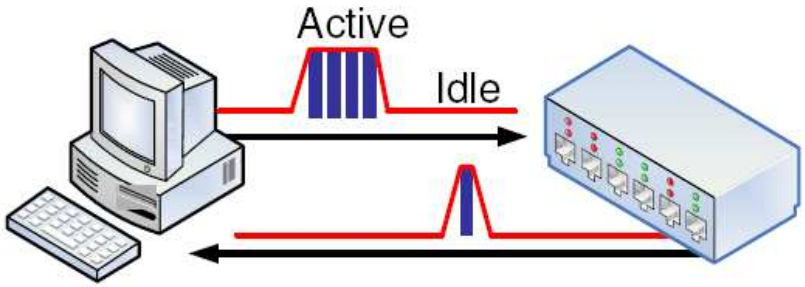
- Overview of Energy Efficient Ethernet (EEE)
- Things to consider

Briefly, what is EEE?

- EEE is a method to facilitate transition to and from lower power consumption in response to changes in network demand
  - In the process of being defined for these copper PHYs
    - 100BASE-TX (Full Duplex)
    - 1000BASE-T (Full Duplex)
    - 10GBASE-T
    - 10GBASE-KR
    - 10GBASE-KX4
    - 1000BASE-KX
  - Uses Low Power Idle (LPI) to save energy

What is Low Power Idle?

- Concept: Transmit data as fast as possible, return to Low-Power Idle
- Saves energy by cycling between Active and Low Power Idle
  - Power reduced by turning off unused circuits during LPI
  - Energy use scales with bandwidth utilization

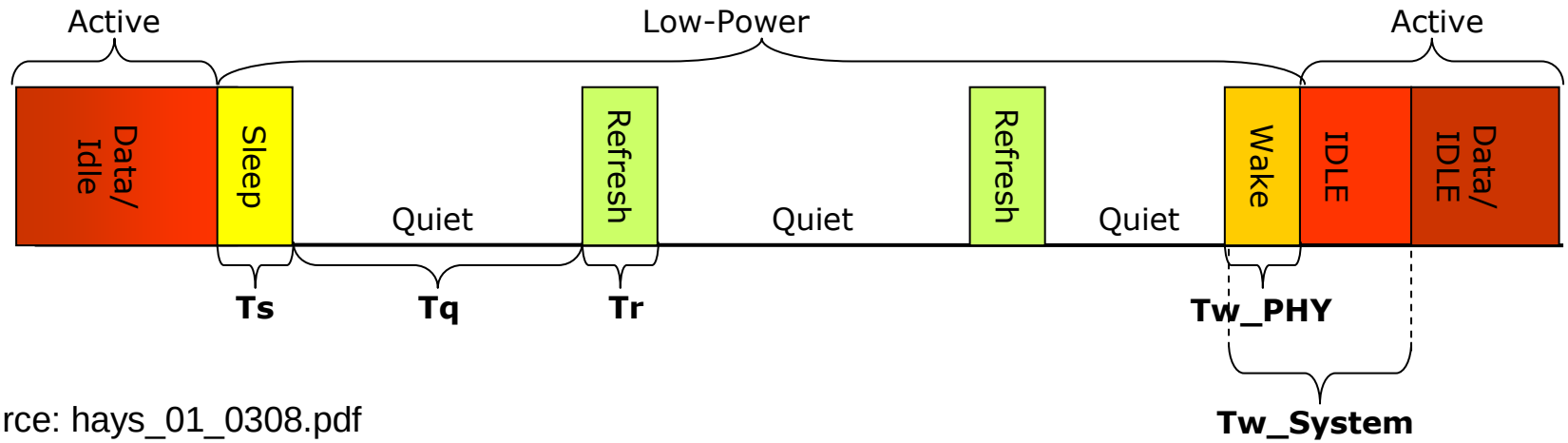


Source: hays_01_0308.pdf

What is Low Power Idle?

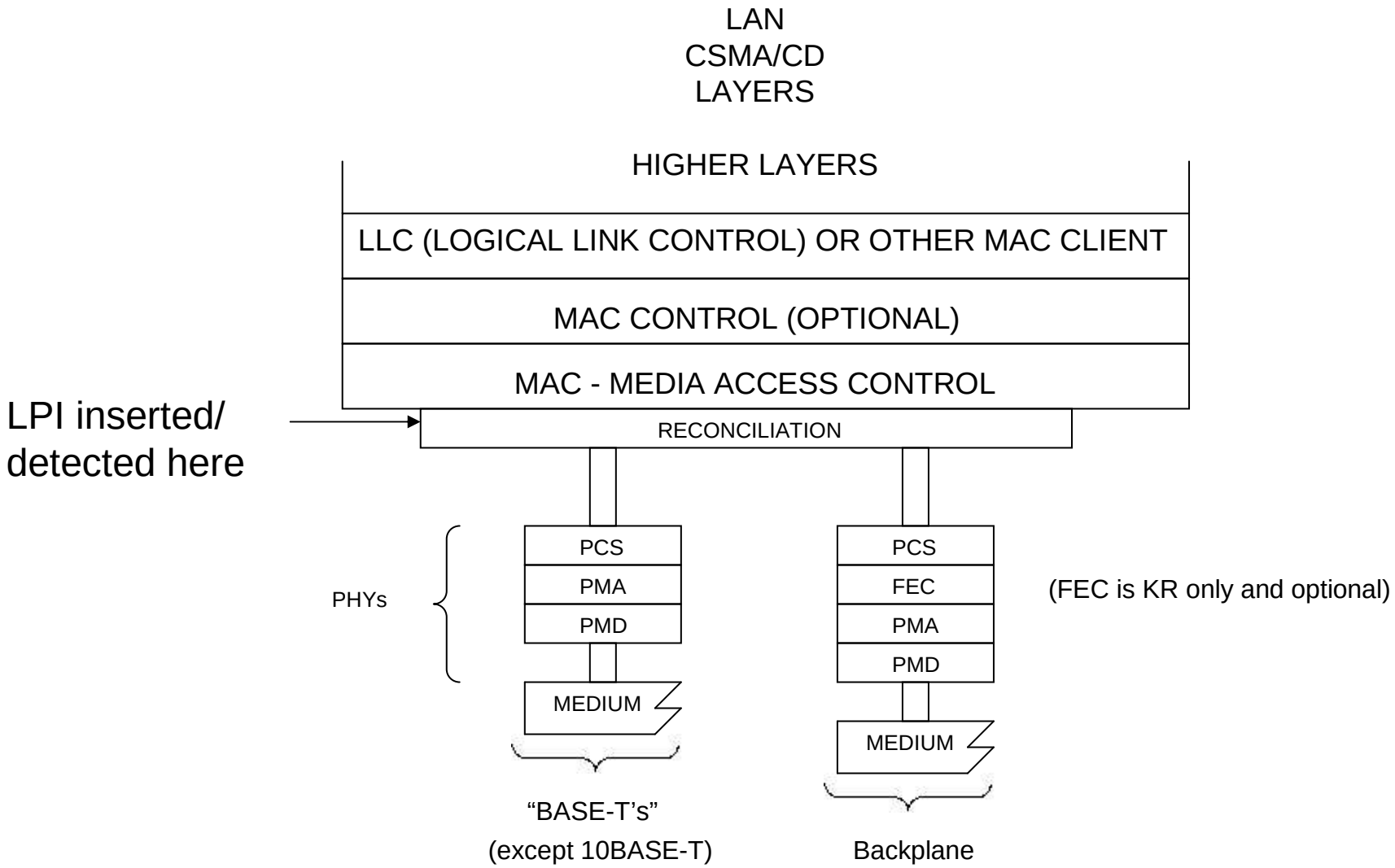
- Definitions

Term	Description
Sleep Time (Ts)	Duration PHY sends Sleep symbols before going Quiet.
Quiet Duration (Tq)	Duration PHY remains Quiet before it must wake for Refresh period.
Refresh Duration (Tr)	Duration PHY sends Refresh symbols for timing recovery and coefficient synchronization.
PHY Wake Time (Tw_PHY)	Duration PHY takes to resume to Active state after decision to Wake.
System Wake Time (Tw_System)	Wait period where no data is transmitted to give the receiving system time to wake up.

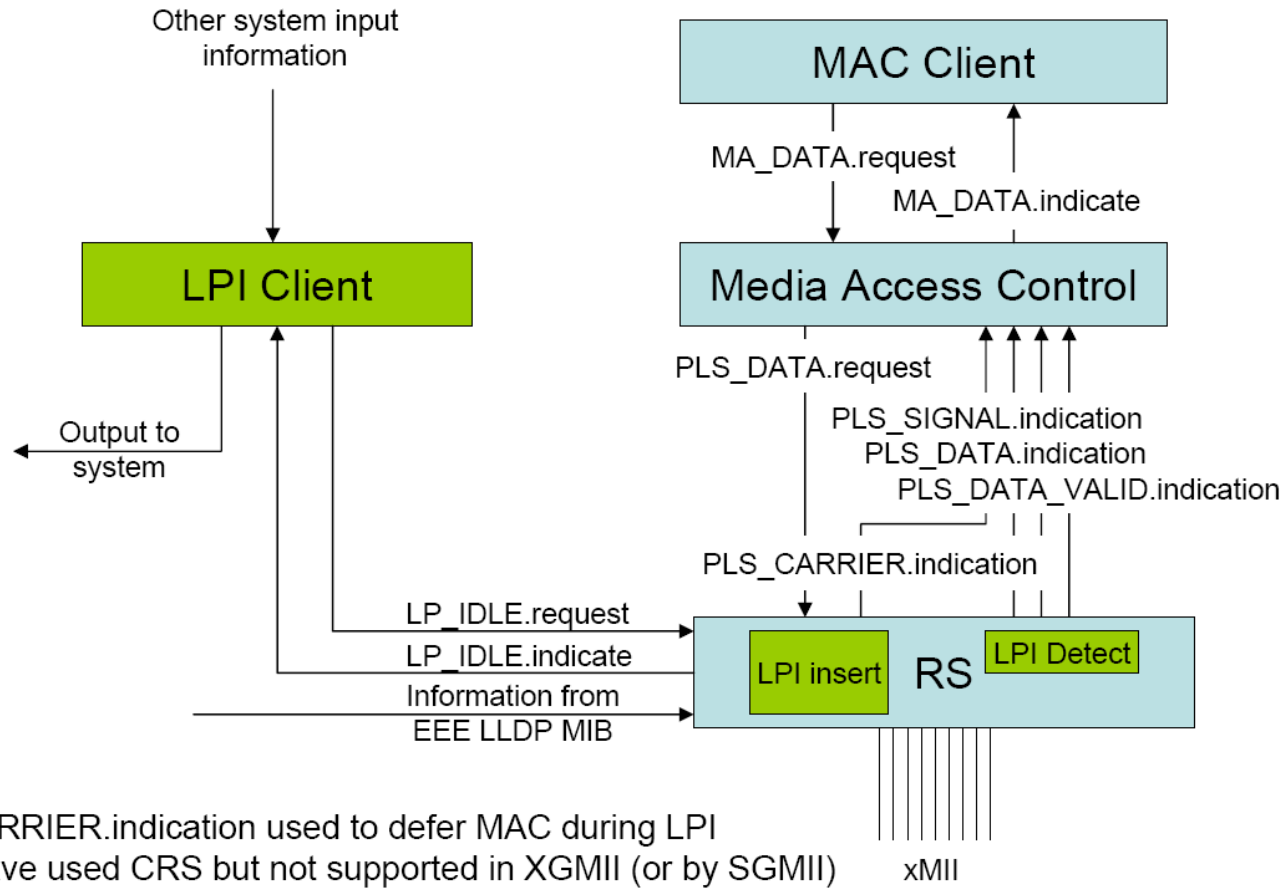


Source: hays_01_0308.pdf

What is Low Power Idle?



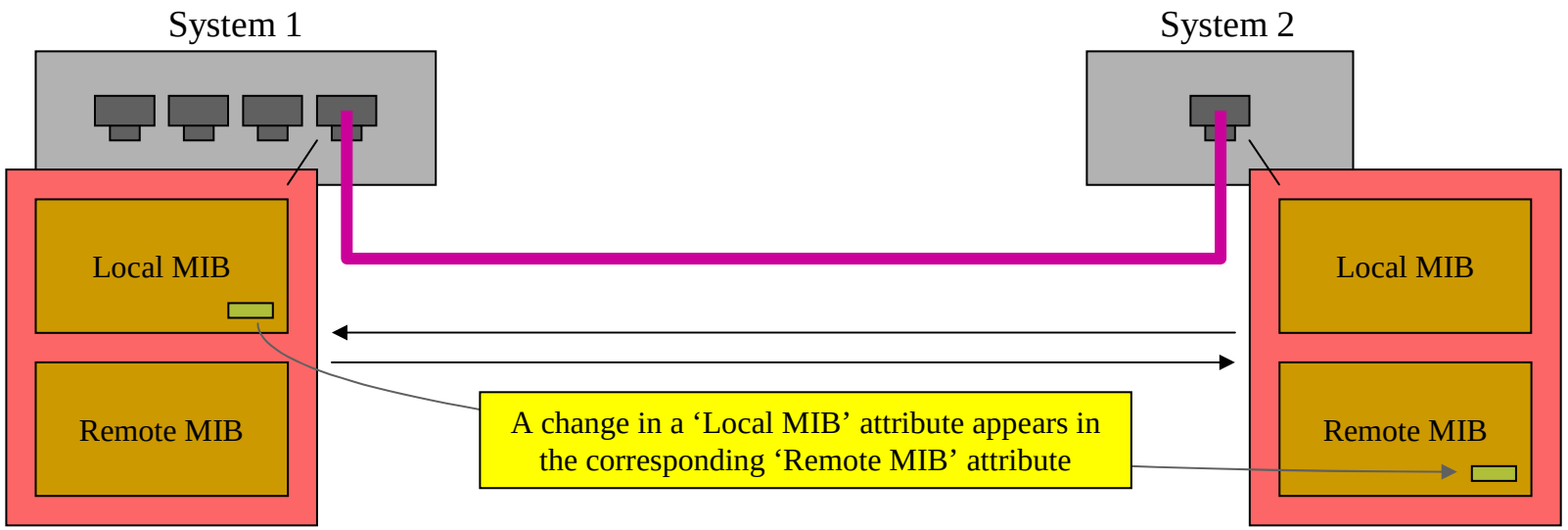
What is Low Power Idle?



Source: law_01_1108.pdf

Link partner communication

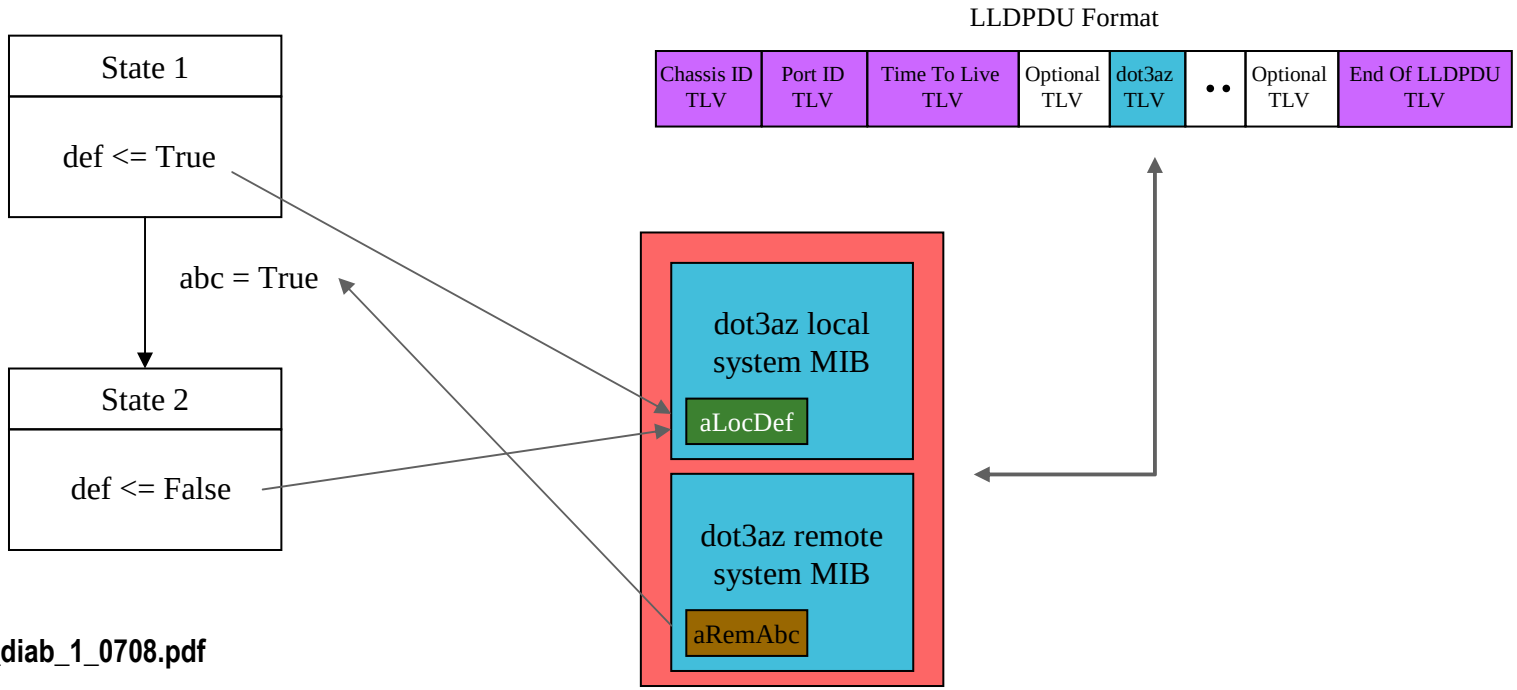
- Use auto-negotiation to notify link partner of EEE capabilities
- Use LLDP after the link is up
- Operates over a point to point link
- Completely enclosed protocol
  - We define data, it gets transported
    - We don't get to make changes to the protocol
- Data in ‘Local MIB’ transported to ‘Remote MIB’
  - Transported by TLVs (type, length, value)



Source: law_01_0508.pdf

LLDP and State diagrams

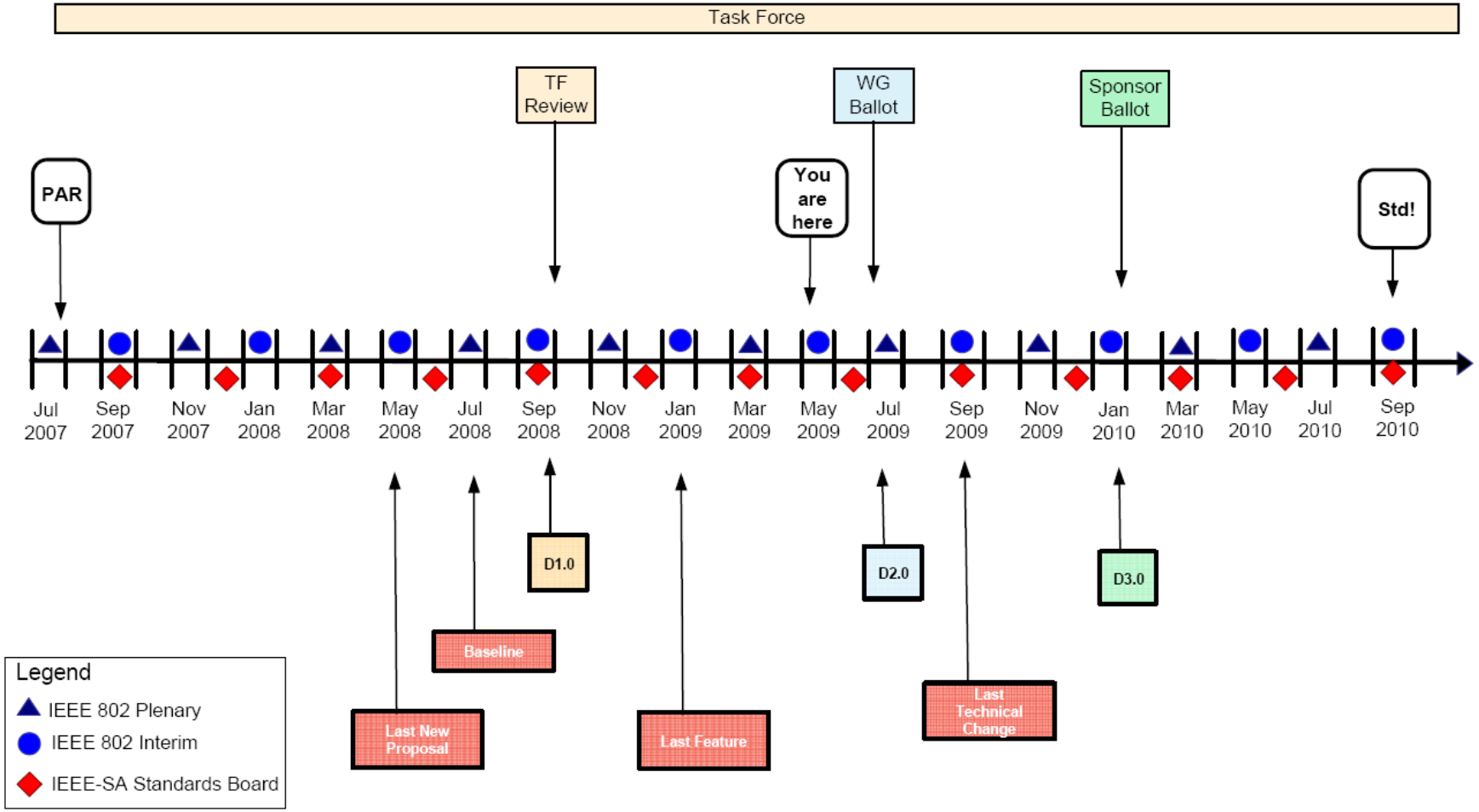
- Can't map directly to TLV contents
  - Map through objects in dot3az local and remote MIB
  - Define MIB attribute to variable mapping
  - Allows .3 layers to take action based on variable changes



Source: joint_diab_1_0708.pdf

Current status

- We are preparing Draft 1.4 in the Task Force



Things to consider

- Latency introduced by EEE
  - Sleep and Wake
  - When at Low Power Mode, PHY device is not available immediately for data transmission request.
  - The max PHY recovery time Tw is different for different PHY types and can be increased by systems using L2
  - Architecture of 802.3az uses carrier deference to indicate availability of medium
- TSSG may be able to use this information to help mitigate latency introduced by Low Power Idle

Thank You!