EEE Capabilities Negotiation Proposal
Revision 2

May 2008, Munich

Robert Hays
Intel Corporation

Energy Efficient Ethernet

Thanks to the Following Contributors

- Hugh Barrass (Cisco)
- Niall Fitzgerald (LSI)
- Adam Healey (LSI)
- Brian Murray (LSI)
- Jacobo Riesco (LSI)
- Mario Traeber (Infineon)
- Aviad Wertheimer (Intel)

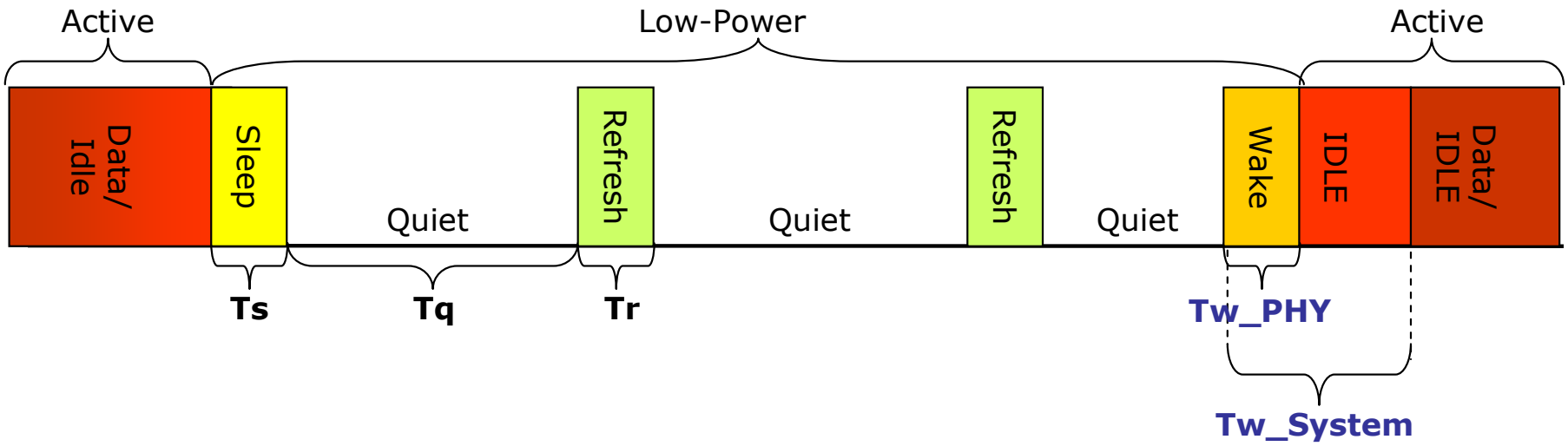
Take Aways from Hays_01_0508 Discussion

1. We have consensus on the Negotiation framework
   - Auto-Neg for advertising EEE capabilities & PHY parameters
   - LLDP for updating dynamic system parameters
2. There are two Tw parameters: Tw_PHY & Tw_system
   - We should to consider whether Tw_PHY is fixed or negotiated
   - Tw_system should be an asymmetric parameter, e.g. different for Rx & Tx, which would provide power savings and simplify negotiation
3. Advertising PHY parameters in the LLDP frame would allow future fiber optic EEE support

Note: Technical changes from Hays_01_0508 highlighted in blue text

EEE Timing Parameters

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.



EEE Negotiation Framework

- Auto-Negotiate EEE capabilities and PHY parameters (required)
  - Ts specified for each PHY
  - Tq, Tr combinations advertised as PHY refresh cycle
    - "Reduced energy" Tq:Tr duty cycle = n:1
    - "Lowest energy" Tq:Tr duty cycle > n:1
    - PHYs advertise highest Tq:Tr ratio supported and resolve to lower value
  - Tw_PHY negotiated or specified value TBD
    - Tw_PHY_{max} to be specified as required value
    - $Tw\_PHY_{max} = Tw\_system_{min}$
    - PHYs advertise lowest Tw_PHY supported and resolve to higher value

- LLDP used to change system parameters (optional)
  - Tw_system_{min} and Tw_system_{max} to be specified as boundary conditions
  - Independent values for Tw_system for Rx and Tx
  - Link partners exchange Tw_system Rx request and Tx capability and resolve to lower value for each

EEE Auto-Neg Extension (1 of 3)

Use Message Next Page (Page 0) with reserved 0x0A message code for EEE Technology Next Page Message code (Annex 28C and Annex 73A).

Message Count#	M 10	M 9	M 8	M 7	M 6	M 5	M 4	M 3	M 2	M 1	M 0	Message Code Description
10	0	0	0	0	0	0	0	1	0	1	0	EEE Technology Message Code. EEE capability 2 unformatted next pages to follow.
11 …	0	0	0	0	0	0	0	1	0	1	1	Reserved for future Auto-Negotiation use

EEE Auto-Neg Extension (2 of 3)

Use Unformatted next page (Page 1) to define IEEE 8023.az support.

Bit	Bit definition
U10	Next page - 1
U9:U7	Reserved, transmit as 0
U6	10GBASE-KR EEE support (0 = no, 1 = yes)
U5	10GBASE-KX4 EEE support (0 = no, 1 = yes)
U4	Reserved, transmit as 0 (for future 1000BASE-KX)
U3	10GBASE-T EEE support (0 = no, 1 = yes)
U2	1000BASE-T EEE support (0 = no, 1 = yes)
U1	100BASE-TX EEE support (0 = no, 1 = yes)
U0	Reserved, transmit as 0 (for future 10BASE-T)

EEE Auto-Neg Extension (3 of 3)

Use Unformatted next page (Page 2) to define PHY refresh cycle.

Bit	Bit definition
U10	Next page (to enable support of future technologies)
U9:U7	Reserved, transmit as 0
U6	10GBASE-KR PHY refresh cycle (0 = Reduced Energy, 1 = Lowest Energy)
U5	10GBASE-KX4 PHY refresh cycle (0 = Reduced Energy, 1 = Lowest Energy)
U4	Reserved, transmit as 0 (for future 1000BASE-KX)
U3*	10GBASE-T PHY refresh cycle (0 = Reduced Energy, 1 = Lowest Energy)
U2	1000BASE-T PHY refresh cycle (0 = Reduced Energy, 1 = Lowest Energy)
U1	100BASE-TX PHY refresh cycle (0 = Reduced Energy, 1 = Lowest Energy)
U0	Reserved, transmit as 0

EEE LLDP Frame

Bytes	Content	Value	Description
6	MAC DA	01-80-C2-00-00-0E	LLDP_Multicast address
6	MAC SA		MAC address of sending station or port
2	Ethertype	88-CC	LLDP Ethertype
9	Chassis ID TLV		Mandatory TLV (see 802.1AB)
9	Port ID TLV		Mandatory TLV (see 802.1AB)
4	Time To Live TLV		Mandatory TLV (see 802.1AB)
2	TLV type/Length	127/8	TLV Type and Length
3	OUI	00-12-0F	802.3 OUI
1	Subtype	TBD	Energy Efficient Ethernet 802.3 subtype
4	System Wake Times	See next slide	System wake time advertisement
1	EEE PHY parameters	See next slide	PHY refresh duty-cycle parameters
2	End Of LLDPDU TLV	00-00	Mandatory TLV (see 802.1AB)
14	Padding + CRC		

EEE LLDP Frame Fields

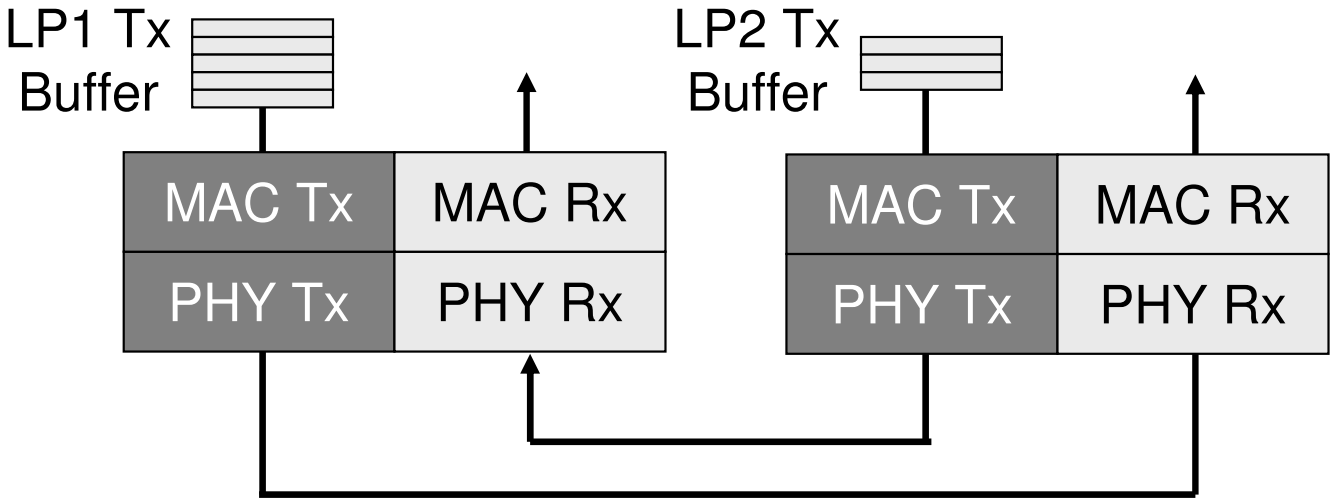
System Wake Field

Bit	Content	Value	Description
31:16	Transmit Tw_system capability	Valid Range TBD	System Wake Time, in microseconds, the system is capable of waiting before it must transmit data. For example, this could be a function of Tx system buffer depth.
15:0	Receive Tw_system request	Valid Range TBD	System Wake Time, in microseconds, the system is requesting it's link partner to hold off data transmission to enable the system to wake from the desired power-saving state. For example, this could be a function of system resume latency from the desired system power-saving states.

EEE PHY Parameters Field

Bit	Content	Value	Description
7:1	Reserved	0x0	Reserved (for potential Tw_PHY value)
0	Refresh duty-cycle		Quiet (Tq) & Refresh (Tr) time ratios for PHY power optimization 0 = "Reduced energy" duty cycle (Tq:Tr = n:1) 1 = "Lowest energy" duty cycle (Tq:Tr > n:1)

Tw_system Update Process with LLDP



1. If LP1 wants to change Tw_system, it sends EEE LLDP frame containing local Transmit Tw_system capability and Receive Tw_system request
2. LP2 responds with EEE LLDP frame containing remote Trasmit Tw_system capability and Receive Tw_system request
3. Each link partner sets its actual Transmit Tw_system value to the lower value of remote Receive Tw_system request and local Transmit Tw_system capability
4. EEE operation may resume with new Tw_system values

Opens

- Auto-neg state machine modifications
- EEE Control & status registers
  - MDIO registers for each PHY
  - MII extension registers
- 802.1AB spec changes for EEE MIB extension & TLVs
- PHY parameter values
  - Ts values
  - Tq, Tr combinations
  - Tw_PHY_{min} and Tw_PHY_{max}
- System parameter values
  - Tw_system_{min} and Tw_system_{max}

Thank You!

- Questions?