Energy Efficient Ethernet

EEE for 400G

Objective Proposal

IEEE 802.3 400G Gb/s Study Group

Michael J. Bennett
Lawrence Berkeley National Laboratory

Wael William Diab
Broadcom Corporation

Topics

- Overview of EEE recent work in recent projects
- Objective proposal

Contributors and Supporters

- Pete Anslow, Ciena
- Kousi Balasubramanian, Cisco
- Thananya Baldwin, Ixia
- Hugh Barrass, Cisco
- Yakov Belopolsky, Bel Stewart Connector
- Brad Booth, Independent
- Matt Brown, AppliedMicro
- Martin Carroll, Verizon
- Steve Carlson, High Speed Design
- Joseph Chow, Realtek
- Mandeep Chadha, Vitesse
- David Chalupsky, Intel
- Chin Chang, Huawei
- Robert Coenen, Reflex Photonics
- Ken Christensen, U. of South Florida
- Curtis Donahue, UNHIOL
- Dan Dove, AppliedMicro
- Dave Dwelley, Linear Technology
- Dave Estes, UNHIOL
- Alan Flatman, LAN Technologies
- Mike Grimwood, Broadcom
- Bob Grow, RMG Consulting
- Mark Gustlin, Xilinx
- Marek Hajduczenia, ZTE
- Adam Healy, LSI
- Jack Jewell, Independent/ Commscope
- David Law, HP
- Valerie Maguire, Siemon
- Jeff Maki, Juniper Networks

Contributors and Supporters

- Edwin Mallette, Bright House Networks
- Andy Moorwood, Infinera
- Gary Nichols, Cisco
- Ron Nordin - Panduit
- Bruce Nordman, LBNL
- Dave Ofelt, Juniper Networks
- Velu Pillai, Broadcom
- Randy Rannow, Apichip
- Pedro Riviriego, Nebrija University
- Kapil Shrikhande, Dell
- Peter Stassar, Huawei
- Nathan Tracy, TE
- Ludwig Winkel, Siemens
- George Zimmerman, CME Consulting

Energy Efficient Ethernet

- Energy Efficient Ethernet (EEE) is a set of techniques to reduce energy used by an Ethernet device during periods of low link utilization
  - Motivation is for energy proportionality

- A set of EEE interfaces were specified in IEEE Std 802.3az-2010 (now IEEE Std 802.3-2012) and new ones since in ongoing projects

Emerging Ethernet Projects with Energy Efficiency Objectives

- IEEE P802.3bj 100G Cu and Backplane Task Force
  - *EEE objective added and PAR modified to reflect EEE*
  - *Adds EEE to new P802.3bj interfaces **and** Cu interfaces from IEEE Std 802.3ba-2010 (now IEEE Std 802.3-2012)*
- IEEE P802.3bm 40G and 100G Next Generation Optics
  - *EEE objective added and PAR modified to reflect EEE*
  - *Adds EEE to new P802.3bm interfaces **and** optical interfaces from IEEE Std 802.3ba-2010 (now IEEE Std 802.3-2012)*
- IEEE P802.3bn EPoC (EPON Protocol over Coax)
  - *EEE objective included in project*
- IEEE P802.3bp: Reduced Twisted Pair Gigabit Ethernet (RTPGE)
  - *EEE objective included in project*
- IEEE P802.3bq: 40GBASE-T
  - *EEE objective included*

IEEE Std 802.3az-2010™ Overview

- Now part of IEEE Std 802.3-2012
- Low Power Idle on a variety of interfaces
  - 100BASE-TX, 1000BASE-T, 10GBASE-T, 10GBASE-KR, 10GBASE-KX4
  - EEE does not equal LPI but encompasses it
    - E.g. 10BASE-Te is EEE but doesn't use LPI
- Use LLDP for link-partner communications
  - Capability exchange for certain interfaces
    - No Auto Negotiation required with optical links
  - Enable enhanced system savings

New EEE Modes in P802.3bj and P802.3bm

- LPI modes supported for 40 Gb/s and 100 Gb/s
  - Deep sleep: the transmitter ceases transmission during the quiet period similar to IEEE Std. 802.3az-2010™ (LPI)
  - Fast wake: the transmitter continues transmitting signals between the sleep and wake states enabling the receiver to resume operation faster compared to deep sleep

- Considerations for the Optical interfaces
  - LPI exchange; physical layer stays up (Fast Wake only)
  - Capability exchange via LLDP; no need for autoneg

Motivation for EEE in 400Gb/s Ethernet

- With EEE broad market potential is reinforced
- We have retroactively added it for prior speeds
- We are adding it to the generation of Ethernet that precedes 400Gb/s Ethernet
- By the time this project is done (assuming it becomes a project) most of Ethernet will have the EEE option
  - Upper layer subsystems are being architected to take advantage of energy savings. E.g. switches and controllers

Motivation for EEE in 400Gb/s Ethernet

- Data center operators are very much interested in using power efficiently as energy-use impacts operational expense
  - E.g. Google spent ~$200M on energy in 2010
    - Note that Google's data centers are roughly 50% more efficient than others
- Data Center Operators want energy-proportional equipment
- Larger data centers use optical links
  - "Likely a lot of value in figuring out EEE for optical links"

Source: http://www.ethernetalliance.org/wp-content/uploads/2012/02/EATEF_Panel-3_Power_12_0216.pdf slides 51,52,56

Motivation for EEE in 400Gb/s Ethernet

- Energy Efficiency is a priority for regulators
  - EU CoC on Energy Consumption of Data Centers
  - EU CoC on Broadband equipment *requires* EEE (effective 2014)
  - Energy Star specs for Small Network Equipment
  - Energy Star Large Network Equipment has started
  - Policy will encourage technologies like EEE
  - Can support that by including EEE in the specification
    - Avoid other non-standardized techniques to reduce energy that may get pushed by various specifications absent EEE
- EEE is effectively a "must" for a new specification

Source: http://www.itu.int/dms_pub/itu-t/oth/09/05/T09050000010004PDFE.pdf
Source: http://www.energystar.gov/index.cfm?c=new_specs.small_network_equip

Why EEE *at the start of projects?*

- Very difficult and time consuming task to retrofit EEE into completed specifications
- Run the risk of breaking things
- Much more efficient to consider EEE in the initial specification

How could this apply to 400Gb/s Ethernet?

- These are just examples…
- Lowest hanging fruit
  - Leverage work from P802.3bj and P802.3bm
  - Use Fast Wake LPI codewords for signaling – no PMD power-down
- Any other ways to achieve energy efficiency with or in conjunction or instead of LPI
  - That's up to the SG/TF to determine, once objective is adopted
- This is a contribution-driven organization ...
  - Let's include EEE in our scope of work

Objective Proposal

Proposed text for objective motion:

Move that the 400Gb/s Ethernet Study Group adopt the following objective:

Specify optional Energy Efficient Ethernet (EEE) capability for 400Gb/s PHYs

(>= 75% required) Y: N: A:

Questions?

Thank You!

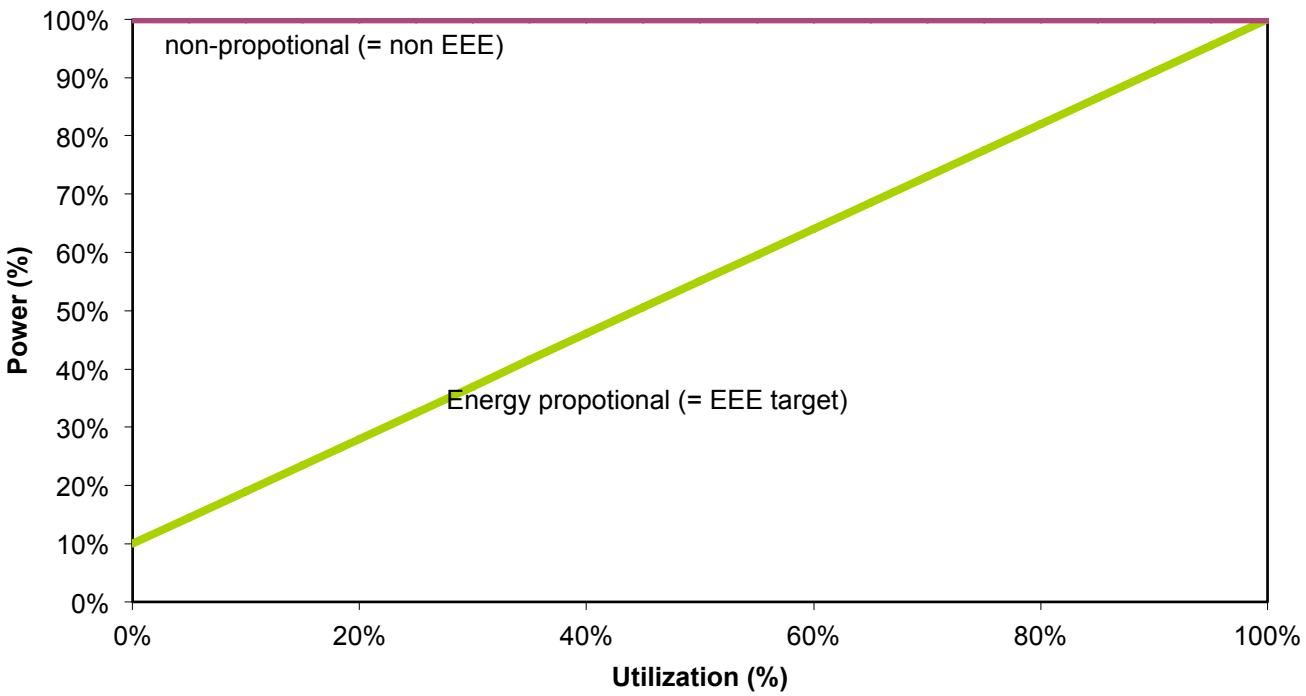
Backup

Motivation for EEE in 400Gb/s Ethernet

- EEE could help make the datacenter more energy proportional to load[1].
- End users are asking developers to "make better energy proportionality a primary design objective" for future systems[1].
- Savings for the IEEE 802.3az PHY alone should be around 90% and energy reduced by up to 70% for the NIC when in LPI mode[2].
  - much greater savings possible in systems using LLDP
    - See dove_02_05_08.pdf (slide 5)

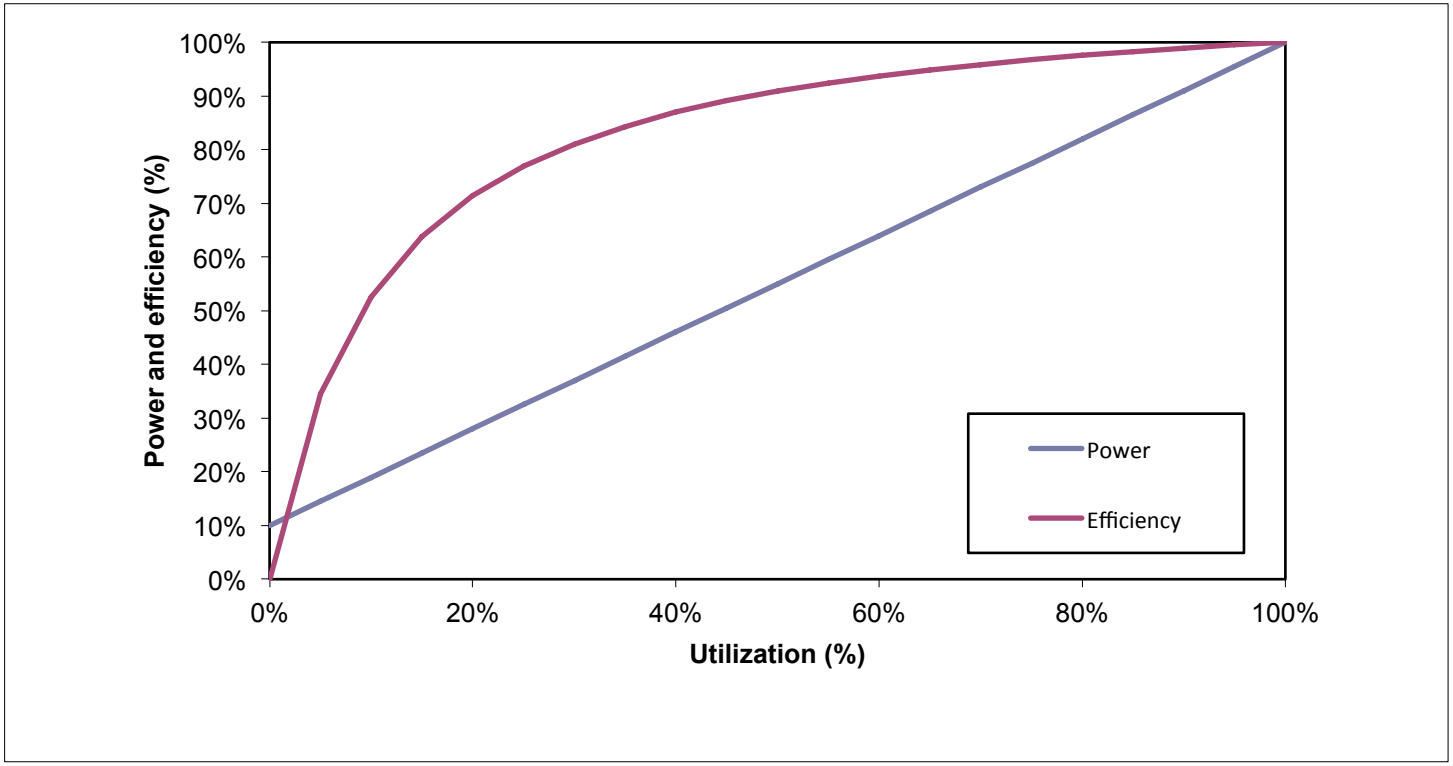
1. L. Barroso and U. Hölzle, The Case for Energy-Proportional Computing. Computer, 40(12):33-37, December 2007
2. P. Reviriego, K. Christensen, J. Rabanillo, and J. A. Maestro, 'An Initial Evaluation of Energy Efficient Ethernet' in IEEE communications letters, VOL. 15, NO. 5, May 2011

Energy Proportionality Illustration



Thanks to Ken Christensen, University of South Florida for this contribution

Energy Proportionality Illustration



- Note that if the power curve is non-linear (e.g., convex) it is possible for efficiency to be greater than 100%. This does suggest that the term "efficiency" is incorrect.

Thanks to Ken Christensen, University of South Florida for this contribution