An Overview of Energy-Efficient Ethernet

Reduced Twisted-Pair Gigabit Ethernet Study Group

Michael J. Bennett
Lawrence Berkeley National Laboratory

Wael William Diab
Broadcom Corporation

Contributors and Supporters (partial list)

Topics

- Overview of Energy-Efficient Ethernet
  - Low Power Idle
  - Link Partner Communications

- Observations and Next Steps

What is Energy-Efficient Ethernet?

- Energy Efficient Ethernet (EEE) is a method to reduce energy used by an Ethernet device during periods of low link utilization

- Specified in IEEE 802.3az-2010™

- The premise for EEE is that Ethernet links have idle time and thus opportunity to save energy

- Specified for copper interfaces

    - "BASE-T's'

    - Backplane (except 40G)

- The method is called Low Power Idle (LPI)

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
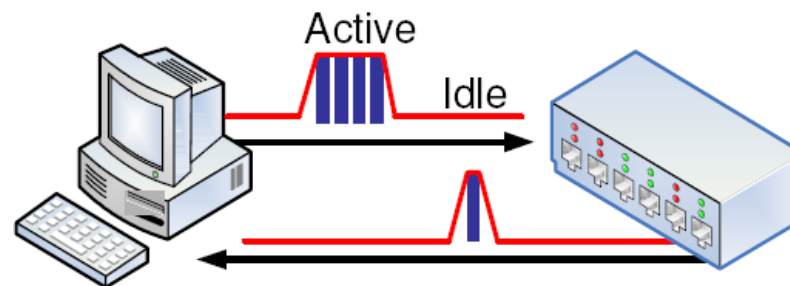
Example: link utilization

- Snapshot of a File Server with 1 Gb Ethernet link
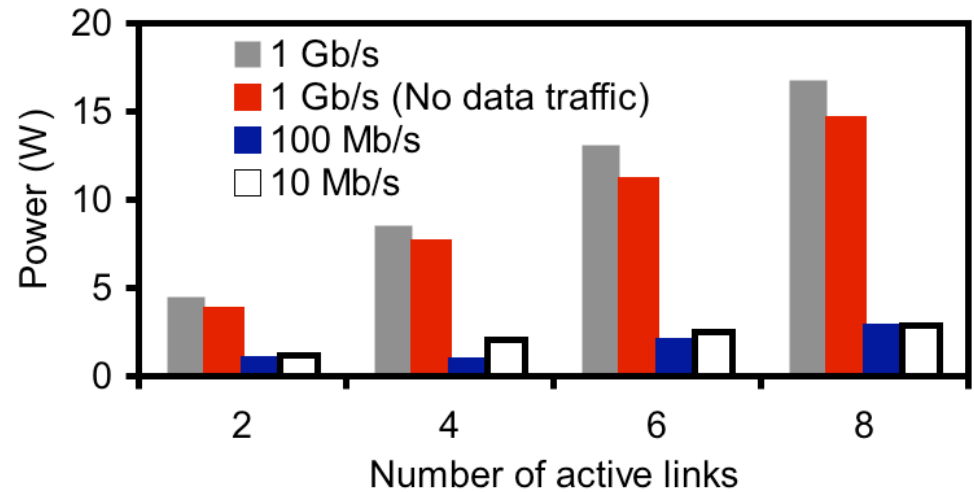 - Shows time versus utilization (trace from LBNL)

File Server Bandwidth Utilization Profile

Start time 12:33 PM 2/8/2007 (30 min)

utilization <=1.0 %
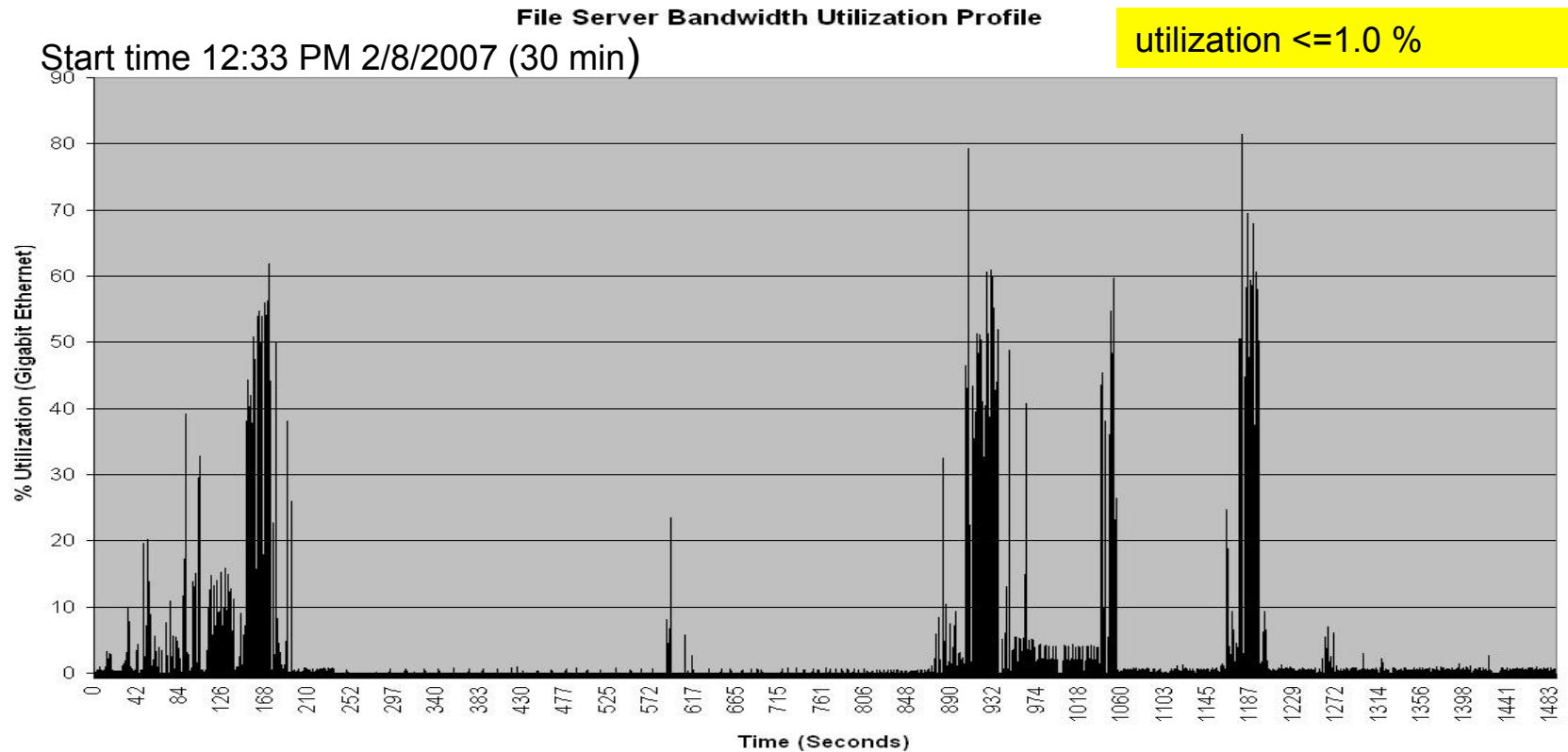
Link power

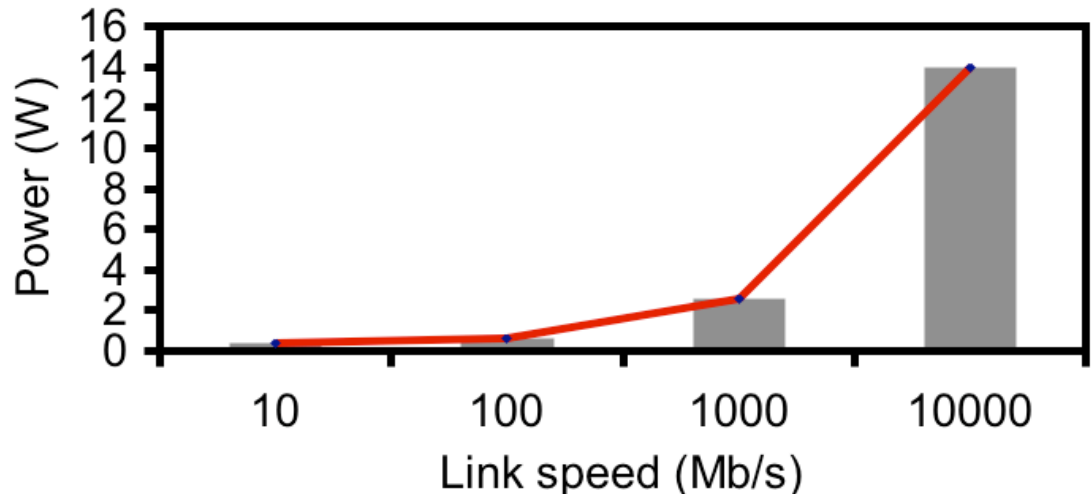## Results from (rough) measurements

— **all incremental AC power**

— **measuring 1st order**
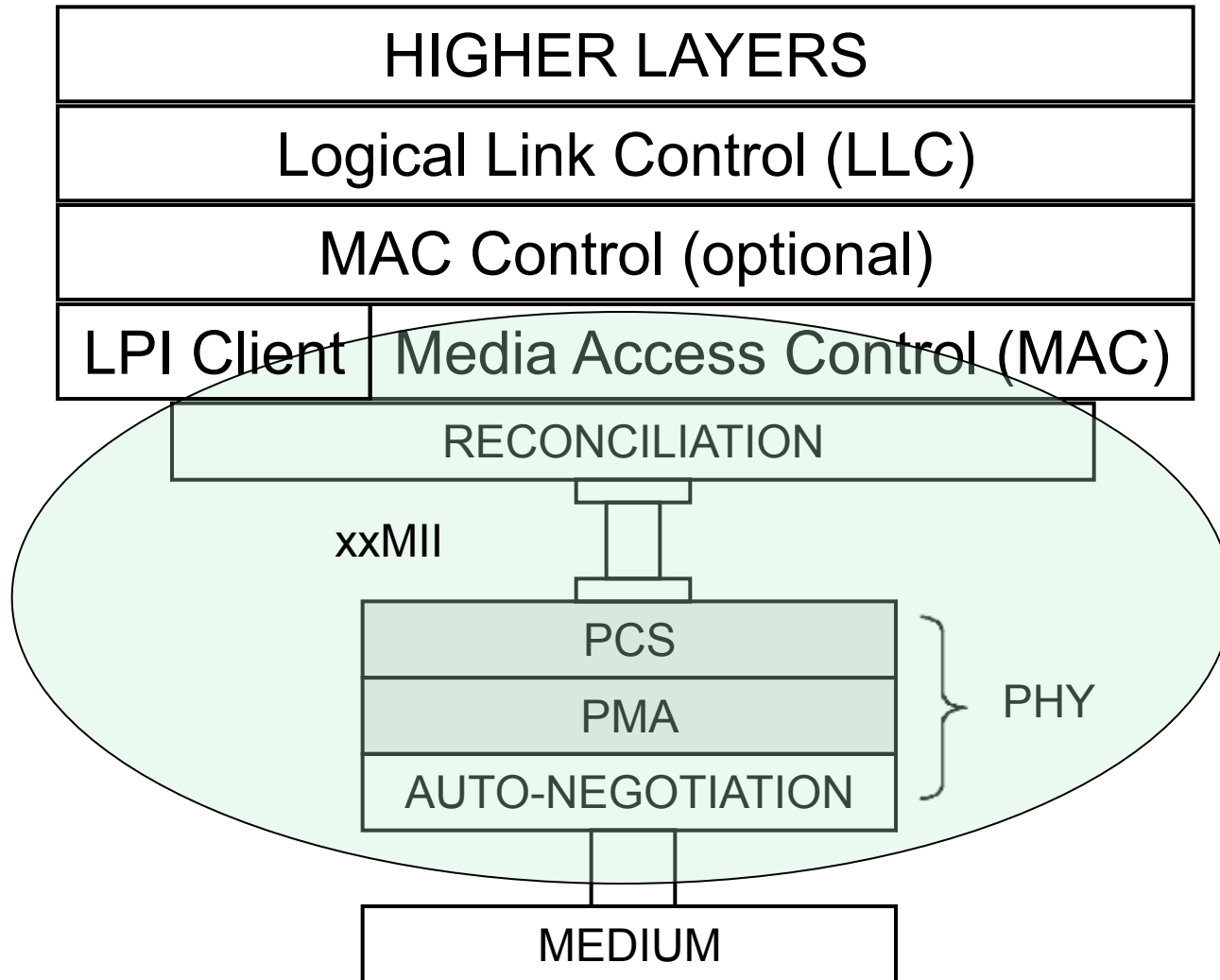
- Typical switch with 24 ports 10/100/1000 Mb/s

- Various computer NICs averaged

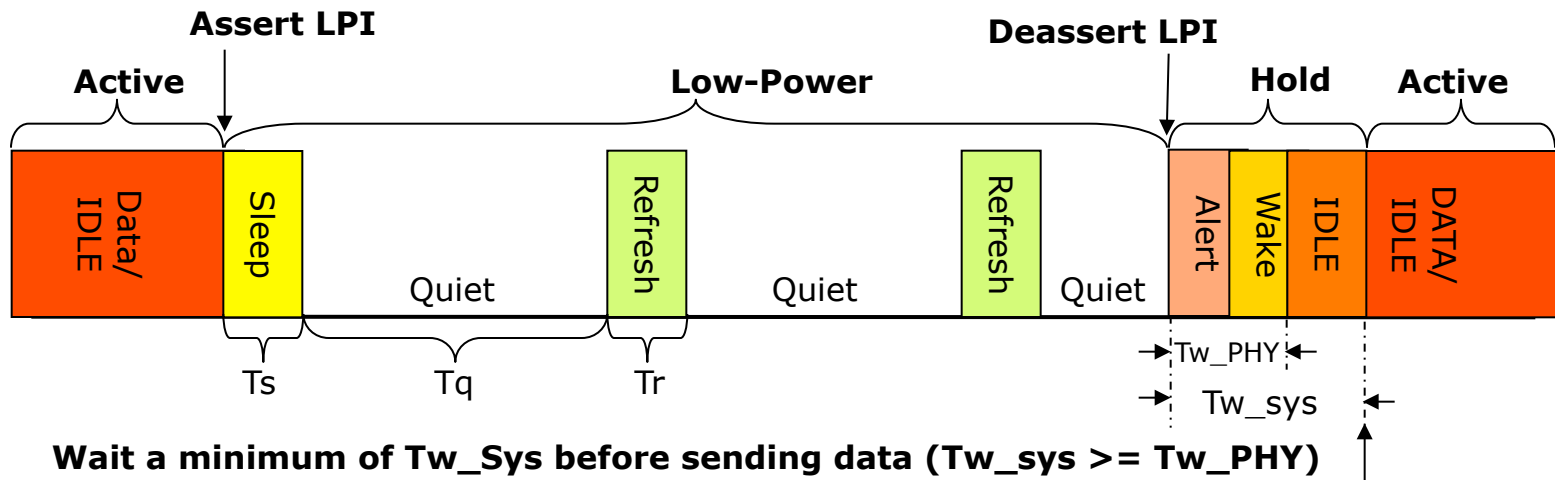Note: Measurement data from 2006

Where EEE Fits

HIGHER LAYERS
Logical Link Control (LLC)
MAC Control (optional)

LPI Client	Media Access Control (MAC)

RECONCILIATION

xxMII

PCS

PMA

AUTO-NEGOTIATION

} PHY

MEDIUM

LPI Overview

Assert LPI

Deassert LPI

Active ⟵ Data/IDLE · Sleep ⟶ **Low-Power** ⟵ Quiet · Refresh · Quiet · Refresh · Quiet ⟶ **Hold** ⟵ Alert · Wake · IDLE ⟶ **Active** ⟵ DATA/IDLE

Ts   Tq   Tr

Tw_PHY

Tw_sys

Wait a minimum of Tw_Sys before sending data (Tw_sys >= Tw_PHY)

- LPI – PHY non-essential circuits shut down during idle periods
- During power-down, maintain coefficients and sync to allow rapid return to Active state
- Wake times (Tw_PHY) for Twisted-Pair PHYs:
  - 100BASE-TX: <= 30 usec
  - 1000BASE-T: <= 16.5 usec
  - 10GBASE-T: <= ~8 usec (2 modes)

Link partner communication

- Uses auto-negotiation to notify link partner of EEE capabilities

- Uses Link Layer Discovery Protocol (LLDP) to notify link partner of parameter changes

  - E.g. control policy

    - User can choose energy savings preferred over performance or vice versa
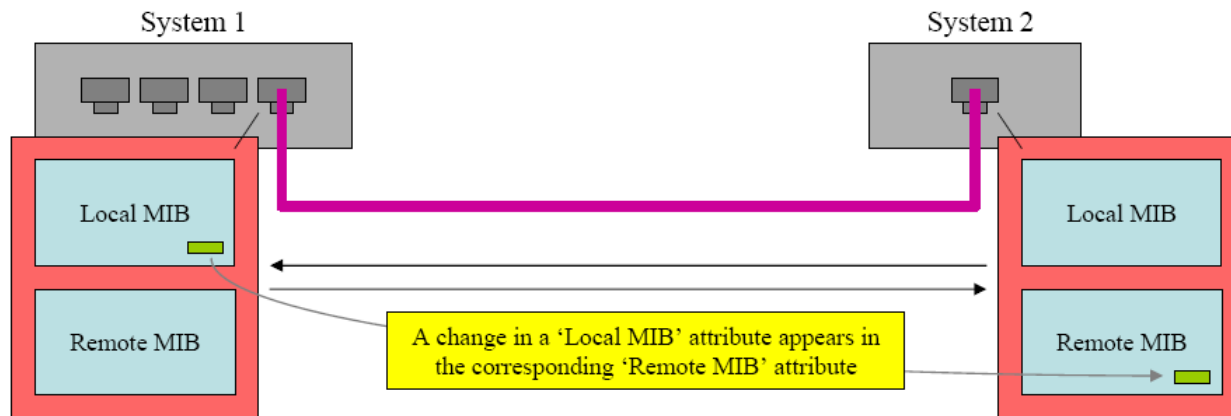
Observations and Next Steps

- Low standby power mentioned in consensus building CFI preso
    - http://www.ieee802.org/3/RTPGE/public/mar12/CFI_01_0312.pdf#Page=22
- Is there interest in pursuing a EEE solution as part of the RTPGE work to reduce power consumption during IDLE?
    - If so, would interest be to
        - Carry protocol to allow higher layers to take advantage of IDLE
        - Suspend operation at the physical layer
        - Both of the above
        - Utilize a different scheme
- If a suspension at the PHY layer is needed, what acquisition times (wake-up times are satisfactory)
    - E.g. similar to 1000BASE-T, longer, shorter etc.
- Leverage of EEE if similar LPI scheme used *and PHY is suspended*
    - Some portion of EEE can be leveraged (e.g. protocol and layer 2)
    - Some will be specific to PHY defined and will be different from 1000BT

Thank You!