IEEE 802.3az: The Road to Energy Efficient Ethernet

K. Christensen, P. Reviriego, B Nordman, M. Bennett, M. Mostowfi, J.A. Maestro

Presented by: Jordi Cucurull

Department of Computer and Information Science (IDA)
Linköpings universitet
Sweden

LiU expanding reality

Outline

- Why Energy Efficient Ethernet?
- The 802.3az Standard
- Packet coalescing
- Evaluation

Outline

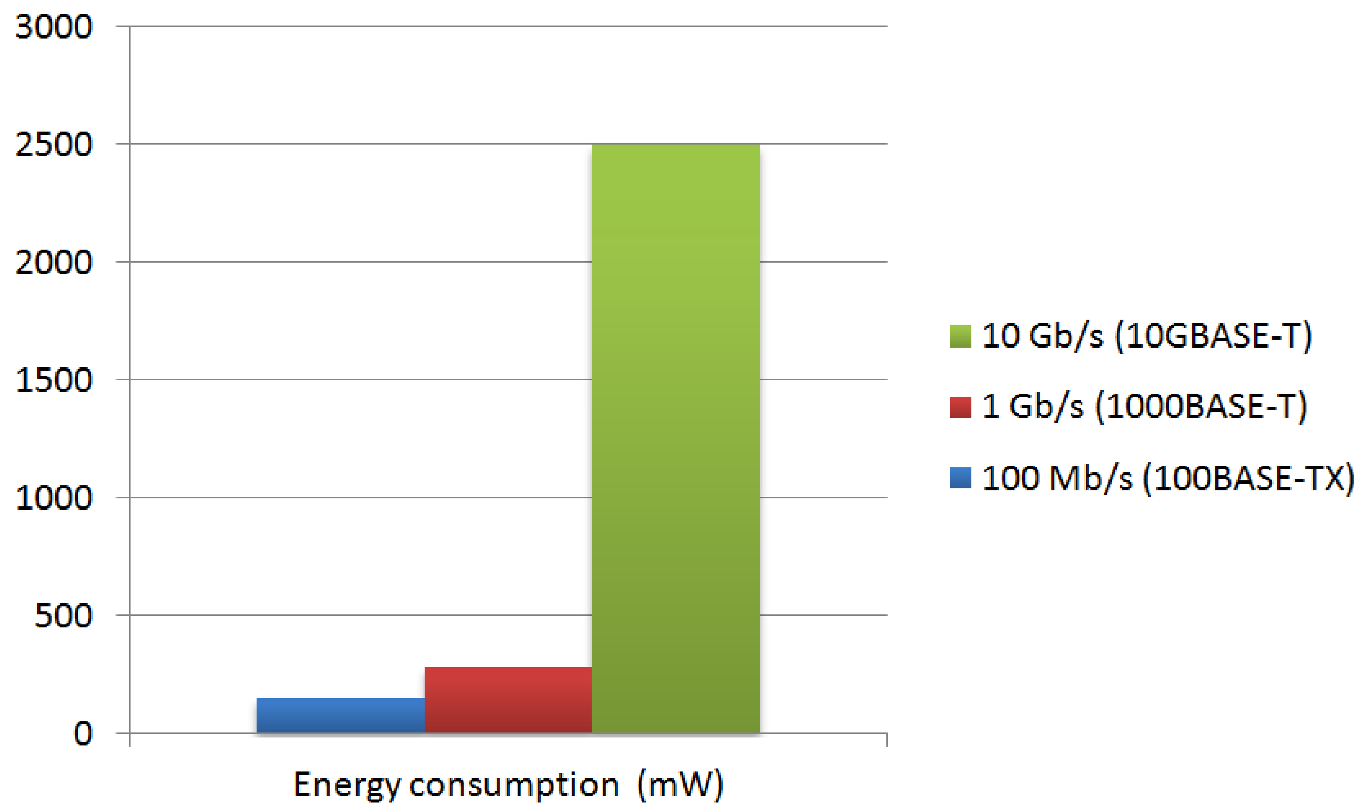
- **Why Energy Efficient Ethernet?**
- The 802.3az Standard
- Packet coalescing
- Evaluation

Energy consumption of Ethernet

- Ethernet interfaces are highly spread
  - Present on desktops, notebooks, servers, TVs…
  - 1 billion in US and 3 billion worldwide (2010)

- Four different data rates on UTP cable
  - 10 Mb/s (10BASE-T)
  - 100 Mb/s (100BASE-TX)
  - 1 Gb/s (1000BASE-T)
  - 10 Gb/s (10GBASE-T)

Energy consumption of Ethernet



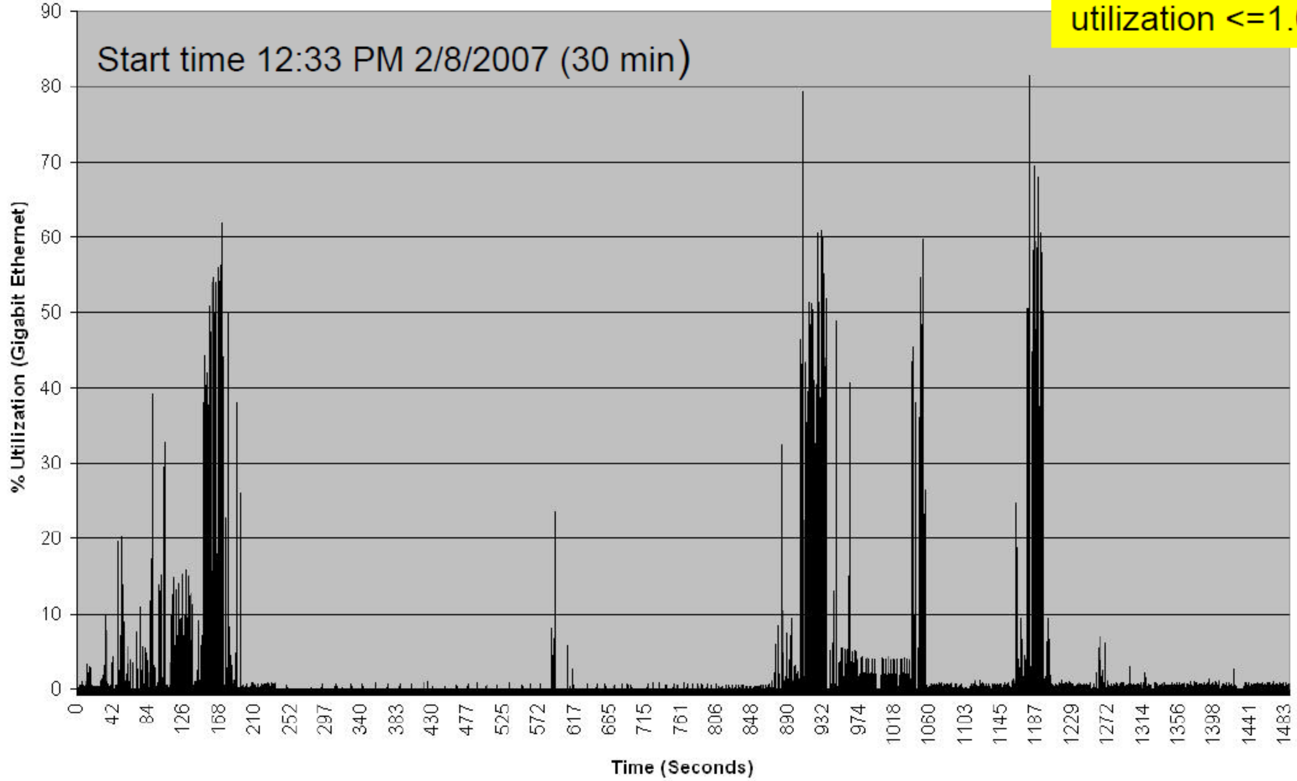
Source: Marvell Semiconductor Inc.

Energy consumption of Ethernet

- Energy consumption dependent on link data rate
  - Higher data rate more complex physical interfaces
    - To increase speed
    - To keep low bit error rate at 100m distance
- Energy consumption independent on link usage
  - When no data to send an IDLE signal is transmitted

Low utilisation of server links

utilization <=1.0 %



File server with 1Gb/s Ethernet

H. Barrass, M.Bennett, W.W.Diab, D.Law, B.Nordman, G.Zimmerman. Tutorial Energy Efficient Ethernet. 2007

Outline

- Why Energy Efficient Ethernet?
- **The 802.3az Standard**
- Packet coalescing
- Evaluation

History of the standard

- July 2005 – Idea of creating Energy Efficient Ethernet (EEE)
- November 2006 – Panel presentation in 802.3 Working Group to describe rationale for EEE
- September 2007 – P802.3az Task Force is formed
- October 2008 – First draft of the standard produced
- September 2010 – IEEE Std 802.3az approved

Two approaches to save up energy

- Low Power Idle
  - Switch ports to low power mode during idle periods
- Adaptive Link Rate
  - Modify the link rate during periods of low traffic

Both search for energy proportionality!

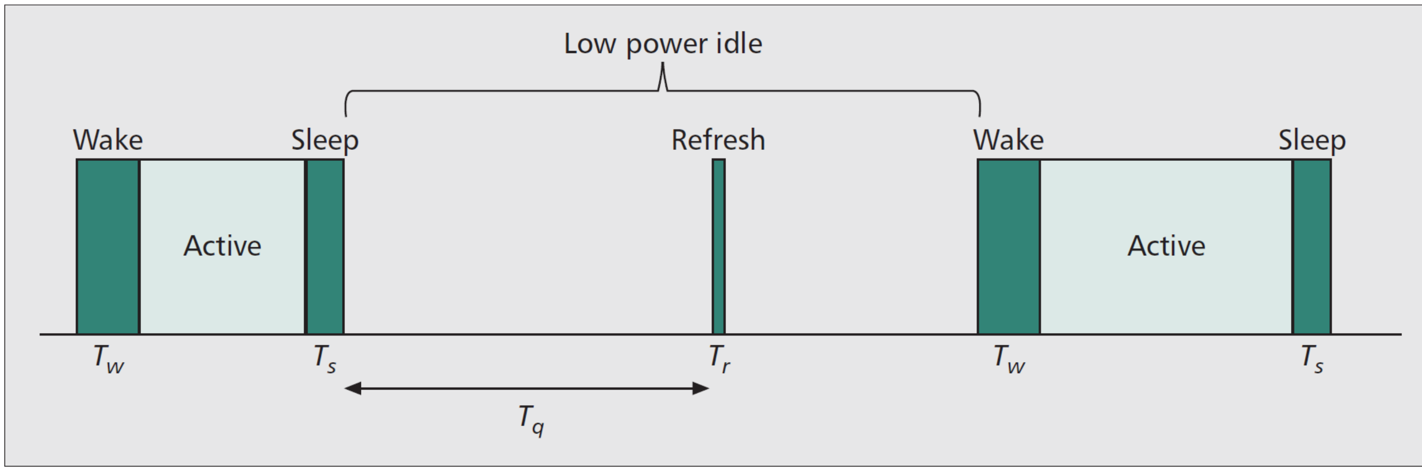
Low Power Idle (LPI)

- Smart switching of interface to low power mode
  - LPI replaces the continuous IDLE signal
  - Energy consumption around 10% of active mode

- Synchronisation is kept with signalling during short periodic refresh intervals (T_r)

- Trade-off energy for latency
  - Transitions between active/sleep modes take time
    - T_s – Time to go to low power (sleep) mode
    - T_w – Time to wake up link

Low Power Idle (LPI)



T_s: Time to low power idle (sleep)

T_w: Time to go to active (wake)

T_q: Interval without signalling

T_r: Refresh signalling interval

Overhead of Low Power Idle

- Overhead is characterised by T_s and T_w
- Example of overhead for 10 Gb/s Ethernet:

	Value
Time to sleep (T_s)	4.48 µs
Time to wake up (T_w)	2.88 µs
Time to send packet of 1500 bytes (TCP DATA)	1.2 µs
Time to send packet of 64 bytes (TCP ACK)	0.0512 µs

Outline

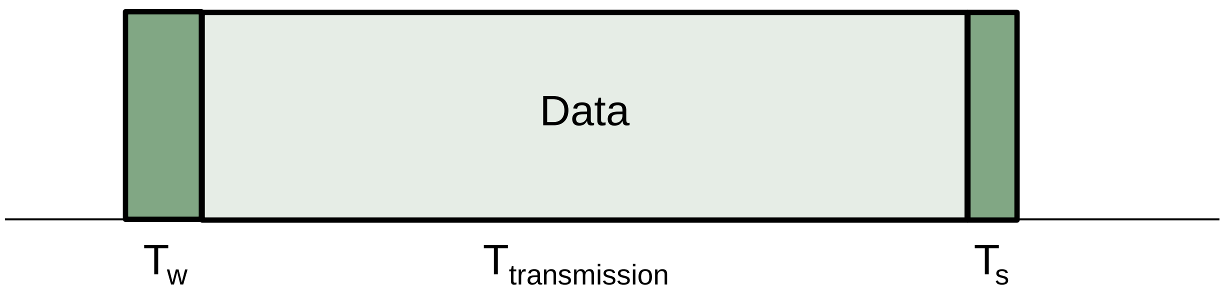
- Why Energy Efficient Ethernet?
- The 802.3az Standard
- **Packet coalescing**
- Evaluation

Performance trade-offs

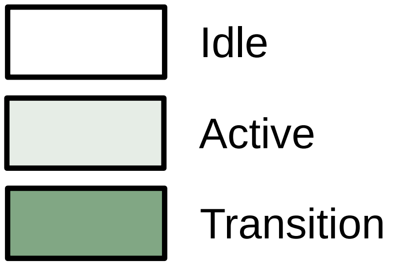
- EEE Energy efficiency as a function of
  - Link utilisation
  - Packet transmission time
  - Distribution of packet interarrival times

A two-case TCP download example

- Transmission of large block of packets

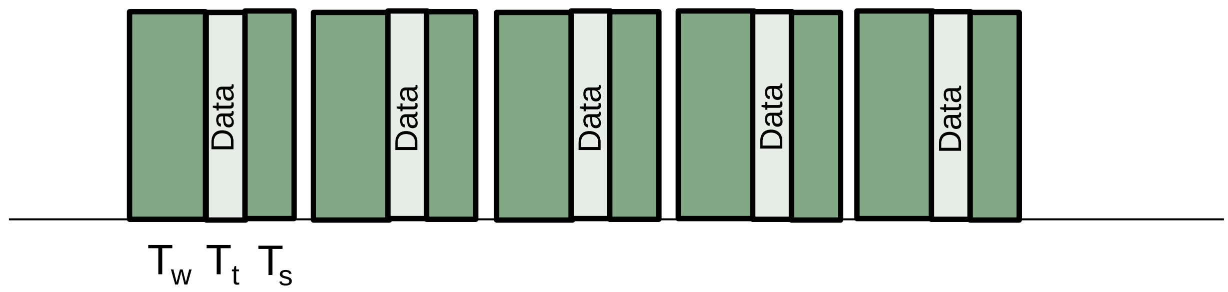


Very low overhead: $T_w + T_s << T_{transmission}$



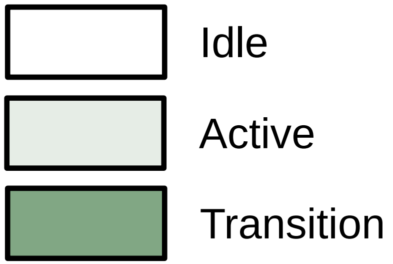
A two-case TCP download example

- Transmission of small packets with gaps between them



Very high overhead: $T_w + T_s > T_{transmission}$

- More latency
- Less energy reduction



How to reduce overhead?

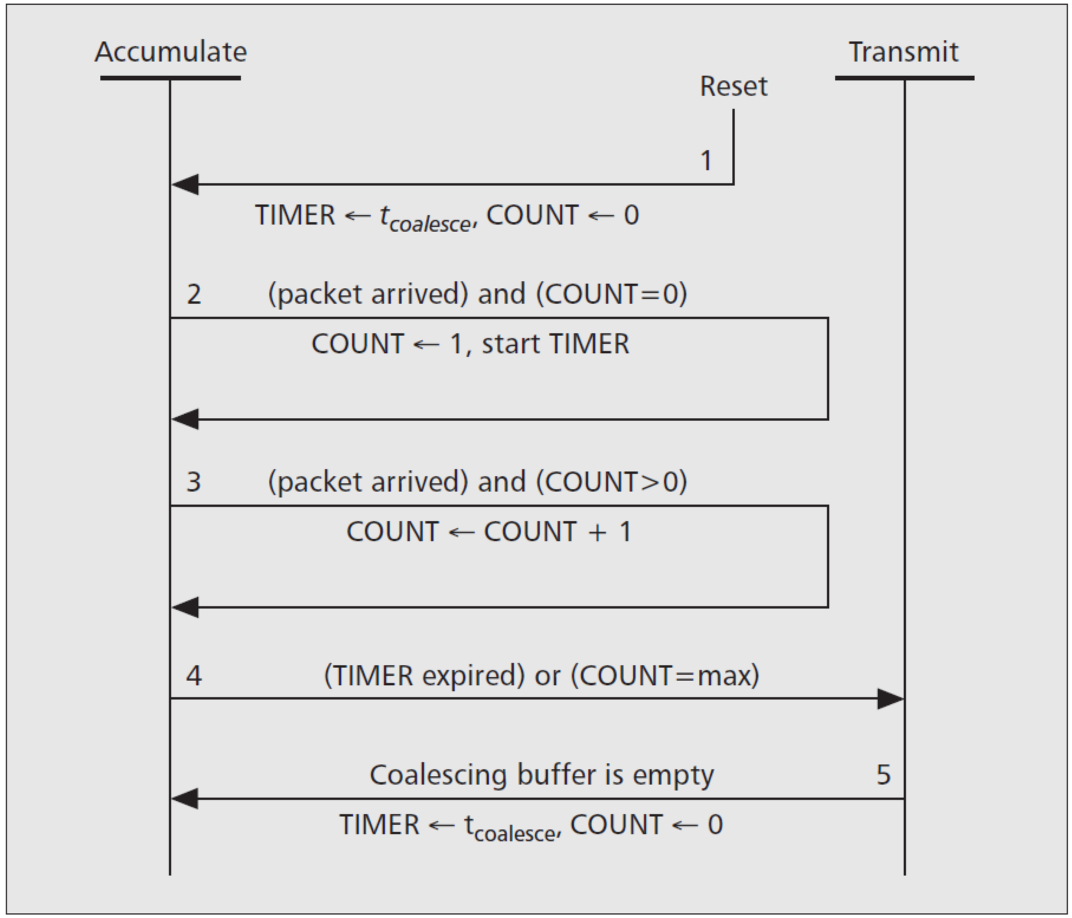
- Buffering packets before sending them
- Packet coalescing
  - Proposed technique to reduce overhead
  - Not part of the standard, but compatible with it

Packet coalescing

- FIFO queue in the Ethernet interface
  - Collects (or coalesces) multiple packets to be sent
  - Packets are sent as a burst over the link

- Coalescing approaches can be based on
  - Packet count
  - Time from first packet arrival
  - Combination of both

Packet coalescing



Finite state machine for packet coalescing

Outline

- Why Energy Efficient Ethernet?
- The 802.3az Standard
- Packet coalescing
- **Evaluation**

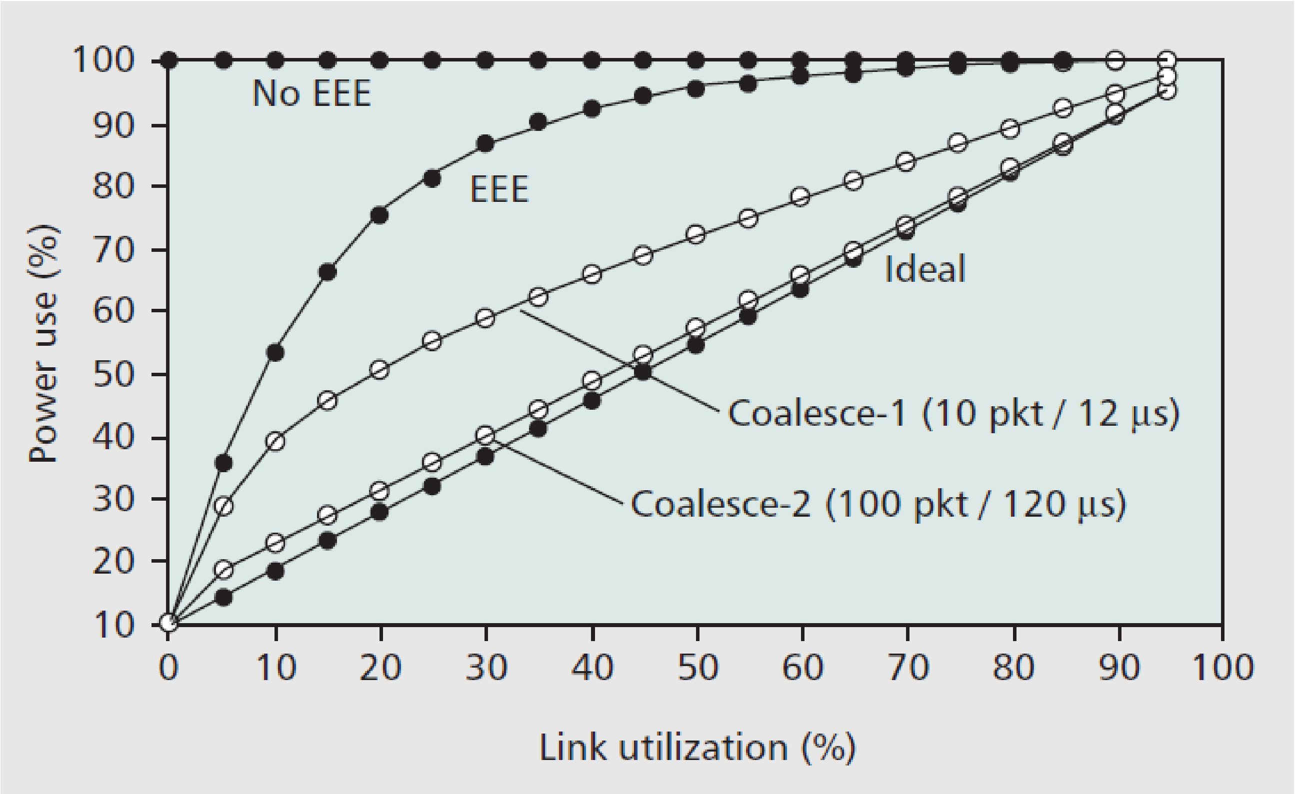
Link simulation

- Packet coalescing analysed on 10 Gb/s simulated link
  - Packets arriving as a Poisson process
  - Packet length of 1500 bytes (T_{pkt} = 1.2 µs)
  - Low power mode consumed 10% of active mode
- Effects measured with different link utilisation
  - Power usage
  - Mean packet delay

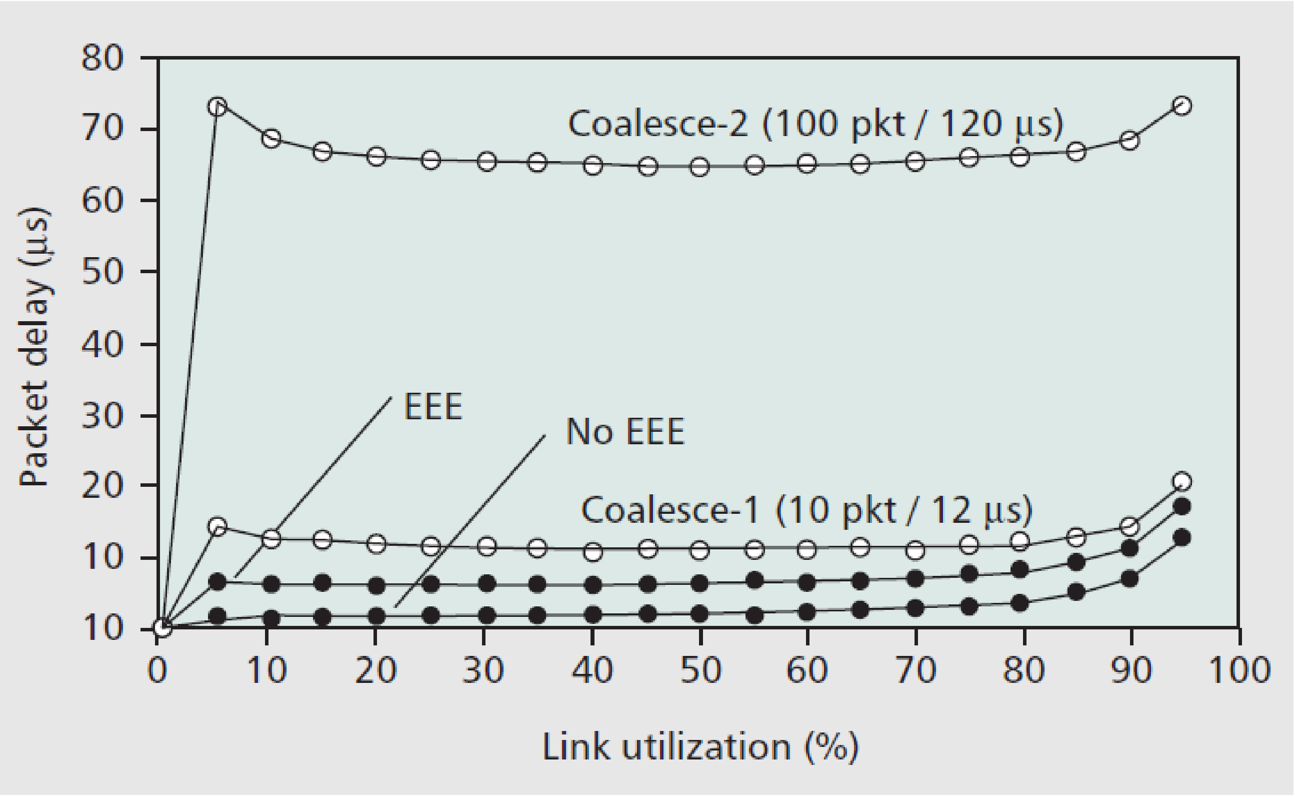
Link simulation

- Different parameters selected
  - Energy Efficient Ethernet link
    - No coalescing
    - Coalescing with two different configurations (timer/counter)
      - Coalesce-1: $t_{coalesce}$ = 12 μs and *max_packets* = 10
      - Coalesce-2: $t_{coalesce}$ = 120 μs and *max_packets* = 100
  - Regular Ethernet link

Link simulation results



Link simulation results

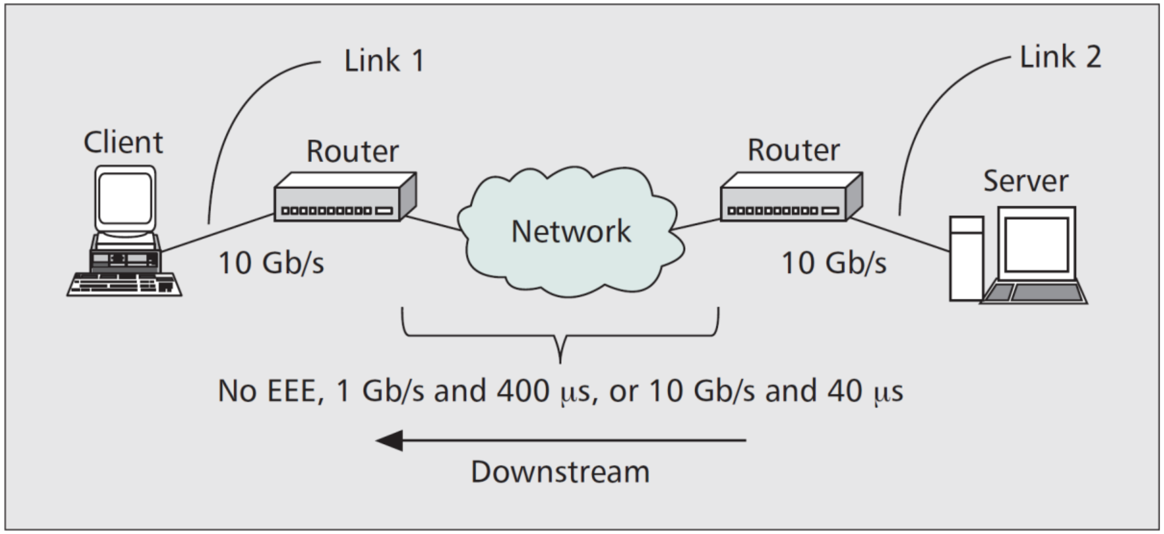


Department of Computer and Information Science (IDA)
Linköpings universitet, Sweden

What is the significance of these delays?

- Internet round trip time is around 10-100 ms
  - An increase of a few tens of microseconds not significant
  - But some time critical applications could be affected
- And if this technology is deployed in the whole network?
  - Then it must be studied.

Network simulation



- Network simulated with NS-2
  - Two core networks (different speed/delay)
  - Buffer of 100 packets at each router
- 1 GB file downloaded with TCP connection
  - 400 packets window size

Energy calculation

- Formula used to compare energy in simulations

$$E = T_w + T_s + T_{transmission} + 0.1 \times T_{idle}$$

Results obtained

10 Gb/s core network with 40 μs of delay

	Download time	Usage Link 1 Up	Energy Link 1 Up	Usage Link 1 Down	Energy Link 1 Down
No EEE	0.843 s	4.0 %	100.0 %	94.9 %	100.0 %
EEE	0.843 s	4.0 %	99.9 %	94.9 %	99.9 %
EEE coalesce-1	0.843 s	4.0 %	50.6 %	94.9 %	99.9 %
EEE coalesce-2	0.847 s	4.0 %	21.3 %	94.5 %	99.5 %

- Download time is almost not affected
- Significant energy savings in up link (ACK channel)

Results obtained

1 Gb/s core network with 400 μs of delay

	Download time	Usage Link 1 Up	Energy Link 1 Up	Usage Link 1 Down	Energy Link 1 Down
No EEE	8.28 s	0.4 %	100.0 %	9.7 %	100.0 %
EEE	8.28 s	0.4 %	65.6 %	9.7 %	74.4 %
EEE coalesce-1	8.28 s	0.4 %	38.0 %	9.7 %	46.7 %
EEE coalesce-2	8.34 s	0.4 %	17.8 %	9.7 %	25.8 %

- Download time is almost not affected
- Significant energy savings in up/down links

Network performance

- Network performance almost not affected
  - But different parameters can lead to worse performance
- Two conditions required for good performance
  - Burst size much smaller than
    - Intermediate buffers in routers and NICs
    - TCP window
  - Coalescing timer much smaller than round trip time

Conclusions

- 802.3az is a fully approved and functional standard
  - Many devices implement the standard
- Energy efficiency of 802.3az devices can be easily improved with packet coalescing
- Millions of dollars can be saved up
  - $410 million savings per year in US
    ($80 additional millions with coalescing)
  - $1 billion savings per year globally