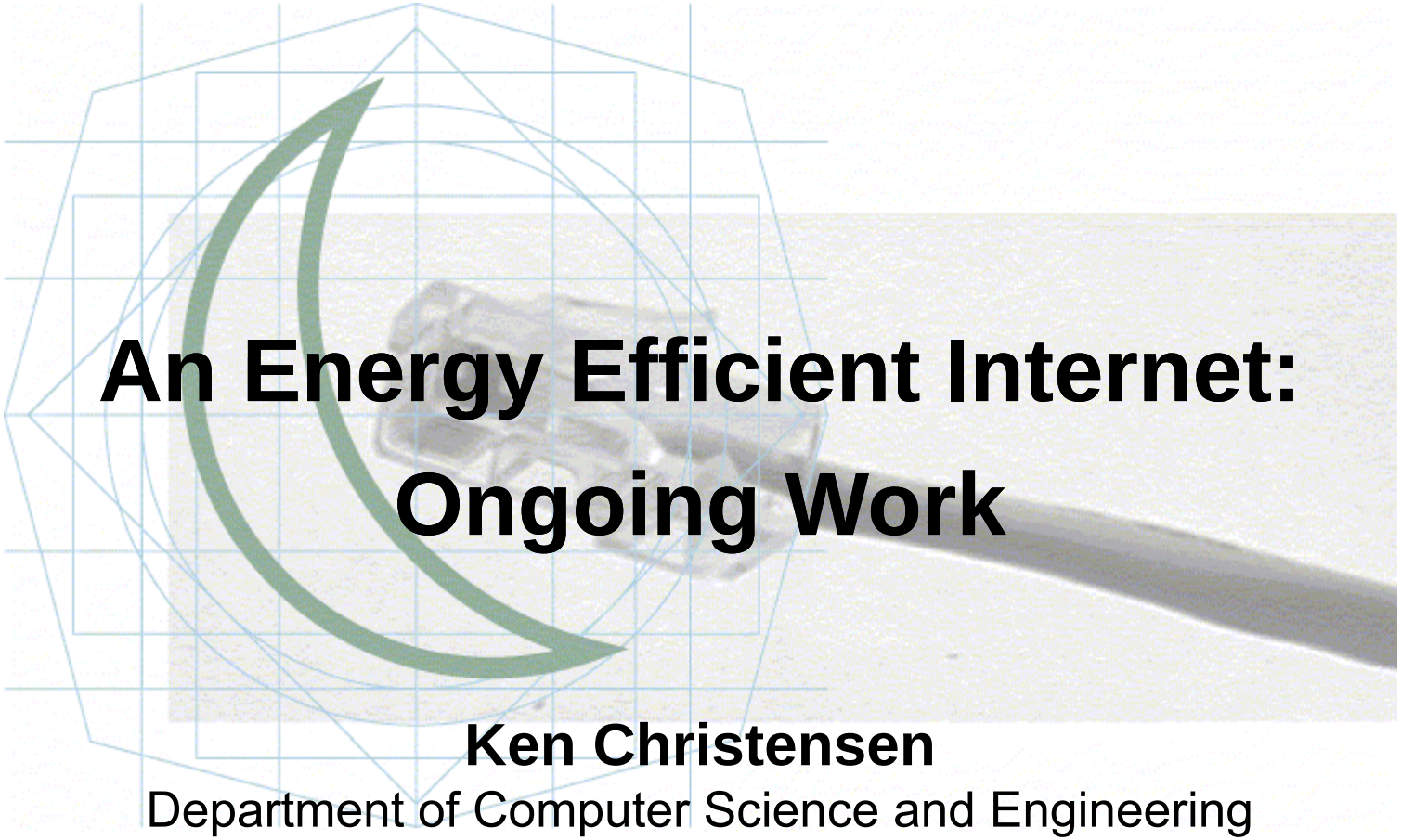


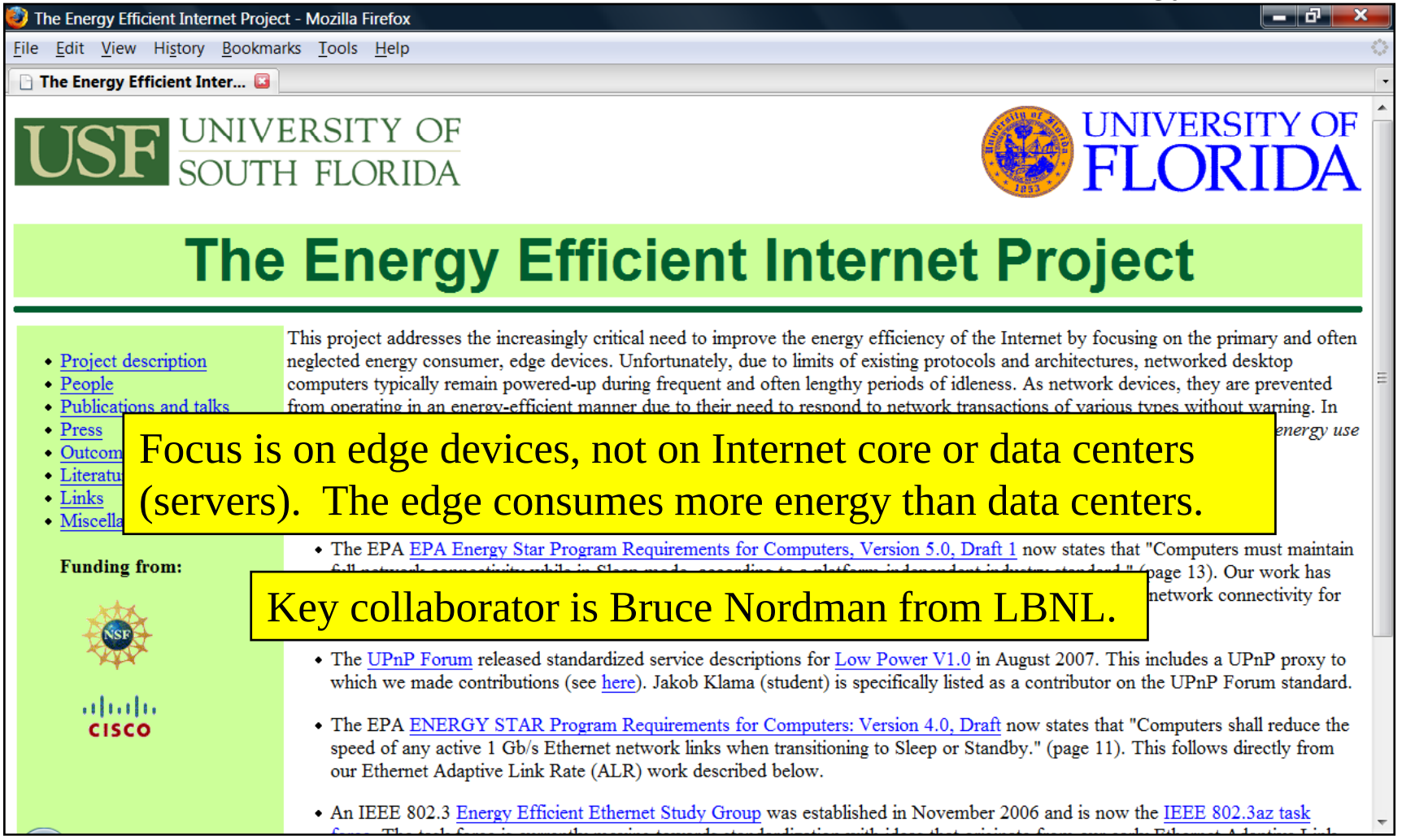
An Energy Efficient Internet: Ongoing Work

Ken Christensen
Department of Computer Science and Engineering
University of South Florida
Tampa, FL 33620
christen@cse.usf.edu

This material is based upon work funded by the National Science Foundation under grant CNS-0520081.

The project

http://www.csee.usf.edu/~christen/energy/main.html



Where this talk is going

- **Energy usage by IT equipment**
- **Reducing *direct* energy use**
- **Reducing *induced* energy use**
- **Some additional efforts**

Direct energy use = energy used by network links and equipment (routers, switches), but not end devices.

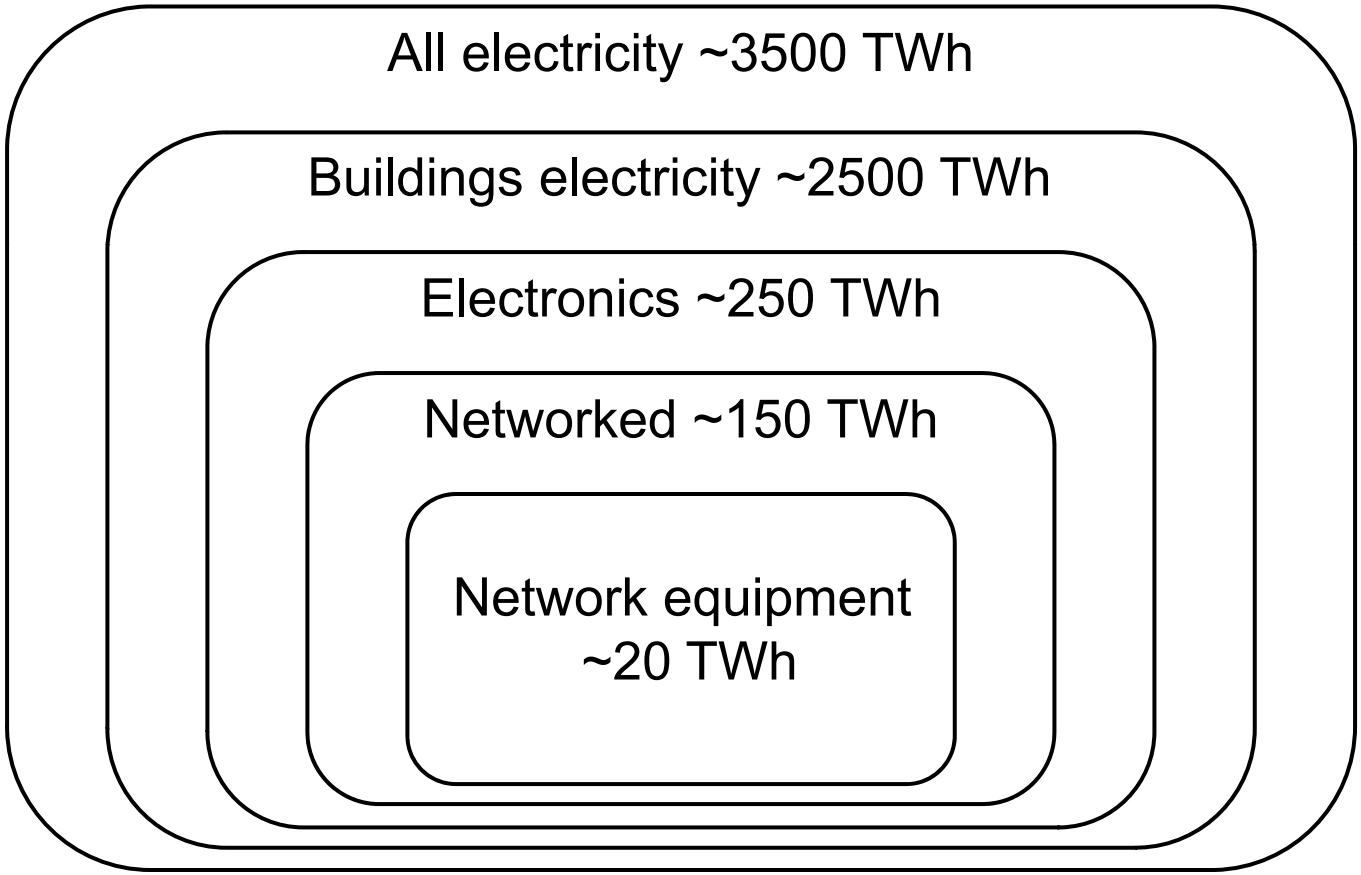
Induced energy use = Increment for higher power state of devices needed to maintain network connectivity.

Electricity production and costs

- **Let's set the stage…**
  - 1 kWh = $0.10 (average US residential cost for electricity)
  - 1 Wyr = $0.88
  - 1 TWh = $100 million
  - 1 TWh = 0.75 million tons of CO_2

Electricity use in the USA

- **2006 US electricity usage*** (not to scale)



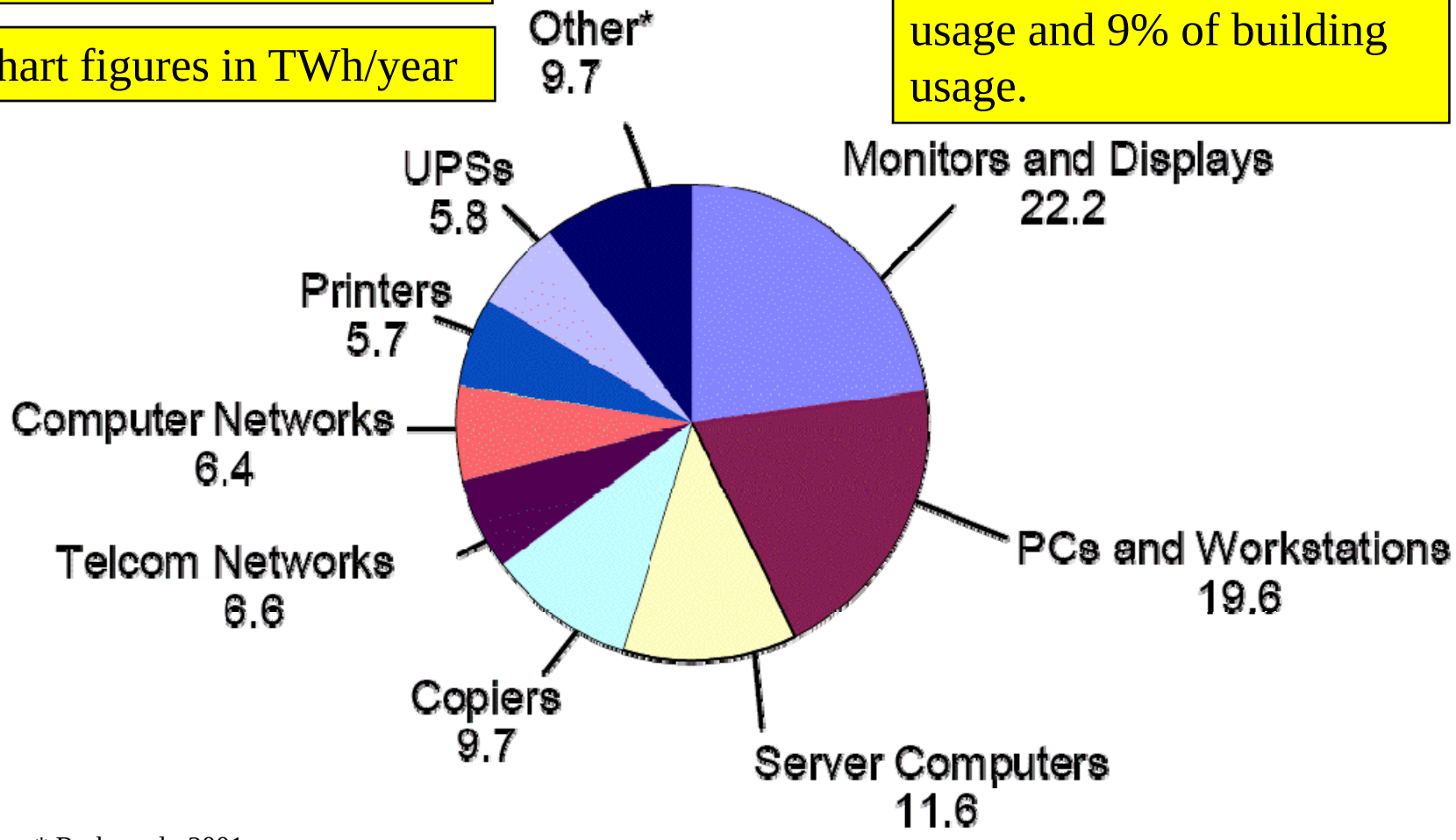
* B. Nordman, "Networks, Energy, and Energy Efficiency," presentation at *Cisco Green Research Symposium*, March 2008.

IT energy use in 2000: All IT equipment*

Commercial buildings only

Chart figures in TWh/year

97 TWh/yr in 2000. This is 3% of national electricity usage and 9% of building usage.



* Roth er al., 2001.

Electricity use by IT equipment today

- **How much electricity do PCs consume?**
  - EPA estimates about 2% of all electricity consumed*
- **How much electricity do data centers consume?**
  - About 1.2% of all electricity consumed**
- **How much electricity does the Internet consume?**
  - "The Internet accounts for 5% of all the power we consume – in a couple of years, that figure will be 10%." ***

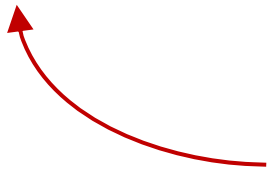
* "EPA Announces New Computer Efficiency Requirements," Release date: 10/23/2006, Contact: Enesta Jones.
** Jon Koomey quoted in InformationWeek, February 15, 2007
** Institute for Energy Efficiency, UC Santa Barbara, 2008.

How much in greenhouse gas?

- **Figure it out at the EPA Greenhouse Gas Calculator**

  http://www.epa.gov/cleanenergy/energy-resources/calculator.html

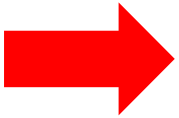
- **One 100W PC on 24/7 for one year is…**
  - 0.88 metric tons of CO2
  - 0.12 passenger cars for one year
  - 77.3 gallons of gasoline consumed
  - 0.09 homes for one year



One PC = about 10% of a home!

More on PC energy impact

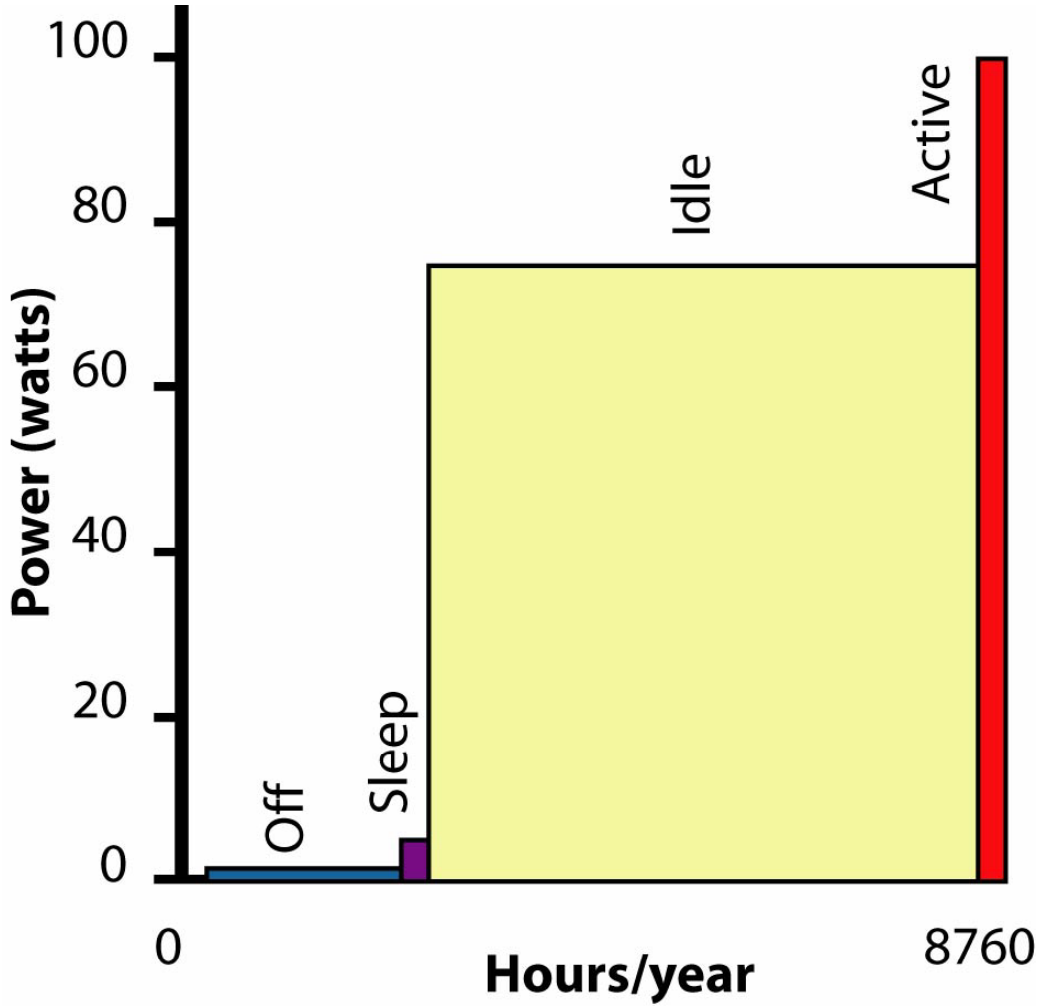
- **The typical US home consumes 10,654 kWh/yr***
- **One 100W PC on 24/7 for one year is…**
  - 876 kWh/yr
  - This is 8.2% of entire home consumption



Each PC adds about 8% to your power bill!

* Energy Information Administration, "U.S Household Electricity Report," July 2005.

Typical commercial PC energy use



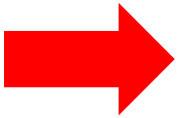
$$P_{on} >> P_{sleep}$$

$$P_{sleep} \cong P_{off}$$

Consumption is driven by on time, not by usage

PE energy use is “induced”

- **Why is the typical PC fully powered on when idle?**
  - For usability reasons?
  - For network connectivity reasons?



This is induced energy use

Where this talk is going

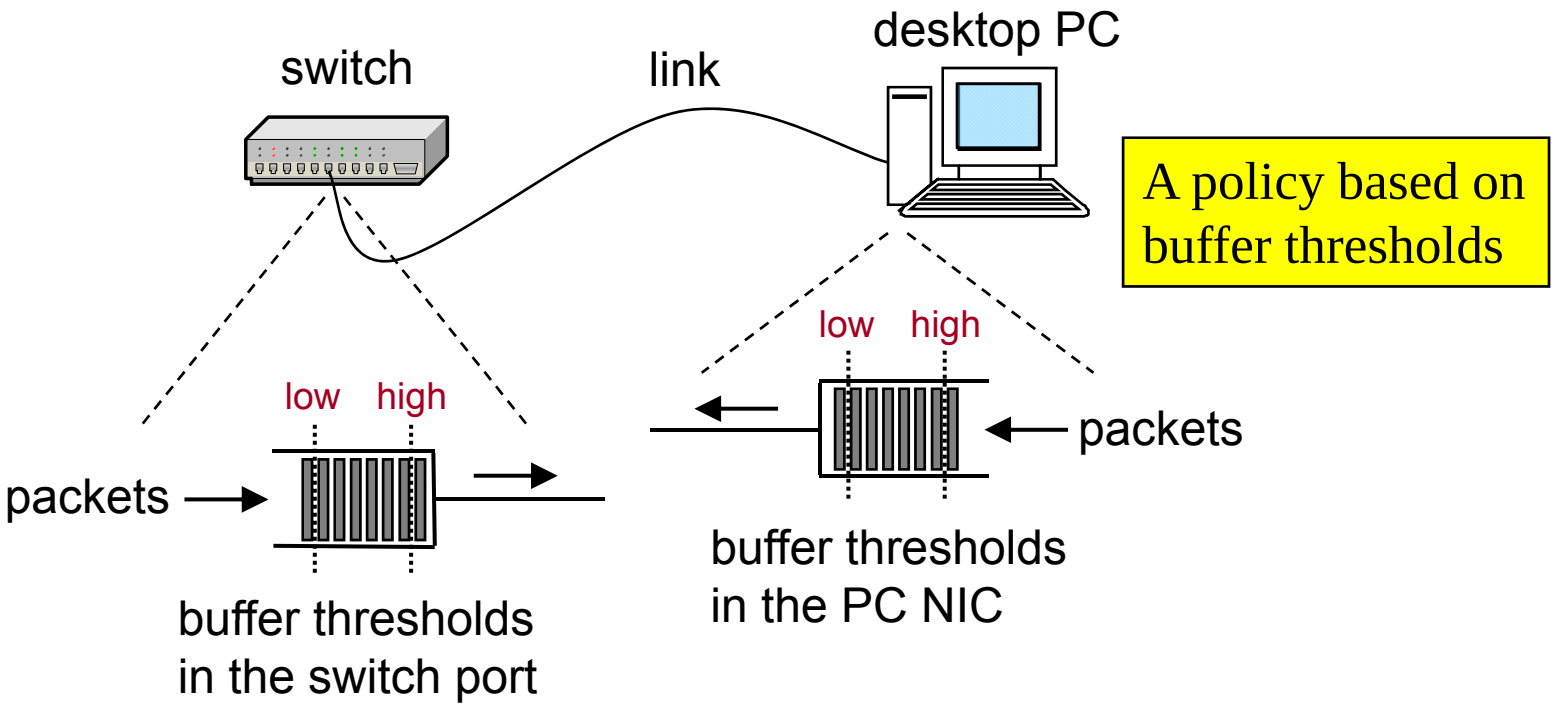
- **Energy consumption by IT equipment**
- **Reducing *direct* energy use**
- **Reducing *induced* energy use**
- **Some additional efforts**

Reducing energy use of links

- **Observation #1: Most edge links are lightly utilized**
  - 1% to 5% on average
- **Observation #2: Higher rates consume more power**
  - About 2 to 4 W for 1 Gb/s versus 100 Mb/s
  - Much more for 10 Gb/s versus 1 Gb/s
- **Idea: Match link data rate with utilization**
- **Key issue: Time to change between data rates**
  - Can buffer overflow occur during transition?
  - What impact might this packet loss have?

Ethernet Adaptive Link Rate (ALR)

- **Two parts to the problem**
  1) *Mechanism* for how to switch link rate
  2) *Policy* for when to change link rate



Ethernet ALR

- **We published the idea and some results**
  - From a 2005 paper and a 2006 whitepaper

2005

INTERNATIONAL JOURNAL OF NETWORK MANAGEMENT
Int. J. Network Mgmt 2005; 15: 297–310
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nem.565

Managing energy consumption costs in desktop PCs and LAN switches with proxying, split TCP connections, and scaling of link speed

By Chamara Gunaratne, Ken Christensen*† and Bruce Nordman

The IT equipment comprising the Internet in the USA uses about $6 billion of electricity every year. Much of this electricity use is wasted on idle, but fully powered-up, desktop PCs and network links. We show how to recover a large portion of the wasted electricity with improved power management methods that are focused on network issues. Copyright © 2005 John Wiley & Sons, Ltd.

1. Introduction

A growing expense and impact of the Internet is its energy use. Current estimates are that 2% of electricity consumption in the USA goes to powering the Internet.[1] In Germany it is estimated that energy consumption by IT equipment will be between 2% and 5% in 2010.[2] The 2% estimate for the USA totals more than 74 TWh/year or $6 billion per year. It is predicted that energy use of IT equipment is growing faster than energy use of any other type within buildings.[3] Much of this energy use is wasted. Energy use by IT equipment is not proportional to utilization of the equipment. A recent study by Lawrence Berkeley National Laboratory (LBNL) showed that 60% of all desktop PCs in commercial buildings remain fully powered-on during nights and weekends[4] with existing power management almost always disabled. Beyond the PC are the Ethernet link and workgroup switch. At present, these energy consumers have almost no means of power management. Existing Internet protocols including discovery and routing are also 'energy unaware'; future protocols need to be made energy aware. For existing protocols that cannot be changed, methods of accommodating current operation must be developed. In previous work we have shown that there exists the potential for savings of *billions of dollars per year* in the USA alone.[5–7] These savings are summarized in Section 6 of this paper. Energy costs are a part of the total cost of ownership of an IT operation. Savings in these costs are of interest to IT managers and companies are beginning to respond with network management products (such as Verdiem with its centralized power management controller[8]) to address this need.

An efficient device consumes energy proportional to its output or utility. Thus, an idle or lightly utilized PC or Ethernet link should not consume the same energy as one that is highly utilized. In this paper, we develop several new methods to reduce energy consumption of PCs,

Chamara Gunaratne is a graduate student in the Department of Computer Science and Engineering at the University of South Florida.

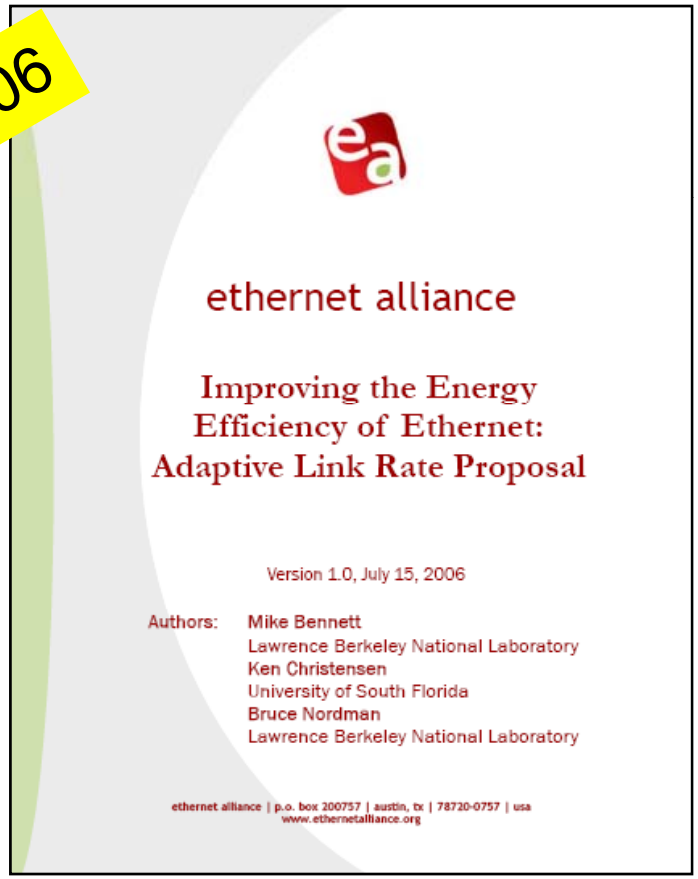
Ken Christensen is an Associate Professor in the Department of Computer Science and Engineering at the University of South Florida.

Bruce Nordman is a Principal Research Associate in the Energy Analysis Department, Environmental Energy Technologies Division, of Lawrence Berkeley National Laboratory Berkeley, California.

*Correspondence to: Ken Christensen, Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620, USA
†E-mail: christen@csee.usf.edu

Copyright © 2005 John Wiley & Sons, Ltd.

2006



ethernet alliance

Improving the Energy Efficiency of Ethernet: Adaptive Link Rate Proposal

Version 1.0, July 15, 2006

Authors: Mike Bennett
Lawrence Berkeley National Laboratory
Ken Christensen
University of South Florida
Bruce Nordman
Lawrence Berkeley National Laboratory

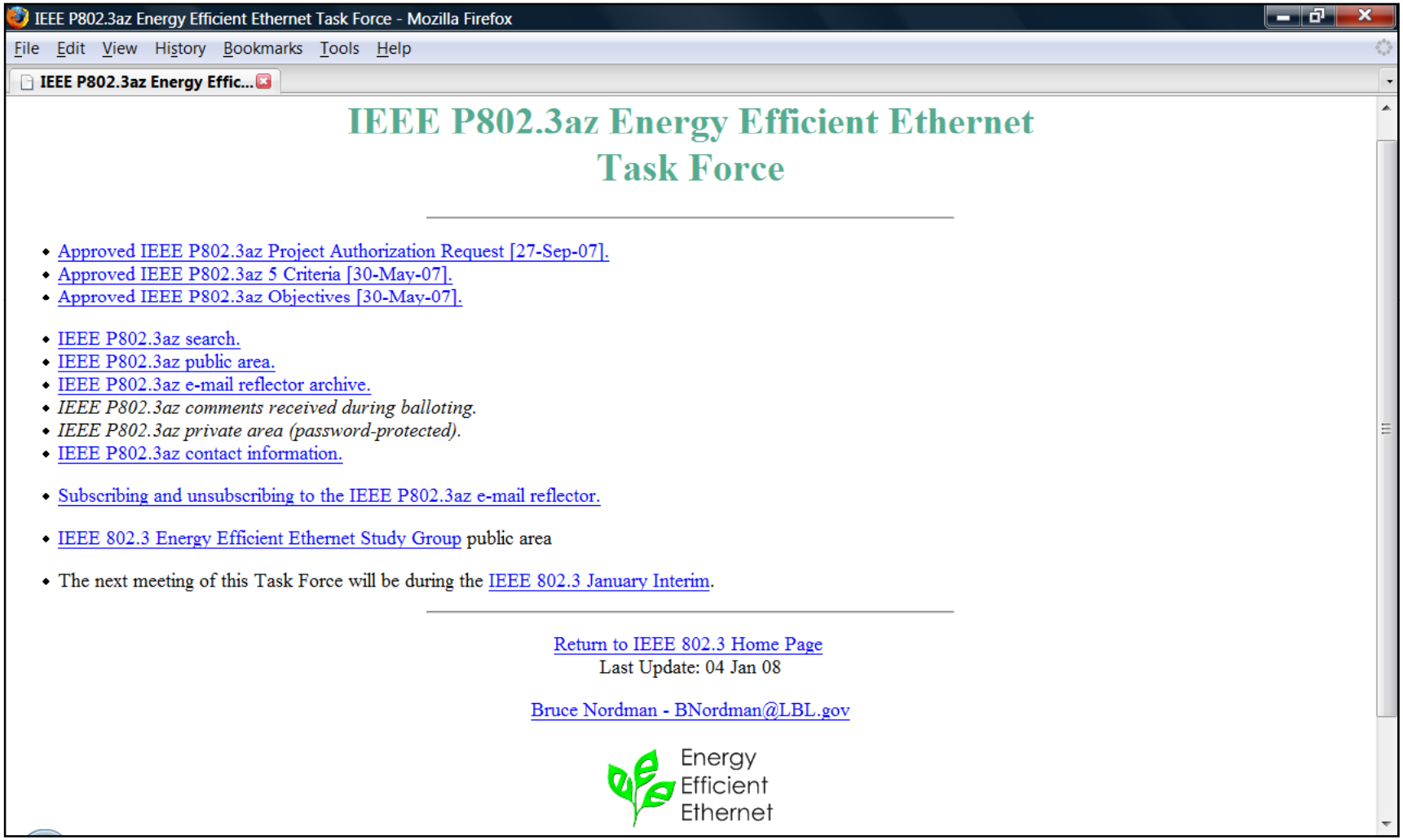
ethernet alliance | p.o. box 200757 | austin, tx | 78720-0757 | usa
www.ethernetalliance.org

Work done by other people…

- **ALR found its way into an IEEE 802.3 study group**
  - Mike Bennett from LBNL as chair
- **Got named "Energy Efficient Ethernet"**
- **Became Rapid PHY Selection (RPS)**
- **Much discussion on switching times**
- **Much work on mechanisms**
- **Some work on policies**
  - My students and I did work here

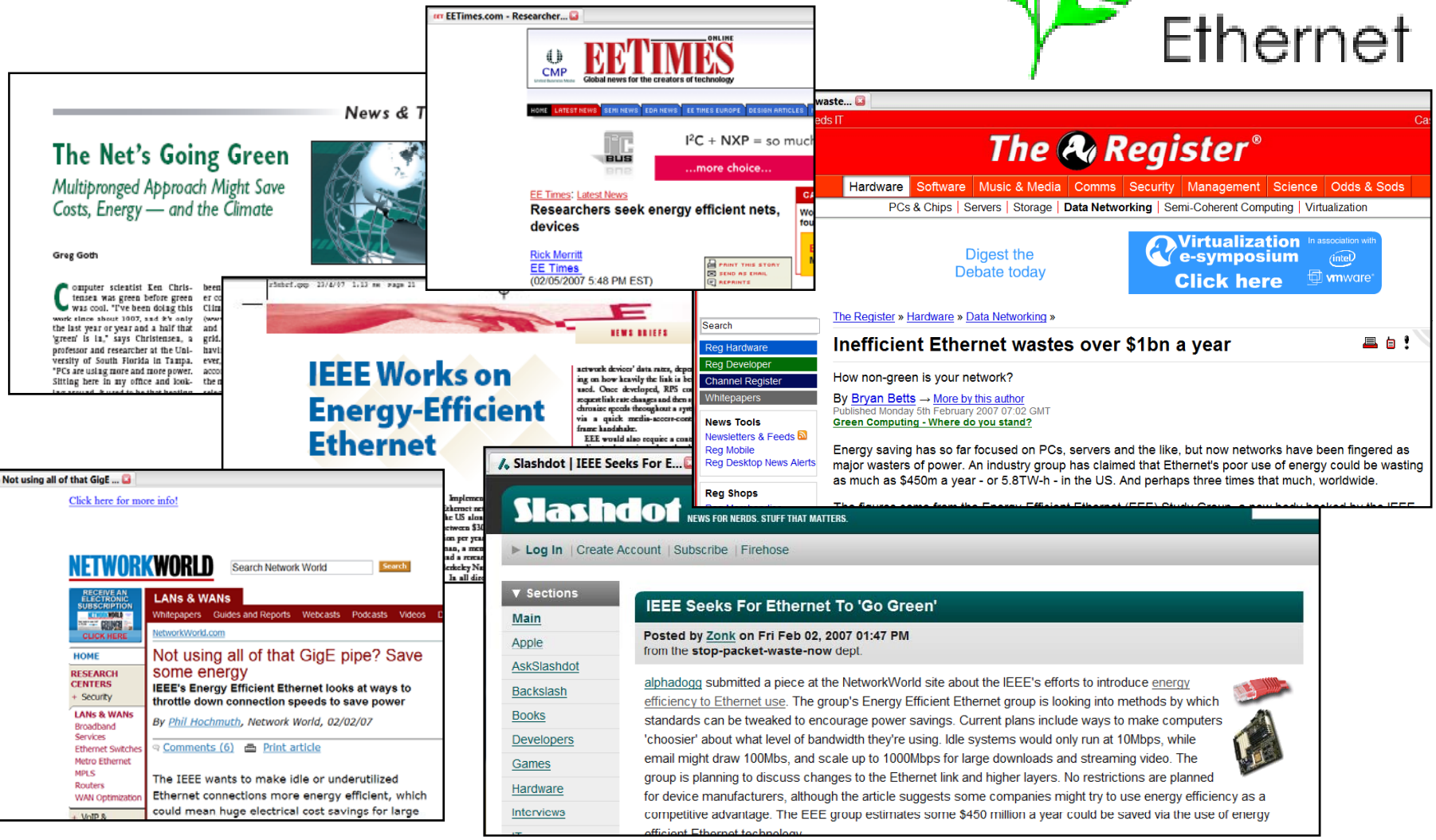
IEEE 802.3az task force

http://www.ieee802.org/3/az/index.html



Some press and a logo…

Energy * Efficient Ethernet



* Logo by Glen Kramer of Teknovus, Inc. (full permission for use granted via email dated January 27, 2007)

Where is EEE now…

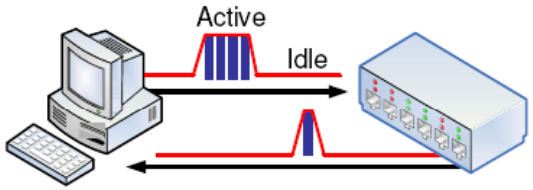
- **EEE is moving forward to becoming a standard**
  - PAR was approved in fall 2007
  - 2009 timeline for completion
  - Is in EPA Energy Star Version 5.0 Draft 1 for computers
- **Current proposal is for "Active-Idle" approach**
- **Active-Idle idea is from Intel**
  - Use a low-power idle between packets
  - Switch to high data rate when a packet is queued
  - Can stay in high data rate if desired
  - About *10 microseconds* to transition out of low-power idle

Active-Idle…

- **Slide from November 2007 802.3az meeting…**

7

Active/Idle Toggling with 0BASE-x Concept



- Principle: Transmit data at fastest rate then return to idle
  - Energy savings come from power cycling between active/idle states
- Active/Idle toggling could be used *instead* of PHY rate shifting
  - Offers the best energy efficiency on links with lower utilization
  - Integrates well with existing PC power management schemes (e.g. ACPI)
  - Clock & power gating (on/off) is easier than rate shifting
- Asymmetrical operation would provide even better energy efficiency
  - Each direction could enter active & idle states independently
  - Most end-node traffic is heavily weighted toward either send or receive
  - Tx & Rx data paths already operate independently above the PHY

Energy Efficient Ethernet

intel

EEE could enable deeper savings

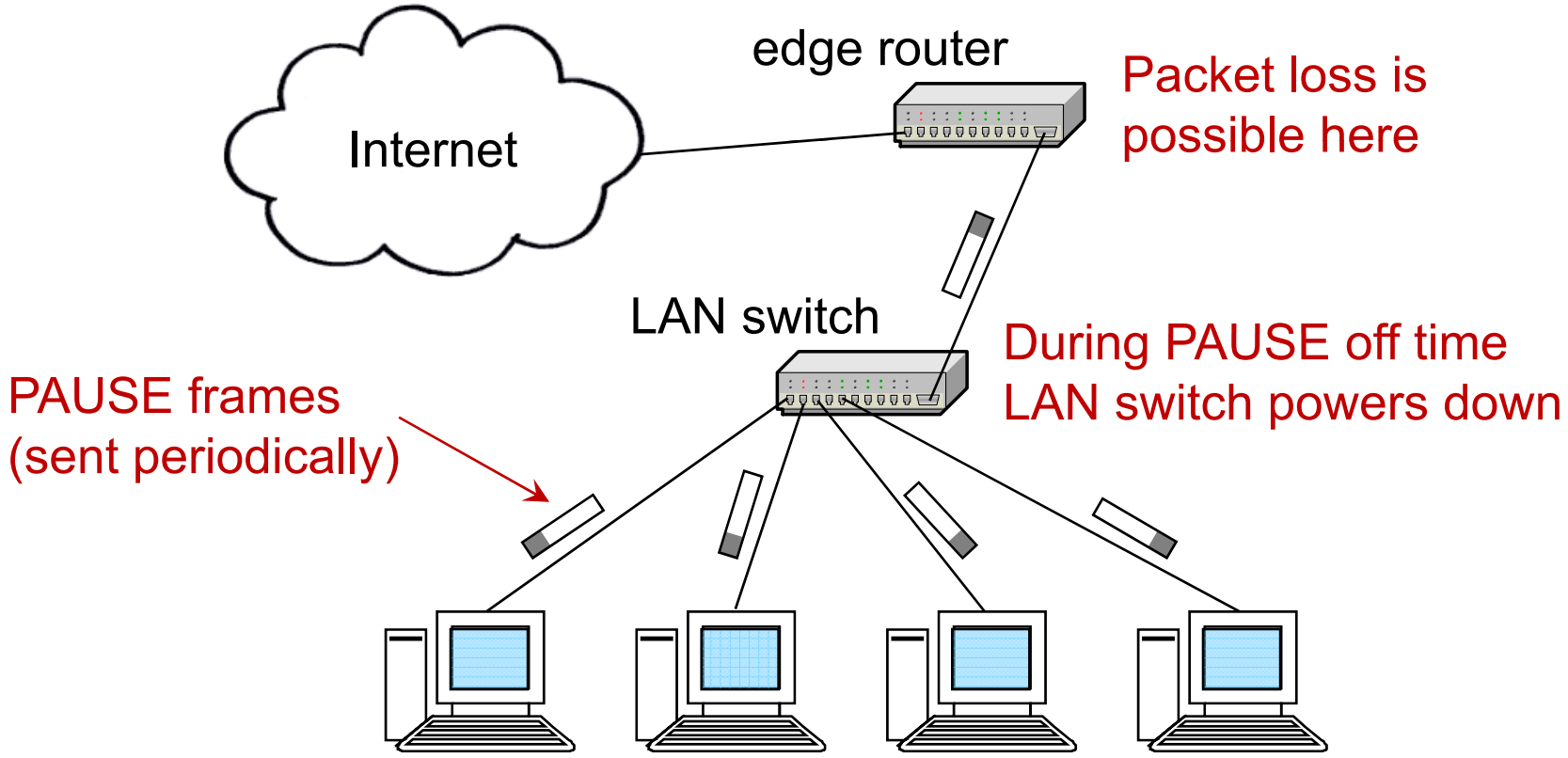
- **Greater savings than from link and PHY/MAC**
  - Savings *within* switches, routers, and servers
- **Can a reduced link data rate allow for…**
  - Powering-down of components?
  - Clocking-down of components?
- **Is some sort of signaling needed to control transition time of link data rate?**

What is next?

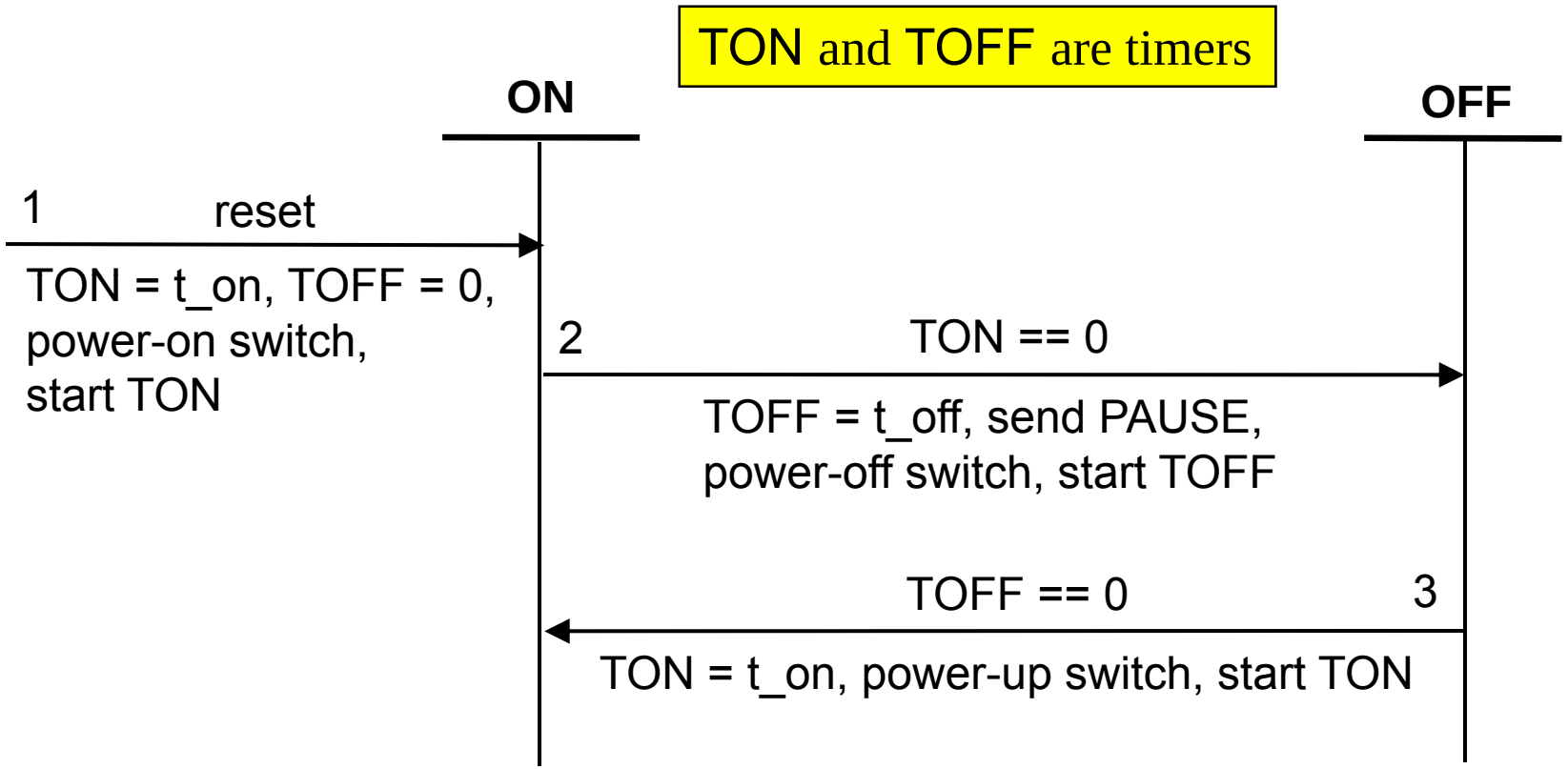
- **EEE requires both ends to participate**
  - Full EEE deployment is still many years off
- **Can we do something simpler?**
  - Something backwards compatible?
- **<u>Idea</u>: PAUSE Power Cycle (PPC)**
  - Use PAUSE to proactively cycle links on and off
  - During link off time power-down LAN switch

PAUSE Power Cycle (PPC)

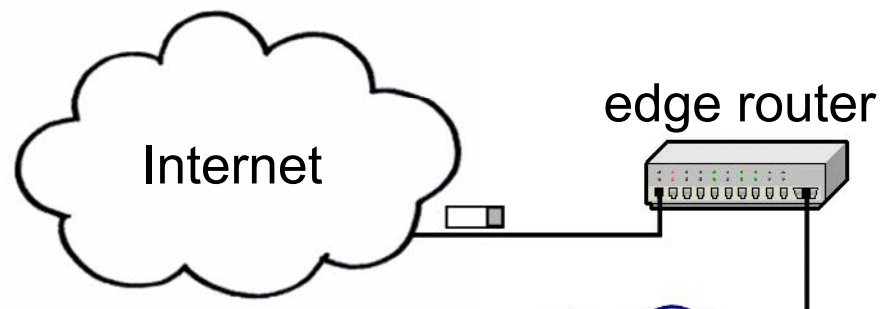
- **Basic idea is to periodically send PAUSE frames**
  - Power "stuff" down during PAUSE interval



PPC Finite State Machine



PPC animation



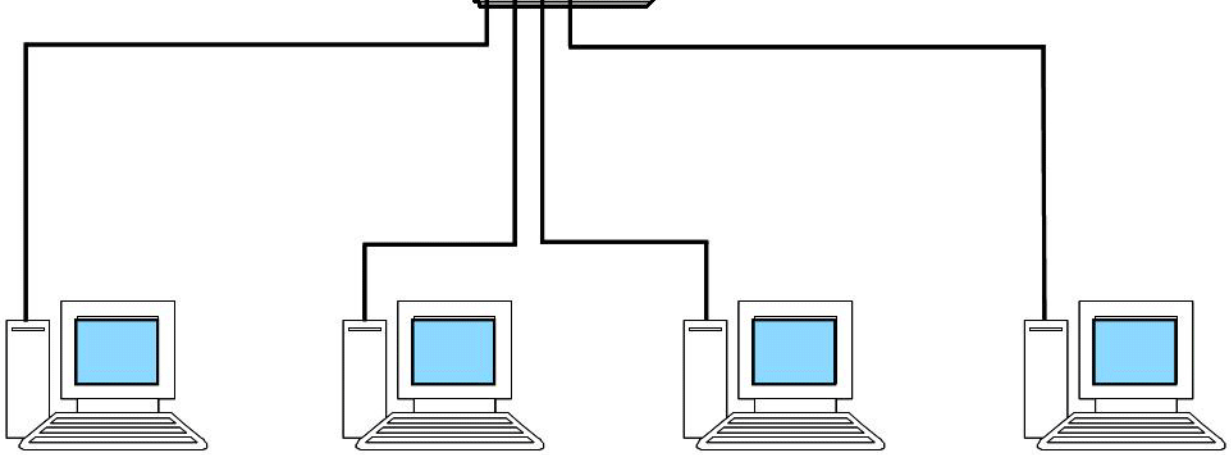
edge router

Packet loss is possible here

PAUSE frames (sent periodically)

LAN switch

During PAUSE off time LAN switch powers down



PPC parameters

- **Key parameters**
  - t_{off} = time off (the PAUSE quanta time)
  - t_{on} = time on
  - D = duty cycle

$$D = \frac{t_{on}}{t_{on} + t_{off}} \qquad t_{on} = \frac{D \cdot t_{off}}{1 - D}$$

- **Time between PAUSE frames is t_{on} + t_{off}**

Does PPC work?

- **We have *emulated* PPC in a test bed**
  - Test bed looks sort of like the previous figure
  - Use a PC to send PAUSE packets through a repeater
  - All links were 100 Mb/s
- **Used human subjects to assess effects**
  - For file transfer, web surfing, and real-time video
  - Subjective evaluation (MOS score approach)
    - Rating from 5 (excellent) to 1 (unacceptable)
- **Real time video was 10 frames/sec (about 1 Mbps)**
  - Axis 2100 camera

PPC evaluation

- **Experimented with 50% duty cycle**
  - t_{off} = 50, 100, and 300 milliseconds
  - $t_{on} = t_{off}$
- **Seven human subjects (all students)**
- **For web browsing**

t_{off} = 50 ms → MOS score of 4.2
t_{off} = 100 ms → MOS score of 3.9
t_{off} = 200 ms → MOS score of 3.1

Not really sure what this means ☹

PPC evaluation continued

- **For real time video**
  - t_{off} = 50 and 100 ms had no problems!
  - t_{off} = 300 ms had some problems (t_{off} > interframe time)

Artifact

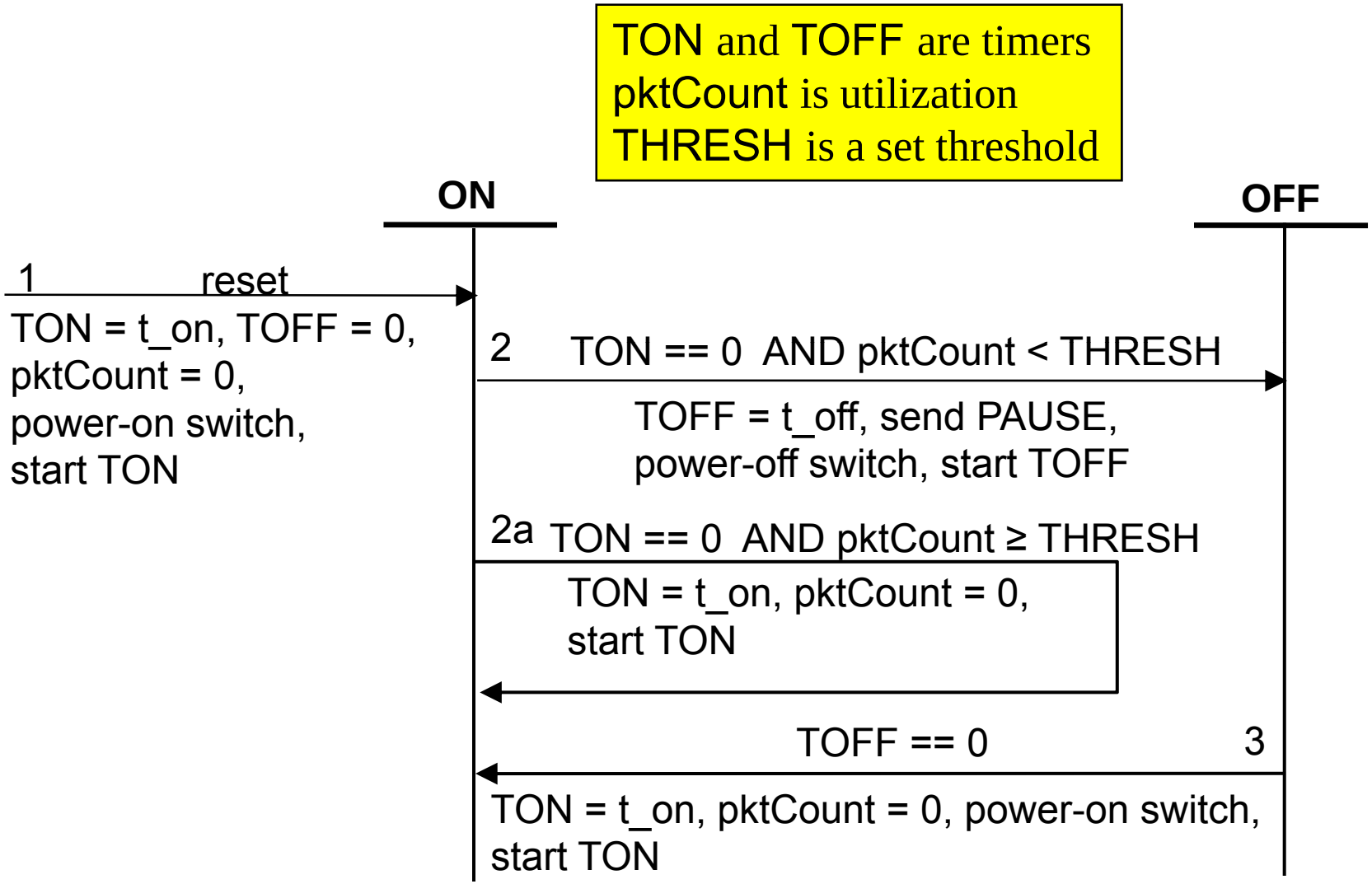
PPC next steps

- **Better evaluation – Experiment**
  - Better user study (look for just noticeable thresholds)
  - Measure packet loss at edge router in test bed
- **Better evaluation – Simulation**
  - ns2 simulation for larger and more realistic networks
    - Assess impact of PPC on higher layer protocols/apps
- **Explore an adaptive policy**
  - Change when to sleep based on utilization?
  - Change t_{off} and/or t_{on} based on utilization?

Simple adaptive PPC policy

- **Basic adaptive policy description**
  - If utilization is high, do not sleep and stay powered on
  - t_{on} and t_{off} timer values are fixed
- **Policy characteristics**
  - Handles high utilization well
  - Does not handle low utilization well
    - Could sleep more

Simple adaptive PPC policy FSM



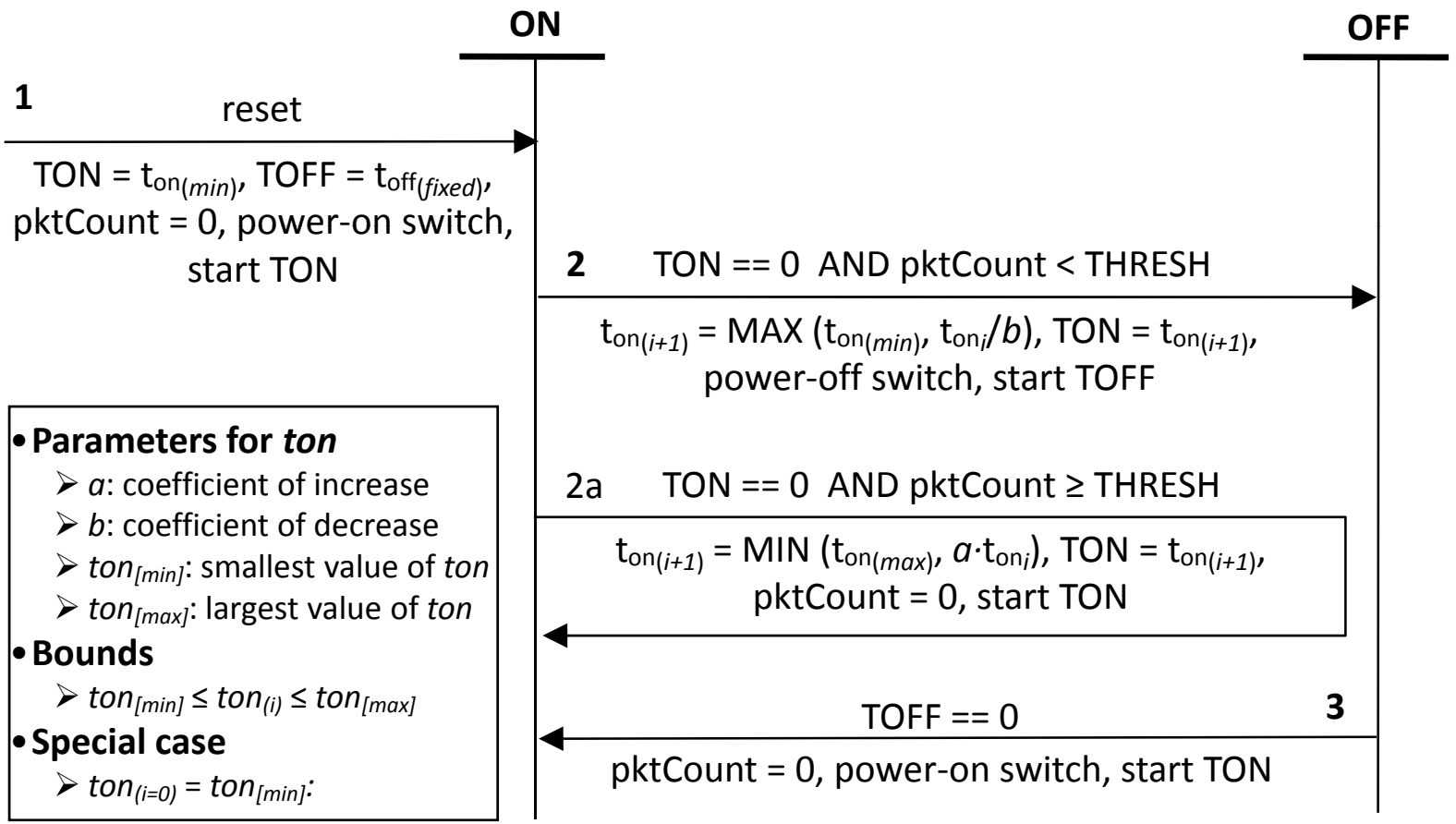
Complex adaptive PPC policy

- **Policy for changing t_{on}**
  - Change t_{on} based on utilization
  - If utilization is high, increase t_{on}
  - If utilization is low, reduce t_{on}

- **Policy characteristics**
  - t_{on} trails and adapts to utilization
  - Based on additive increase multiplicative decrease

Complex adaptive PPC policy FSM

Still under work…



- **Parameters for *ton***
  - *a*: coefficient of increase
  - *b*: coefficient of decrease
  - $ton_{[min]}$: smallest value of *ton*
  - $ton_{[max]}$: largest value of *ton*
- **Bounds**
  - $ton_{[min]} \leq ton_{(i)} \leq ton_{[max]}$
- **Special case**
  - $ton_{(i=0)} = ton_{[min]}$:

Where this talk is going

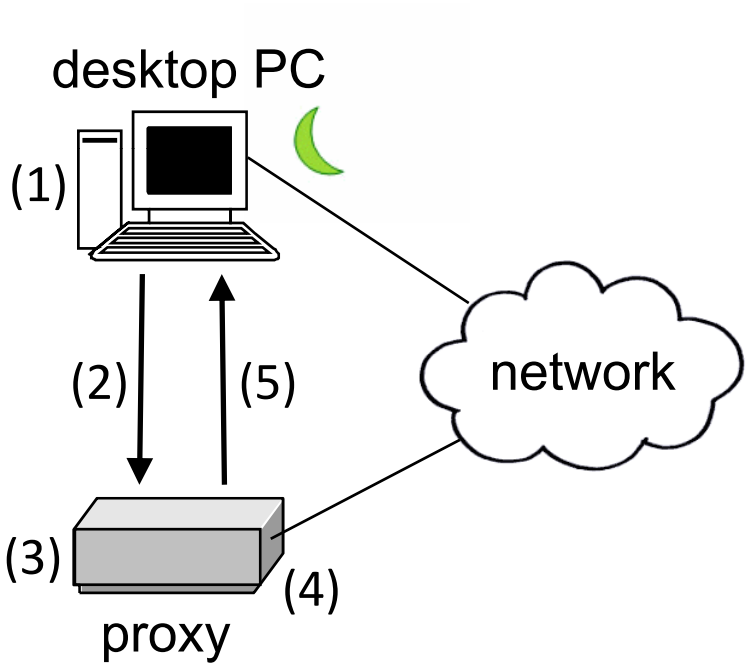
- **Energy consumption by IT equipment**
- **Reducing *direct* energy use**
- **Reducing *induced* energy use**
- **Some additional efforts**

Reducing energy use of edge devices

- **Observation: Devices are often left fully powered-on to maintain network connectivity or presence**
  - Usually devices are not in active use
- **Idea: A low-power network connectivity proxy**
  - Maintains *full network connectivity* for a sleeping device
  - Enable devices to sleep more often
  - Not related to existing Wake-on-LAN
- **Key issues: Lots of issues…**
  - Definition of "connectivity"
  - Wake-ups (not too many, not too few)
  - Packet loss (during wake-up)

Network connectivity proxy

- **Proxy covers for sleeping device**
  - Proxy could be in local NIC or in the network



Steps:

1) PC determines it is time to sleep

2) PC state transferred to proxy

3) PC sleeps, proxy maintains presence

4) Proxy determines need to wake-up PC

5) PC awakes and proxy state transferred

Key protocols that proxying might cover

- **Layers 1 and 2**
  - Already covered by PHY/MAC
- **Layer 3**
  - ARP, ND, ICMP, IGMP, IPSec, etc.
- **Layer 4**
  - TCP connection request (SYN)
  - TCP connections (keep-alives)
- **Higher layers**
  - NetBIOS, SMB, DHCP, SNMP, SSDP, VPN, SSH, etc.
  - Application semantics

Proxying at higher layers

- **Need to support network applications**
- **Two examples:**

  1) Proxying standard for low power UPnP
     - UPnP uses distributed discovery (SSDP)
     - SSDP is lightweight

  2) Current work in proxying for P2P
     - P2P has lots of query traffic, but downloads are rare
     - Query traffic is lightweight

Network connectivity from the EPA

- **Future EPA Energy Star Program Requirements**
  - Version 5.0, Draft 1* (for computers)

> "Computers must maintain full network connectivity while in Sleep mode, according to a platform-independent industry standard."

* From http://www.energystar.gov/index.cfm?c=revisions.computer_spec

Proxying for UPnP – standard

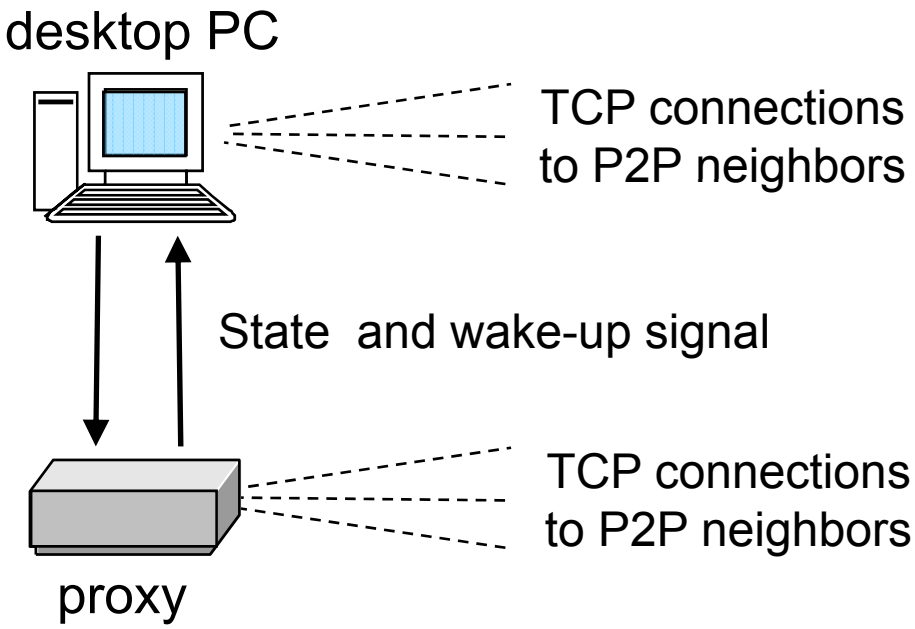
- **UPnP Low Power Architecture (from UPnP Forum)**
  - Version 1.0, August 28, 2007

> "**UPnP Basic Power Management Proxy:** This node will act on behalf of sleeping devices and make sure that the devices are discoverable if they are in low power state. This node will store methods for waking the UPnP Low Power devices."

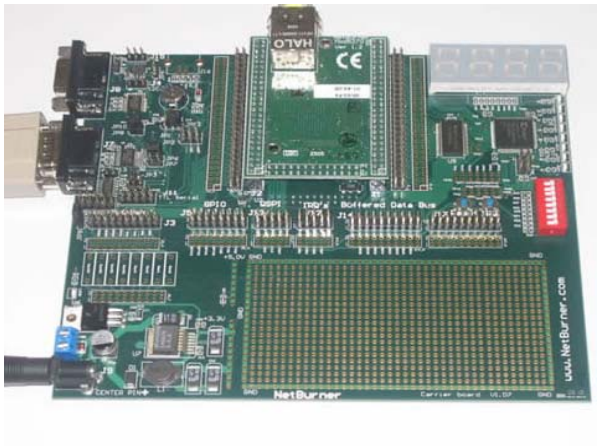
* From http://www.upnp.org/specs/lp/UPnP-lp-Architecture-v1-SDCP-20070828.pdf

Proxying for P2P – in the lab

- **Approach #1 for Gnutella**
  - Move query handling to microcontroller when PC sleeps
  - Wake-up PC when a GET comes in

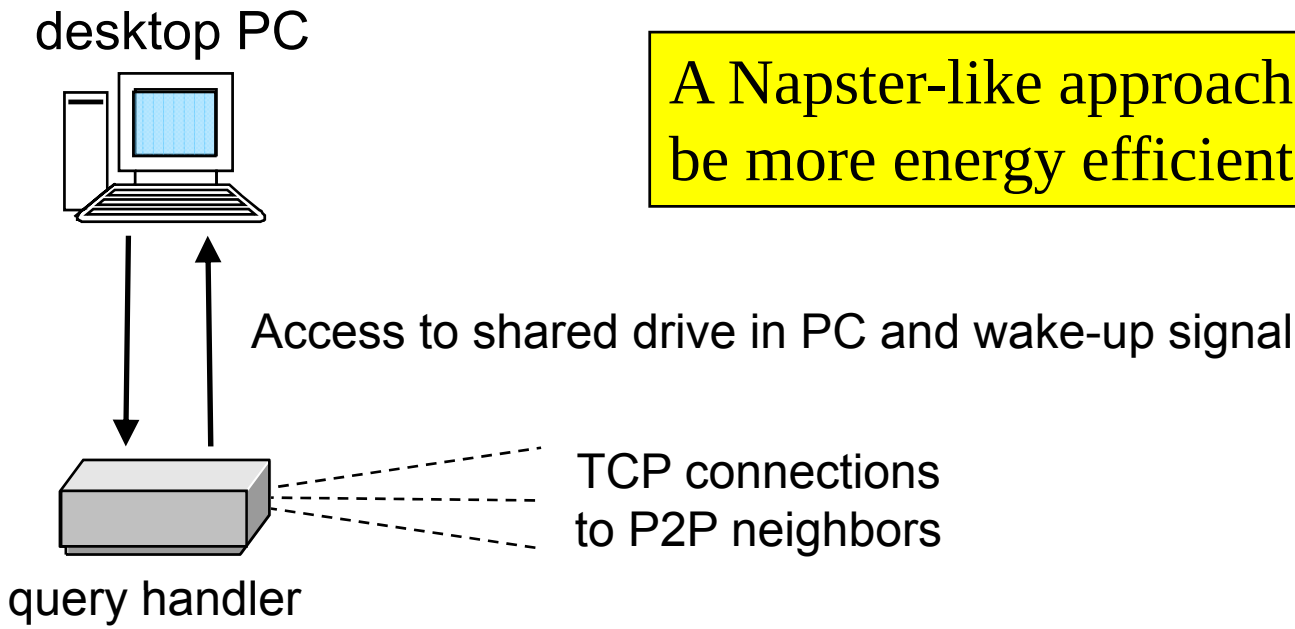


NetBurner microcontroller



Proxying for P2P – in the lab continued

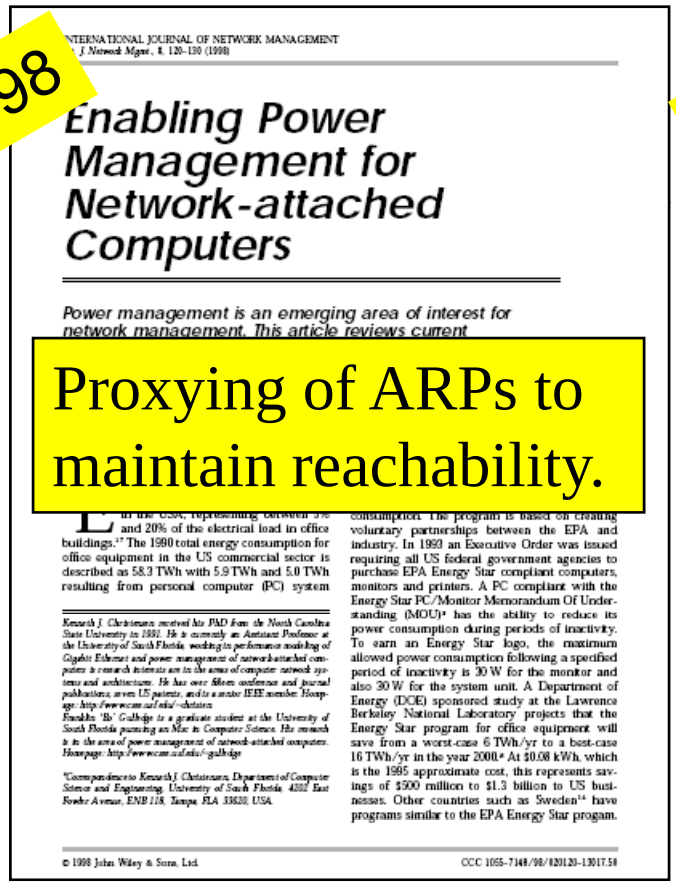
- **Approach #2 for Gnutella**
  - Permanently split query handling and file storage
  - Query handling always running in a smaller device
  - Use a shared drive in PC and wake-up when needed



Proxying for reducing energy use

- **Started to think about proxying 10 years ago**
  - Now hoping to define a direction towards a standard

1998



INTERNATIONAL JOURNAL OF NETWORK MANAGEMENT
Int. J. Network Mgmt., 8, 120–130 (1998)

Enabling Power Management for Network-attached Computers

Power management is an emerging area of interest for network management. This article reviews current

Proxying of ARPs to maintain reachability.

and 20% of the electrical load in office buildings.[1] The 1990 total energy consumption for office equipment in the US commercial sector is described as 58.3 TWh with 5.9 TWh and 5.0 TWh resulting from personal computer (PC) system

consumption. The program is based on creating voluntary partnerships between the EPA and industry. In 1993 an Executive Order was issued requiring all US federal government agencies to purchase EPA Energy Star compliant computers, monitors and printers. A PC compliant with the Energy Star PC/Monitor Memorandum Of Understanding (MOU)[3] has the ability to reduce its power consumption during periods of inactivity. To earn an Energy Star logo, the maximum allowed power consumption following a specified period of inactivity is 30 W for the monitor and also 30 W for the system unit. A Department of Energy (DOE) sponsored study at the Lawrence Berkeley National Laboratory projects that the Energy Star program for office equipment will save from a worst-case 6 TWh/yr to a best-case 16 TWh/yr in the year 2000.[4] At $0.08 kWh, which is the 1995 approximate cost, this represents savings of $500 million to $1.3 billion to US businesses. Other countries such as Sweden[14] have programs similar to the EPA Energy Star progam.

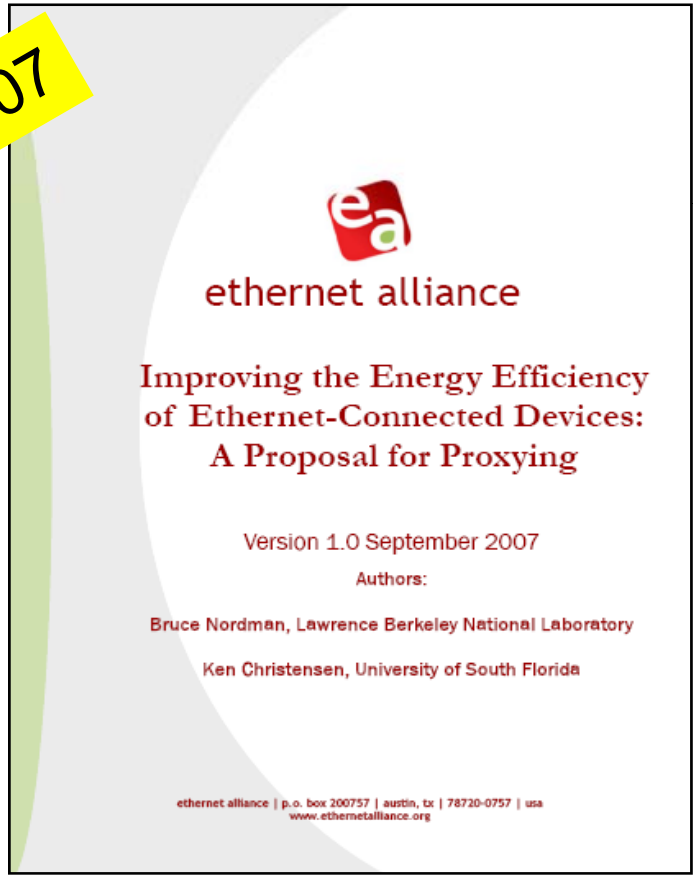
Kenneth J. Christensen received his PhD from the North Carolina State University in 1991. He is currently an Assistant Professor at the University of South Florida working in performance modeling of Gigabit Ethernet and power management of network-attached computers in research interests are in the areas of computer network systems and architectures. He has over fifteen conference and journal publications, seven US patents, and is a senior IEEE member. Homepage: http://www.csee.usf.edu/~christen

Franklin 'Bo' Gulbrigt is a graduate student at the University of South Florida pursuing an MSc in Computer Science. His research is in the area of power management of network-attached computers. Homepage: http://www.csee.usf.edu/~gulbrigt

Correspondence to Kenneth J. Christensen, Department of Computer Science and Engineering, University of South Florida, 4202 East Fowler Avenue, ENB 118, Tampa, FLA 33620, USA.

CCC 1055-7148/98/020120-13$17.50

2007



ethernet alliance

Improving the Energy Efficiency of Ethernet-Connected Devices: A Proposal for Proxying

Version 1.0 September 2007

Authors:

Bruce Nordman, Lawrence Berkeley National Laboratory

Ken Christensen, University of South Florida

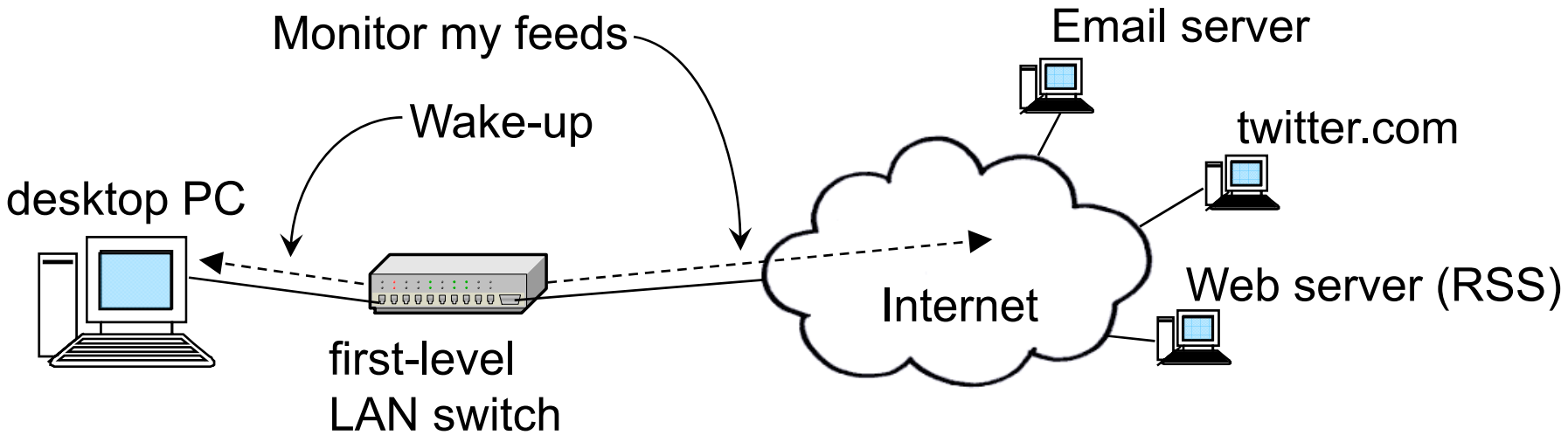
ethernet alliance | p.o. box 200757 | austin, tx | 78720-0757 | usa
www.ethernetalliance.org

What is next?

- **Can "intelligence" in the network help the edge?**
- **Explore making power state of edge devices known**
- **Idea: Support in first-level LAN switch**
  - Not age-out entries in ARP cache for sleeping PCs
  - Filter packets to eliminate spurious wake-ups
  - Reply to simple protocols
  - Generate packets for simple protocols
  - Wake-up PC when appropriate
  - Maybe even assist applications to sleep

Proxying in a switch

- **Can a switch filter packets for a sleeping PC?**
  - Use existing packet inspection capabilities
- **Can a switch assist applications in a sleeping PC?**
  - Monitor email, twitter, IM, RSS feeds, etc.
  - Wake-up PC when something interesting happens



Next steps for switch proxying

- **First…**
- **What applications should be target for this?**
  - A protocol and application classification is necessary
- **Classification Criteria:**
  - Is the protocol/application widely used?
  - Is it chatty?
  - Can we divide the functionality of the application? This is useful to put certain functionalities in the proxy

Next steps continued

- **Second…**
- **Existing capabilities that could be useful?**
  - Deep Packet Inspection (DPI) used in many routers
  - Existing protocol agents running on routers
- **DPI**
  - We plan to use it to detect traffic from applications previously classified
- **Existing Protocol Agents**
  - Structure can be used for proxiable applications

Next steps continued

- **Third…**
- **Prototype the proxy-in-a-switch**
- **Use open source router implementations**
  - Vyatta
  - XORP
- **Develop protocol agents to run on above routers**

Where this talk is going

- **Energy consumption by IT equipment**
- **Reducing *direct* energy use**
- **Reducing *induced* energy use**
- **Some additional efforts**

SNMP Power MIB

- **We want to expose and control power state**
- **Can we use SNMP to do this?**
- **Developing a Power MIB for desktop PC**
  - Power management capabilities
  - Power management settings
  - Total time for idle, busy, and sleep
  - Current elapsed time for idle, busy, or sleep
  - Statistics on wake-up events (network, user, etc.)
  - Statistics on sleep events
  - Actual power and energy use if a meter is installed?!

Green telnet

- **Telnet (SSH, etc.) ties state to TCP connection**
  - This effectively prevents client from going to sleep
- **Can we disconnect and buffer data in the server?**
- **Developing a gtelnetd and client for Linux**
  - Server buffers data when clients goes to sleep
  - Client reconnects when it wakes-up
  - Server delivers buffered data when client reconnects
- **Working on an article for Dr. Dobbs Journal**

Adaptive power management for PCs

- **PC operating systems use inactivity time-out**
  - For example, to put system to sleep
  - Use a fixed value for time-out
- **Can we do better with an adaptive time-out?**
- **Have experimented with using past activity history**
  - Adaptively set the time-out value based on prediction
  - No conclusive results yet
  - Need to do more characterization of PC users

Summary

- **Energy savings can be enabled by the network**

  - **Direct energy savings**
    - Matching link data rate to link utilization

    ➡ How can we exploit this for deeper savings?

  - **Induced energy savings**
    - Letting devices sleep without losing network presence

    ➡ How can we enable this for all end devices?

Acknowledgments

- **Folks who contributed…**
  - Bruce Nordman
  - Francisco Blanquicet
  - Miguel Jimeno
  - Jakob Klamra
  - Jeremy Blackburn
  - Others

Some of my students

Thank you!

Questions?

Ken Christensen
Department of Computer Science and Engineering
University of South Florida
Tampa, FL 33620
christen@cse.usf.edu

Key links

- **Project website**
  - http://www.csee.usf.edu/~christen/energy/main.html
- **Project publications and invited talks**
  - http://www.csee.usf.edu/~christen/energy/pubs.html
- **ALR whitepaper at Ethernet Alliance**
  - http://www.ethernetalliance.org/technology/white_papers/alr_v10.pdf
- **Proxying whitepaper at Ethernet Alliance**
  - http://www.ethernetalliance.org/technology/white_papers/Proposal_for_Proxying_edit.pdf
- **IEEE 802.3az taskforce**
  - http://www.ieee802.org/3/az/index.html
- **UPnP Forum Low Power V 1.0**
  - http://www.upnp.org/specs/lp.asp