Reducing the Energy Consumption of Networked Devices

Bruce Nordman
Energy Analysis
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
bnordman@lbl.gov

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

Topics

- Energy use by IT equipment Part 1
- Overview of power management Part 2
- Reducing network induced energy use Part 3
- Reducing network direct energy use Part 4
- Potential energy savings Part 5
- Summary and next steps Part 6

Background - Key Terms

Network Equipment

- Products whose only function is to enable network communications (Switches, routers, firewalls, modems, etc.)

Networked Device

- An electronic product with digital network connection — a piece of network equipment or end use device.

NIC

- Network Interface Controller.

Energy

- Direct AC electricity consumed by electronic devices. Does not include extra space conditioning energy, UPS, etc.

3

Energy use by IT equipment

❖ **Welcome to Part #1**

In this part…the energy consumption of IT generally and PCs specifically.

4

Current IT energy use: All IT equipment

- **"Big IT" – all electronics**
  - PCs/etc., consumer electronics, telephony
    - Residential, commercial, industrial
  - **200 TWh/year**
  - $16 billion/year
  - Nearly 150 million tons of CO_2 per year

PCs and etc. already digitally networked — *Consumer Electronics* (CE) will be soon

One central baseload power plant (about 7 TWh/yr)

5

Current IT energy use: All IT equipment continued

- **" Little IT" — office equipment, telecom, data centers**
  - **97 TWh/year** (2000) [Roth] — 3% of national electricity; 9% of commercial building electricity

Commercial buildings only

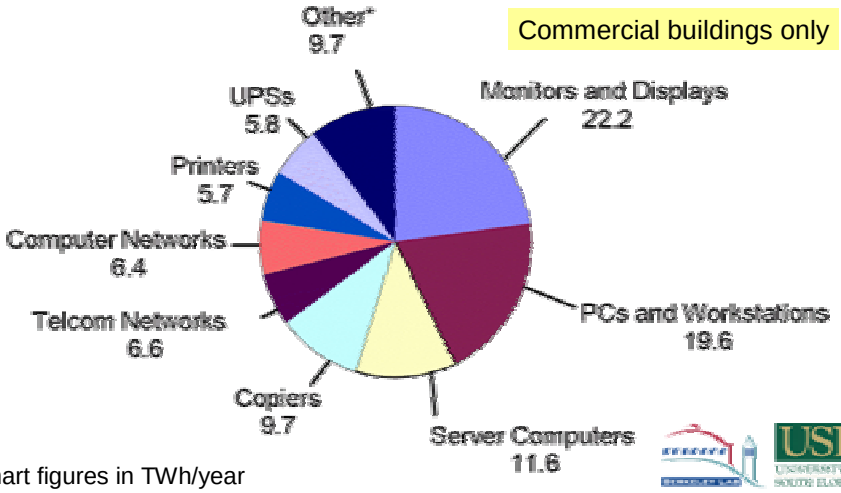


Chart figures in TWh/year

6

PC energy use

❖ **PCs**

- Computing box only — not including displays
- Year 2000:
  - PCs: **31 TWh/year**
  - Servers: 12 TWh/year
- Today:
  - PCs could be **46 TWh/year** and is rising steadily

➔ **$2.4 billion/year**

➔ **$3.7 billion/year**

7

PC energy use: 24/7 PC example

❖ **Bruce's home PC and display***

	On	Sleep	Off
Computer	**57.5 W**	**7.5 W**	**6.0 W**
Display	**17**	**2**	**2**

- Display can power manage – On 20 hours/week; Sleep 148
- Computer can't **(and stay on network)** – On 168 hours/week

❖ **Annual consumption**

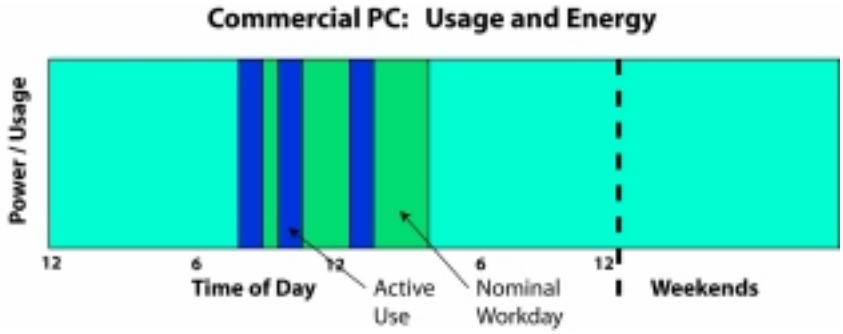
- 540 kWh/year
- ~$70/year

16% of current annual electricity bill

8

* Bruce doesn't leave the PC on 24/7

PC energy use: How PCs use energy



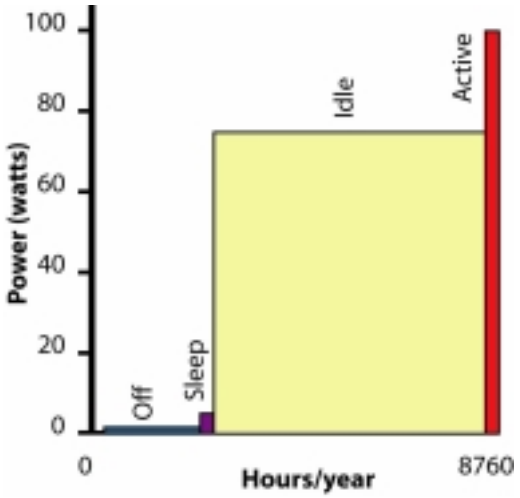
- **Active use is a small part of week**
  - Energy use is not closely related to activity
- **Most commercial PCs are on continuously**
  - Increasingly true for residential PCs
  - Most of time, highly powered but doing little or no work

Savings opportunity!

9

PC energy use: Structure

Typical Commercial PC Annual Energy Use



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

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

Consumption is driven by on-times, not by usage

10

PC energy use: Numbers

- **Power levels**
  - 70 W in On (notebooks 20); 5 W in Sleep; 2 W in Off

Desktop vs. Notebook

- **Usage**

	Portion of Stock "Continuous On"	% Sleeping
Commercial	About 2/3 (2003)	6%
Residential	~20% (2001) and rising*	~10% ?

  - Most home PCs in homes with >1 PC
  - Home broadband penetration rising (~50%)

Now vs. Future

➔ > 50% on 24/7

- **Stock**

Home vs. Office

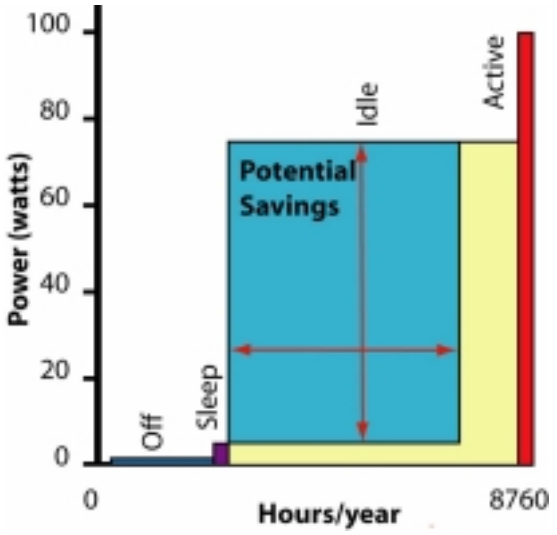
  - Roughly 100 million each residential and commercial

National PC energy today ➔ 46 TWh/year

11

* Half of these on 40-167 hours/week

PC energy use: "Waste" / Savings opportunity



Most of time when idle, could be asleep;
PC savings potential is most of current consumption

12

Energy Star context

- **1992 — Began with PC and monitor power mgmt.**
  - Capability to PM; sleep/off levels
- **1999 — Reduced power levels; addressed network connectivity**
- **Current specification revision process**
  - Power supply efficiency
  - Limits on system "idle" power
  - Network connectivity in Sleep
- **Could play a key role in reducing energy use from networks**

13

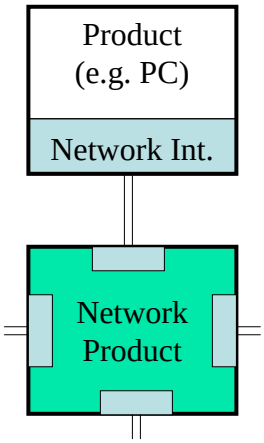
Network equipment energy use continued

- **Switches, Hubs, Routers** (commercial sector only)
  - 6.05 TWh/year — 2000 [Singh] ➔ ~$500 million/year
- **Telecom equipment** (mobile, local, long distance, PBX)
  - 6.1 TWh/year — 2000 [Roth] ➔ ~$500 million/year
- **NICs alone — Quick Estimate**
  - 300 million products with NICs; NIC at both ends
  - 1 W per NIC; Continuous use
  - ➔ 600 MW NIC power; ➔ 5.3 TWh/year
    ➔ > $400 million/year

14

Network direct and induced energy use

- **Network Direct**
  - NICs
  - Network Products
    - Switches, Routers, Broadband Modems, Wireless Access Points, …
- **Network Induced**
  - Increment for higher power state of devices needed to maintain network connectivity (usually On instead of Sleep or Off)
  - Common causes:
    - Can't maintain needed connectivity
    - Too cumbersome to set up or use



15

Overview of power management

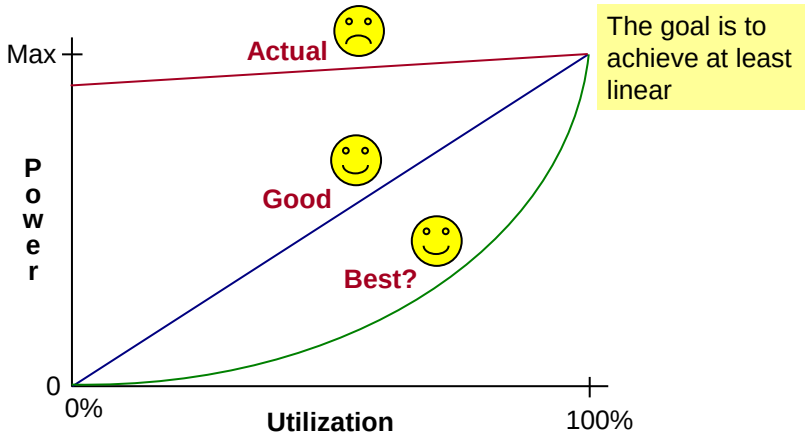
- **Welcome to Part #2**

In this part… an overview of power management, wake on LAN, and current technology directions.

16

Power and utilization

- **Power use should be proportional to utilization**
  - But it rarely is!



17

Basic principles of power management

- **To save energy we can:**
  - Use more efficient chips and components
  - Better power manage components and systems

- **To power manage we have three methods:**

Do less work (processing, transmission)
- Transmitting is very expensive in wireless

Slow down
- Process no faster than needed

Turn-off "stuff" not being used
- Within a chip (e.g., floating point unit)
- Within a component (e.g., disk drive)
- Within a system (e.g., server in a cluster)

18

Power management in PCs

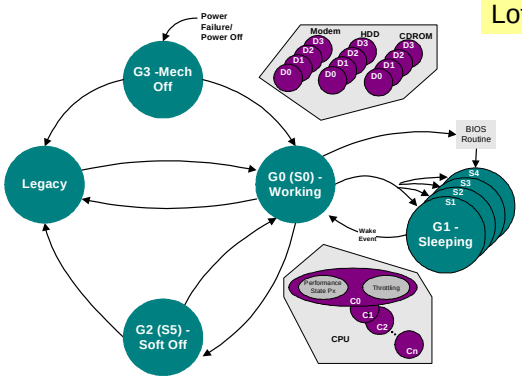
- **PCs support power management (via ACPI)**
  - For conserving batteries in mobile systems
  - For energy conservation (EPA Energy Star compliance)
- **How it works …**
  - Use an inactivity timer to power down
  - Power down monitor, disks, and eventually the entire system
    - Sleep (Windows *Standby)* and *Hibernate*
  - Resume where left-off on detection of activity
    - Mouse wiggle or key stroke to wake-up
    - Wake on LAN packet

19

Power management in PCs continued

- **Advanced Configuration and Power Interface (ACPI)**
  - ACPI interface is built-in to operating systems
    - An application can "veto" any power down

Lots of states!



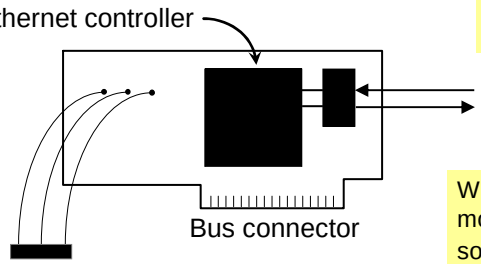
20

* From page 27 of ACPI Specification (Rev 3.0, September 2, 2004)

Wake on LAN

- **Wake on LAN (WOL)**
  - A special MAC frame that a NIC recognizes
    - Called Magic Packet (by AMD)
    - Developed in mid 1990's
    - Intended for remote administration of PCs

All this is now on the motherboard and PCI bus.



WOL implemented in most Ethernet and some WiFi NICs.

Cable and connector for auxiliary power and wake-up interrupt lines

21

Wake on LAN continued

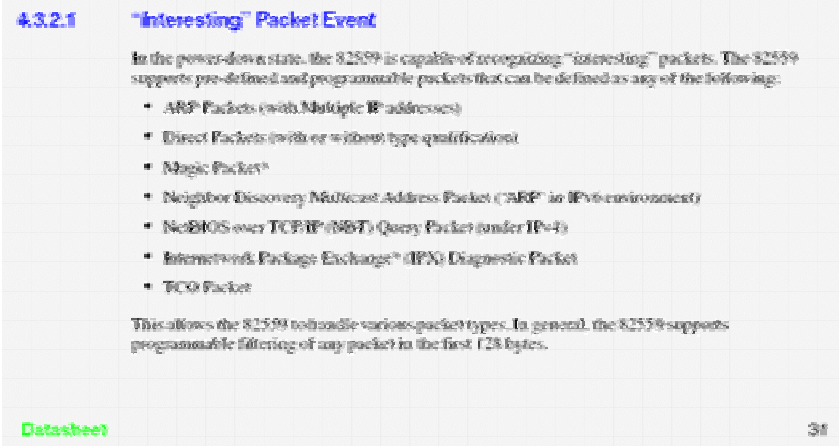
- **WOL has shortcomings…**
  - **Must know the MAC address of remote PC**
  - **Cannot route to remote PC due to last hop router timing-out and discarding ARP cache entry**
  - **Existing applications and protocols do not support WOL**
    - For example, TCP connection starts with a SYN

PC doesn't wake up enough.

22

Directed packet wake-up

- **A better WOL**
  - Wake on interesting packets and pattern matching*



4.3.2.1 "Interesting" Packet Event

In the power-down state, the 82559 is capable of recognizing "interesting" packets. The 82559 supports pre-defined and programmable packets that can be defined as any of the following:

- ARP Packets (with Multiple IP addresses)
- Direct Packets (with or without type qualification)
- Magic Packet™
- Neighbor Discovery Multicast Address Packet ("ARP" in IPv6 environment)
- NetBIOS over TCP/IP (NBT) Query Packet (under IPv4)
- Internetwork Package Exchange™ (IPX) Diagnostic Packet
- TCO Packet

This allows the 82559 to handle various packet types. In general, the 82559 supports programmable filtering of any packet in the first 128 bytes.

Datasheet 31

23

* From page 31 of Intel 82559 Fast Ethernet Controller datasheet (Rev 2.4)

Directed packet wake-up continued

- **Directed packet wake-up has shortcomings…**
  - **Wake-up on unnecessary or trivial requests**
    - "Wake on Junk"
  - **Not wake-up when need to**
  - **Needs to be configured**

PC wakes up too much and/or too little.

24

Reducing network-induced energy use

- **Welcome to Part #3**

In this part… the "sleep-friendly" PC – its motivation, requirements, design, and next steps.

→ Goal is to reduce network *induced* energy use

25

Disabling of power management

- **The issue is disabling of power management in PCs**
- **Why are many PCs fully powered-on "all the time"?**
  - Historically for reasons of poor performance
    - Crash on power-up, excess delay on power-up, etc.
  - Today increasingly for network-related reasons

Increasing number of applications are network-centric

26

Disabling for applications

- **Some applications require a PC to be fully powered-up**
- **Some examples…**
  - File access on a remote network drive
  - Remote access for maintenance
  - Remote access for GoToMyPC or Remote Desktop
  - P2P file sharing
  - Some VPN
  - Some IM and chat applications
- **Some applications disable sleep**
  - No way to know power status of a remote PC
  - No way to guarantee wake-up of a remote PC

27

Disabling for protocols

- **Some protocols require a PC to be fully powered-up**
  - ARP packets – *must respond*
    - If no response then a PC becomes “unreachable”
  - TCP SYN packets – *must respond*
    - If no response then an application is “unreachable”
  - IGMP query packets – *must respond*
    - If no response then multicast to a PC is lost
  - DHCP lease request – *must generate*
    - If no lease request then a PC will lose its IP address

28

Connections are everywhere

- **Permanent connections are becoming common**
  - At TCP level – “keep alive” messages are exchanged
  - At app. level – app. “status” messages are exchanged
    - Must respond at either level or connection can be dropped

Dropped connection returns user to log-in screen (and messages lost!)

PC goes to sleep

29

A traffic study

- **We traced packets arriving to an idle PC at USF (2005)**
  - Received 296,387 packets in 12 hours and 40 minutes

This is 6 pkts/sec

Protocol	% in trace
ARP	52.5 %
UPnP	16.5
Bridge Hello	7.8
Cisco Discovery	6.9
NetBIOS Datagram	4.4
NetBIOS Name Service	3.6
Banyan System	1.8
OSPF	1.6
DHCP	1.2
IP Multicast	1.0

Remaining 2.7% and less than 1% each we found RIP, SMB, BOOTP, NTP, ICMP, DEC, X display, and many others

30

A sleep-friendly PC

What capabilities would a sleep-friendly PC need?

- Maintain network presence with little or no wake-up of PC
- Generate routine packets as needed
- Reliably and robustly wake-up PC when needed
- Not wake-up PC when not needed
- Provide for exposing power state to network
- No changes to existing protocols
  - Only minimal changes to applications
- No change in user experience

31

A sleep-friendly PC continued

❖ **Key functions**

1) Ignore
   - Ignore and discard packets that require no action

2) Proxy
   - Auto-reply to trivial requests without need to wake-up PC
   - Auto-generate packets from "trivial" protocols and applications

3) Wake-up
   - Wake-up PC for valid, non-trivial requests

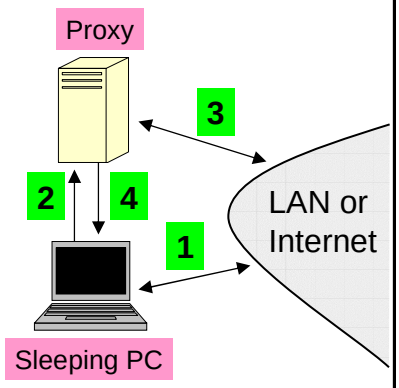
4) Handle TCP connections
   - Prevent permanent TCP connections from being dropped

32

Proxying

❖ **Flow for proxying…**

1. **PC awake; becomes idle**
2. **PC transfers network presence to proxy on going to sleep**
3. **Proxy responds to routine network traffic for sleeping PC**
4. **Proxy wakes up PC as needed**



Proxy can be internal (NIC) or external (in other PC, switch or router, wireless base station, or dedicated device)

33

Options for a Sleep Friendly PC

❖ **Four possible options…**

1) ~~Selective wake-up NICs only~~
   - ~~Such as WOL or direct packet wake-up~~

2) Proxy internal to a NIC
   - We call this a SmartNIC (and includes wake-up)

3) Central proxy in a switch, access point, etc.
   - Build on UPnP proxy idea

4) Very low power fully-operational mode of PC
   - OS and processor active, but operate slowly

SmartNIC is most promising, (3) and (4) can have a role

34

SmartNIC concept

Can we add capability to a NIC such that a PC can remain in a low-power sleep state more than it can today?

Add a few dollars cost to a PC that is recovered in a few months of operation

- **A SmartNIC contains**
  - Proxy capability (*new*)
  - Wake-up capability (*as today and improved*)
  - Ability to advertise power state (*new*)
- **When a PC is powered-down the SmartNIC…**
  - Remains powered-up
  - "Covers" or "proxies" for the PC
  - Wakes-up the PC only when needed
  - Communicates power state as needed

35

Reducing network direct energy use

- **Welcome to Part #4**

In this part… a discussion of how to reduce direct energy use with adaptive link rate.

→ Goal is to reduce network *direct* energy use

36

Power management of a link

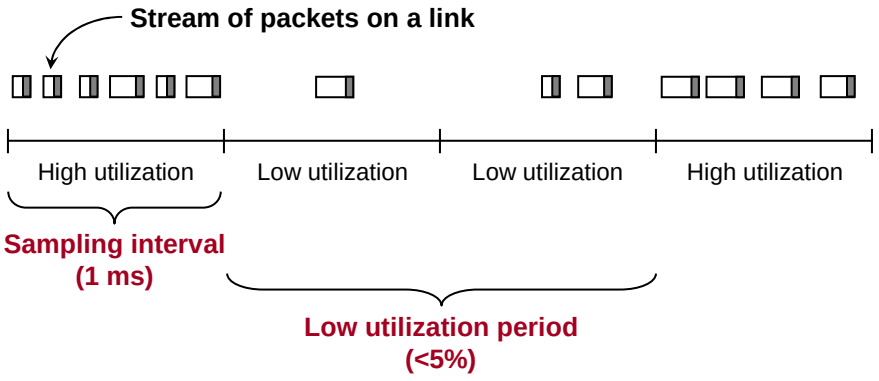
Can we power manage an Ethernet link and NICs?

- **If low utilization, <u>do not need high data rate</u>**
  - Can we switch link data rate?
  - How fast can we switch link data rates?
  - What policies do we use to switch data rates?

37

Low utilization periods

- **Low utilization in a stream of packets**
  - Packets are variable in length (64 to 1500 bytes)



38

Power measurements

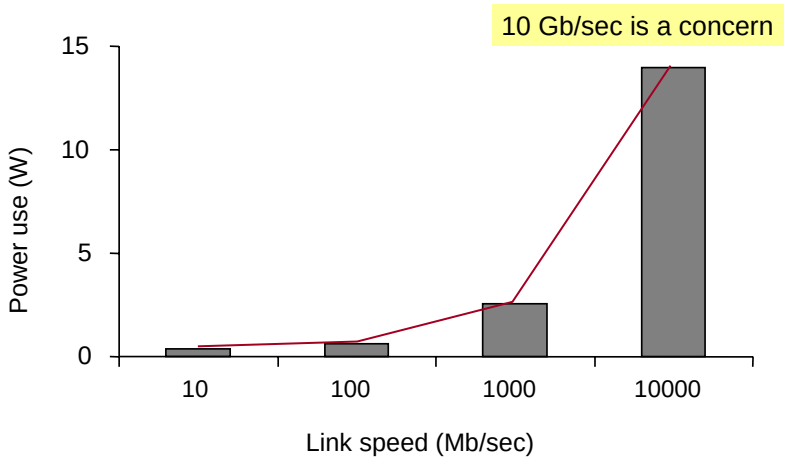
How much power use is direct from the network?

- **We study power consumption due to Ethernet links**
- **We measured…**
  - Cisco Catalyst 2970 LAN switch
  - Intel Pro 1000/MT NIC
- **We studied the specifications for…**
  - Intel 82547GI/82547EI Gigabit Ethernet Controller (NIC)
  - Chelsio N210 10GbE Server Adapter (NIC)

39

Power measurements continued

- **Averages of all measurements**



40

g00.xls

Adaptive link rate (ALR)

- **Automatic link speed switching***
  - For 82547GI/82547EI Gigabit Ethernet Controller

Typical PC NIC

- Automatic link speed switching from 1000Mb/s down to 10 or 100Mb/s in standby
- Low power in standby states
- Supports power-down states without software assistance

Drops link speed to 10 Mb/sec when PC enters low-power state

→ Motivates dropping link data rate if low utilization

41

* From Intel 82547GI/82547EI product information (82547gi.htm)

Adaptive link rate (ALR) continued

Independent of PC power management

Goal: Save energy by matching link data rate to utilization

- **Change (or adapt) data rate in response to utilization**
  - Use 10 or 100 Mb/sec during low utilization periods
  - Use 1 or 10 Gb/sec during high utilization periods
- **Need new *mechanism***
  - Current auto-negotiation is not suitable (too slow)
    - Designed for set-up (e.g., boot-up time), not routine use
- **Need *policies* for use of mechanism**
  - *Reactive policy* possible if can switch link rates "quickly"
  - *Predictive policy* is needed otherwise

ADSL2 has adaptive link rate

42

Adaptive link rate (ALR) continued

How much time is there for power management?

- **We collect and characterize traffic "in the wild"**
- **Traffic collection at University of South Florida (USF)**
  - Three traces from dormitory LAN (3000+ users) in mid-2004
    - USF #1 – The busiest user
    - USF #2 – 10th busiest user
    - USF #3 – Typical user
- **Traffic collection details**
  - All are 100 Mb/sec Ethernet links
  - USF traces are 30 minutes captured with Ethereal

43

Adaptive link rate (ALR) continued

- **Summary of the traces**

Utilization is low

Trace	Total busy time	Total idle time	Total low util time	Utilization at 100 Mb/sec
USF #1	75 s	1759 s	1415 s	4.11 %
USF #2	47	1771	1571	2.63
USF #3	0.55	1801	1799	0.03

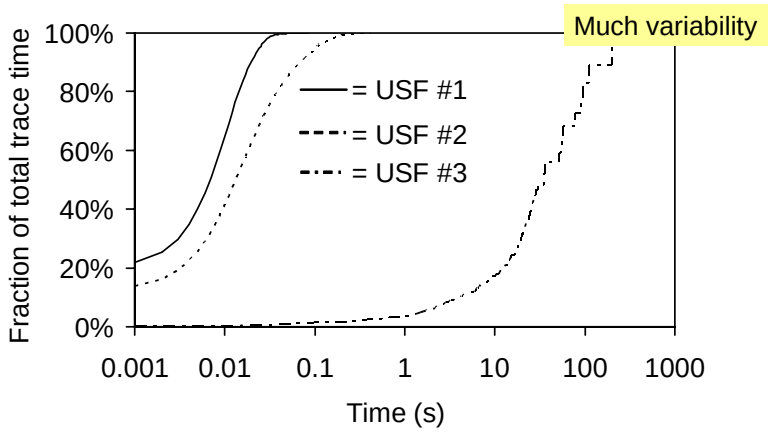
Large variability

Trace	Mean low util period	CoV of low util period
USF #1	0.0060 s	0.91
USF #2	0.0094	1.50
USF #3	1.0892	7.22

44

Adaptive link rate (ALR) continued

- **Fraction of low utilization periods for USF traffic**
  - For USF #1 and #2, most low utilization less than 100ms

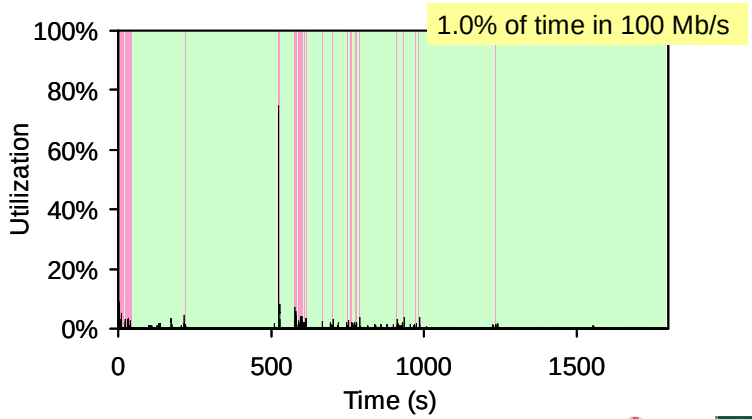


45

g04.xls

Adaptive link rate (ALR) continued

- **Utilization and link speed graphic**
  - Sample USF trace (USF #1)
    - From a simulation of ALR



46

g32.xls

Adaptive link rate (ALR) continued

- **Discussion of results…**
  - Great variation in length of low utilization periods
  - Can achieve energy savings and low delay for all traces
  - Expect that these results will hold for 1 Gb/sec
  - Need to consider energy cost of transition between rates

As with ADSL2, may be very important for MetroEthernet

47

Potential Energy Savings

- **Welcome to Part #5**

In this part… energy savings calculations for the SmartNIC and Ethernet Adaptive Link Rate.

48

Savings Estimates

- **All factors — stock, power levels, usage — not well known and changing**
- **Conclusions rely on magnitude of savings**
  - Not on precise figures
- **Assumptions**
  - 100 million commercial PCs } half desktops
  - 100 million residential PCs } half notebooks
  - Today's power levels
  - Usage patterns — rising # of PCs left on continuously

49

SmartNIC savings

- **First, consider one Continuous-on PC**
  - 40 hours/week in-use
  - 128 hours/week asleep (was fully-on before SmartNIC)
- **Unit Savings**

	Desktop / Notebook
Annual Electricity kWh/year	470 / 100
Annual Electricity $	$37 / $8
4-year lifetime $	**$150 / $32**

- **Stock-wide Savings**
  - Use unit savings for half of stock

➔ **28 TWh/year; $2.3 billion/year**

50

SmartNIC Savings continued

- **Stock-wide average savings**
  - Desktop: $75; Notebook: $16
  - "Budget" for retail cost of SmartNIC hardware
    - Except for notebooks — SmartNIC adds to functionality
- **If SmartNIC adds $5 to system cost, average payback time:**
  - Desktop: About 3 months
  - Notebook: 15 months

Highly cost-effective.

51

Adaptive Link Rate savings: Assumptions

- **"Success" rate: Should be nearly 100%**
  - As the stock of network equipment turns over
  - Does not rely on system sleep status
- **Average on- or asleep-time of whole stock almost 70%**
  - Take 80% of this as low-traffic time
    - ➔ 55% potential reduced data rate time
- **High data rate varies**
  - 1Gb/s - 80% of commercial; 20% of residential (50% average)
  - 100Mb/s - 10% commercial; 70% residential (40% average)

52

Adaptive Link Rate savings: Results

- **Per unit savings** (counts both ends of link)
  - 1Gb/s - 10 kWh/year $3.20 lifetime
  - 100 Mb/s - 3 kWh/year $0.96 lifetime
- **Cost-effectiveness**
  - Hardware cost should be minimal or zero; modest design cost
    - ➔ Very short payback times
- **Stock-wide savings**
  - 1.24 TWh/year

➔ **$100 million/year**

53

Summary and next steps

- **Welcome to Part #6**

In this part… we summarize the key points and discuss the next steps needed to energy savings.

54

IT equipment uses a lot of energy

- **All electronics about $16 billion/year of electricity**
- **PCs about $3.7 billion/year**
- **Many other products becoming PC-like**
- **… and both growing …**

55

Networks induce use of & directly use energy

- **Induce**
  - Many products must stay in a higher power state than otherwise needed to maintain connectivity
    - 802 networks (Ethernet, Wifi, etc.)
    - 1394 networks (some)
    - USB (some implementations)
    - TV set-top boxes (many)
    - and more…
  - Network applications increase on-times
- **Directly use**
  - Network interfaces and network products
  - Combined about $1 billion/year
- **… and both growing …**

56

Networks directly use energy

- **Network interfaces and network products**
- **Combined about $1 billion/year**
- **… and growing …**

57

Large savings potential

- **SmartNIC**
  - Now: $2.2 billion/year
  - Future savings growing
    - More PCs
    - More non-PC products with network connections
    - Longer on-times
    - Growing difference between On and Sleep power
  - Savings highly cost-effective
- **Adaptive Link Rate**
  - Now: $100 million/year
  - Future savings growing
    - More products with network interfaces
    - Higher speeds lead to (much) greater base power level

58

Implementation

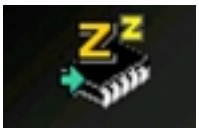
- **Define specification**
  - SmartNIC — IETF (Internet Engineering Task Force)
  - Adaptive Link Rate: IEEE 802.3

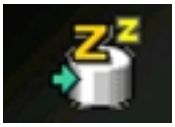
But won't do without external prompting / support

- **Implementation**
  - ENERGY STAR could mandate use once available
  - For industry, less expensive than other options for large energy savings

59

Questions / Comments





Bruce Nordman
Energy Analysis
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
bnordman@lbl.gov

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

60