

# The next frontier for communications networks: power management

Kenneth J. Christensen<sup>a,\*</sup>, Chamara Gunaratne<sup>a,1</sup>, Bruce Nordman<sup>b,2</sup>, Alan D. George<sup>c,3</sup>

<sup>a</sup>Department of Computer Science and Engineering, University of South Florida, 4202 East Fowler Avenue, ENB 118, Tampa, FL 33620, USA

<sup>b</sup>Energy Analysis Department, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory,  
1 Cyclotron Road, 90-4000, Berkeley, CA 94720, USA

<sup>c</sup>Department of Electrical and Computer Engineering, University of Florida, P.O. Box 116200, 327 Larsen Hall, Gainesville, FL 32611-6200, USA

Received 25 June 2004; accepted 25 June 2004

Available online 9 August 2004

## Abstract

Storage, memory, processor, and communications bandwidth are all relatively plentiful and inexpensive. However, a growing expense in the operation of computer networks is electricity usage. Estimates place devices connected to the Internet as consuming about 2%, and growing, of the total electricity produced in the USA—much of this power consumption is unnecessary. Power management is needed to reduce this large and growing energy consumption of the Internet. We see power management as the ‘next frontier’ in research in computer networks. In this paper, we propose methods for reducing energy consumption of networked desktop computers. Using traffic characterization of university dormitory computers, we show that there is significant idle time that can be exploited for power management. However, current Ethernet adapters in desktop computers lack the capabilities needed to allow existing system power management features to be enabled. We address this problem with a proxying Ethernet adapter that handles routine network tasks for a desktop computer when it is in a low-power sleep mode. This proxying adapter can allow existing power management features in desktop computers to remain enabled and have the computer be ‘on the network’ at all times. The energy that we expect can be saved is in the range of 0.8–2.7 billion US dollars/year.

© 2004 Elsevier B.V. All rights reserved.

**Keywords:** Power management; Traffic characterization; Desktop computers; Proxying Ethernet adapter

## 1. Introduction

The amount of electricity consumed by devices connected to the Internet is rapidly increasing. Estimates for the US electricity consumption by the Internet vary widely from less than 1% [15] to a range between 2 and 8% [19]. The lower-end estimate includes the network infrastructure but does not include the devices that attach to the Internet (e.g. servers and desktop computers); the higher estimates include both. Significant growth in this electricity

consumption is expected as personal digital devices and embedded devices proliferate and connect to the Internet. Some extreme projections suggest that 50% of all electricity will be consumed by the Internet and 90 new power plants per year will be needed to satisfy power consumption needs of Internet equipment [14]. However, such extreme projections have been shown to be based on incorrect assumptions [20]. A validated end-of-1999 estimate [19] from Lawrence Berkeley National Laboratory (LBNL) states:

We found that total direct power use by office and network equipment is about 74 TWh/year, which is about 2% of total electricity use in the US. When electricity used by telecommunications equipment and electronics manufacturing is included, that figure rises to 3% of all electricity use (Koomey, 2000). More than 70% of the 74 TWh/year is dedicated to office equipment for commercial use. We also found that power management

\* Corresponding author. Tel.: +1-813-974-4761; fax: +1-813-974-5456.

E-mail addresses: [christen@csee.usf.edu](mailto:christen@csee.usf.edu) (K.J. Christensen), [pgunarat@csee.usf.edu](mailto:pgunarat@csee.usf.edu) (C. Gunaratne), [bnordman@lbl.gov](mailto:bnordman@lbl.gov) (B. Nordman), [george@hcs.ufl.edu](mailto:george@hcs.ufl.edu) (A.D. George).

<sup>1</sup> Tel.: +1-813-974-4761; fax: +1-813-974-5456.

<sup>2</sup> Tel.: +1-510-486-7089.

<sup>3</sup> Tel.: +1-352-392-5225.

currently saves 23 TWh/year, and complete saturation and proper functioning of power management would achieve additional savings of 17 TWh/year. Furthermore, complete saturation of night shut down for equipment not required to operate at night would reduce power use by an additional 7 TWh/year.

Since 1999, the ‘Internet’ energy consumption has increased for three reasons. They are: (1) more absolute number of devices; (2) higher active power of these devices; and (3) more hours per day usage. The larger number of devices since 1999 is due in part to the increased types of devices (including devices for mobile Internet access and embedded systems for direct machine-to-machine communications in the Internet). The large absolute value of electricity consumed supports the need for improving the energy efficiency of the Internet. If energy savings can be achieved, then energy efficiency of the Internet should be pursued with the potential for great cost savings and environmental benefits. Most of the LBNL expected savings would come from desktop computers, printers, and other networked devices. In a modern workplace, desktop computers and other electronic devices are all networked together and connected to the Internet.

A major factor that prevents the proper functioning of power management is the need to maintain network connectivity at all times. Many workers need to access the data and resources on their office desktop and server computers while physically being elsewhere and outside normal working hours. Office and home desktop computers are often left powered-on during nights and other idle periods (both short and long in duration), so they can be accessed for file sharing and maintenance, keep their IP address (if DHCP administered), and so on. Studies conducted by LBNL in 2000 showed that existing power management features are disabled in a majority of office computers [35]:

Of the 1280 computers audited (desktop and desksides only; does not include servers or portables) 44% were off, 3% were in low-power mode, and the remaining 54% were on (these numbers do not total 100% because of rounding). As noted above, there is not always a clear indication that a computer is in low-power mode. The reader should therefore use caution in drawing conclusions about power management use in computers based on these results although we can conclude that at least 5% of the computers left on were in power management mode.

The actual percentage of machines effectively power managing was not likely much higher than the 5% cited above, and in a more recent study [28] in which low-power status was directly measured, the results were similar. We believe that a solution to the problem of disabling of existing power management features lies in developing

a proxying Ethernet adapter that can wake-up a desktop computer in sleep mode when the resources of the computer are needed and otherwise handle routine protocol messages without waking-up the computer. Thus, the desktop computer can maintain a presence on the network without being fully powered-on at all times. Characterization of the traffic generated by networked desktop computers is the first step in addressing this problem. Traffic can be studied from inter-flow, intra-flow, and protocol (semantic) perspectives. The goal is to derive insights that can be used to develop better power management methods and to identify the capabilities needed for a proxying Ethernet adapter. In this paper, new directions in traffic characterization for power management and the development of a proxying Ethernet adapter are described. This paper builds on previous work presented in Ref. [8].

The remainder of this paper is organized as follows. In Section 2, the power consumption of the Internet and of desktop computers is described. Section 3 focuses on power management of desktop computers and also describes the Environmental Protection Agency (EPA) Energy Star program. In Section 4, we describe a traffic characterization study done at the University of South Florida. Section 5 describes a simple time-out based power management scheme and applies it to the collected traffic data. In Section 6, we detail the design and implementation of a proxying Ethernet adapter. Section 7 summarizes the paper and describes future work.

## 2. Power consumption of devices connected to the Internet

Quantifying and finding ways to reduce the energy consumption of the Internet is gaining interest in the research community. In Ref. [15], Gupta and Singh calculate that the physical networking infrastructure of the Internet (i.e. hubs, switches, and routers) consumed 6.05 TWh in 2000 and it is expected to increase by another 1 TWh by 2005. Several ideas are proposed for reducing this energy consumption including putting network interfaces and switch components into sleep modes under low-load conditions and modifying routing algorithms to minimize the number of routes between source–destination pairs under low-load conditions. Energy-aware routing would aggregate traffic to fewer routes and allow interfaces and equipment on unused routes to be put into low-power sleep modes.

In wireless and ad-hoc networking, minimizing energy consumption has been a long-standing research problem. In these networks, battery resources are the primary constraint to useful deployment. Of particular interest are routing protocols where nodes minimize unnecessary transmission. In wireless networks, transmitting packets consume significant energy due to the non-directional nature of wireless transmission. A survey of energy-efficient wireless network protocols is presented in Ref. [18].

Data centers and server clusters represent a highly visible ‘hot spot’ with respect to energy consumption. In a server farm, a large number of high-performance server computers are colocated in one facility. Electricity is needed for both powering the server computers and for cooling the data center facility. Power consumption of individual servers [22] and of clusters of servers [5–7] is being addressed. Methods to power-down servers in a cluster as a function of arriving workload are being studied in Refs. [5,7]. In Ref. [25], the energy consumption patterns of data centers are studied and it is shown that the growth in electricity consumption has not dramatically increased where data centers are located when compared to the overall growth in electricity power consumption. The conclusion is, to some degree, that data centers are relocating energy consumption from office building to server farm data centers and not adding to the overall US energy consumption [25].

Power consumption of desktop computers greatly exceeds that of the much fewer server computers and is also increasing. Newer LCD monitors require less power than older CRT monitors, but system units with their faster CPUs and graphics adapters are increasing in power consumption. For example, a 1999 Dell GX200 933-MHz P3 computer with 256 MB RAM with a 21-in. CRT monitor measured a constant 36 W for the system unit and 96 W for the monitor. A 2002 Dell 2350 1.8-GHz P4 computer with 256 MB RAM with a 19-in. LCD monitor used a varying 60–85 W for the system unit and 27 W for the monitor. The Dell 2350 computer drew 6 W when powered off and 7 W when in ‘standby’ mode (sleep) with Microsoft Windows XP. If a computer and monitor drawing 100 W is added to the household and is ‘always-on’ (i.e. powered-on 24 h/day for 365 days/year), it will result in an increase of about 8.5% (!) in the typical household power bill (based on an average household consumption of 10,219 kWh/year [21]). This is a significant and measurable increase in power consumption for a household and thus also power generation demands for the country. Despite significant research in power management at the chip and operating system levels, the trend is towards continued growth in overall system power consumption. Even with some level of reduction or slowing of growth, power consumption of a fully powered-on desktop computer will never be as low as that of a desktop computer in a lower power state.

### 3. Power management and Energy Star for desktop computers

In this section, we review the basic tenets of power management including time scales of interest, describe power management for desktop computers, and present the issue of maintaining network connectivity at all times.

#### 3.1. Time scales of power management

The typical home or office desktop computer is not in active use 24 h/day. There is significant scope for reducing the energy consumed by the computer by using power management to shutdown the system or its components when they are idle (not in use). Idle periods can be defined over multiple time scales. For each time scale, different power management methods are possible. With respect to network connectivity the time scales are:

- CPU and instruction level (nano to microseconds)—circuits within a CPU can be turned-off
- Inter-packet or intra-flow (micro to milliseconds)—the CPU can be halted or turned-off
- Inter-flow (seconds to hours)—the entire computer system can be turned-off

A flow can be a TCP connection bounded by connection establishment (SYN) and termination (FIN or RST). Each of these time scales requires different approaches to implement power management. In this paper, we focus on the largest time scale (inter-flow). From a network view, inter-flow idle times are a conservative estimate of the available idle time for power management. Even within flows there will be idle periods between packet transmissions that can be used for power management. Methods have been explored to bunch packet transmissions to increase idle time [3]. With respect to network activity, power management is achieved by:

- Predicting, controlling, and making the best use of idle times
- Increasing the predictability of existing idle times
- Creating additional idle time by bunching and/or eliminating traffic
- Having a smaller processor substitute for a larger processor whenever possible

Power management is already in the forefront for the research community in the areas of mobile systems and server clusters. Mobile systems—including emerging sensor networks—are constrained in their capabilities by a lack of sufficient power from batteries. Power management research is ongoing in many different areas including:

- Chip level; where circuits within a chip are powered-off when not used [24,31]
- Protocol level; where new routing protocols minimize the time a wireless transmitter must be on [30]
- System level; where components or subsystems are powered-down when not in use [4]

#### 3.2. Energy Star and other power management standards

In many cases, desktop computers already have energy savings features such as EPA Energy Star [12] compliant

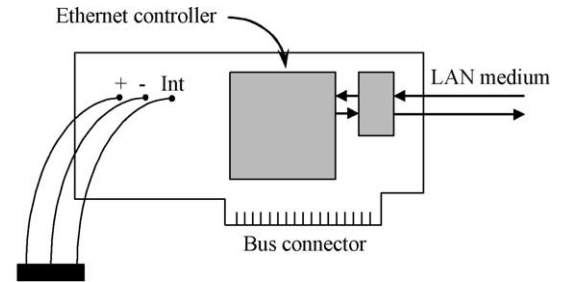
power management features, principally low-power sleep states; however, these features are often disabled. Energy Star is a US government labeling program that promotes energy conscious purchasing by consumers and encourages manufacturers to build more energy efficient products. Energy Star is administered by the US EPA and the US Department of Energy (DOE), and establishes energy consumption specifications for equipment purchased by the US government. Since the US government is one of the largest purchasers of equipment, the specifications often become de facto industry standards (for office equipment). The EPA has established specifications for many types of equipment including office equipment. In the case of computers, after an inactivity period (e.g. 30 min), the computer must enter a low-power sleep state with significantly reduced power consumption. If the computer is network capable, it is expected to respond to any ‘wake events’ conveyed via the network while in sleep mode. Wake events are defined as: "A user, programmed, or external event or stimulus that causes the computer to transition from its low-power or sleep mode to its active mode of operation" [13]. The criteria also states:

A well-designed ENERGY STAR qualified computer will not lose its network connection, which could lead to a loss of data, while in the low-power or sleep mode. Most of the computers that are being manufactured today include a capability to sleep intelligently on a network.

There are a variety of reasons that lead to power management being disabled in desktop computers, but two that arise from network connectivity are a large and increasing portion of the total. The first is that some applications automatically prevent entry into a sleep mode when accessing resources (e.g. having a file open in a shared computer) across a network. For example, if a file on a remote computer is opened using Microsoft Word for Windows XP then neither the local or remote computer will enter into a sleep state. The second reason is that many network administrators and users manually disable power management in order to insure access to their computers at all times.

Several industry-based groups have defined power management and/or network connectivity related standards. Wake-on-LAN (WOL) (also called Magic Packet) [23] is a mid-1990s industry standard that makes it possible for an Ethernet adapter to wake-up a sleeping desktop computer using a specially defined WOL packet. The WOL packet is a standard Ethernet packet with the MAC address (i.e. the Ethernet or ‘hardware’ address) of the target computer repeated 16 times in the data field. Fig. 1 shows a WOL Ethernet adapter, which has three functions:

1. Connection to auxiliary power on the system motherboard to allow adapter operation when the system bus



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

Fig. 1. Ethernet adapter with auxiliary power and wake-up capability.

and processor powers off (i.e. when the computers goes to sleep).

2. Connection to a wake-up interrupt line on the system motherboard that is external to the system bus.
3. The ability to receive all packets when operating on auxiliary power and recognize a special WOL packet. On recognition of a WOL packet, a wake-up interrupt signal is generated that wakes up the sleeping desktop computer.

A computer that needs to use the resources of a currently sleeping computer sends a WOL packet. A WOL packet needs to contain the MAC address of the destination computer (i.e. the computer to be woken-up). Knowledge of MAC addresses beyond a subnet is virtually impossible. Thus, sending a WOL packet across multiple subnets—that is, routing a WOL packet—is not feasible. The WOL packet is also not part of the existing standard TCP connection establishment packet sequence. WOL has not achieved significant use beyond specialized systems management software. Another key issue is that devices do not share power state information and thus do not ‘know’ when to initiate wake-up sequences.

A current industry standard for power management in desktop computers is the Advanced Configuration and Power Interface Specification (ACPI) [2]. ACPI defines a set of common interfaces through which applications, device drivers, the BIOS, and the operating system can exchange data regarding configuration and power status. It also defines sets of standardized states for system devices and transitions between these states. Five separate low-power modes, S1 through S5, are defined. S1 defines a low-latency sleep state while S5 defines a state that requires a complete boot to recover. ACPI also extends WOL by specifying ‘wake events’ based upon matching specific bit patterns for received packets. By defining a bit pattern that is the computer’s IP address it is possible to enable an Ethernet adapter to detect a data packet that is directed at that computer at the IP level. In addition to supporting WOL, Intel Pro 1000 series Ethernet adapters [17] are capable of waking-up the computer on receiving a packet with the IP address of the system or an ARP packet targeted for the system—this is a direct packet wake-up. Direct packet

wake-up may have the shortcoming of triggering unneeded wake-ups and thus preventing a system from sleeping when it should be. Some initial measurements have shown that this may be the case.

### 3.3. Maintaining network presence during power managed periods

A power-managed desktop computer that is also connected to a network suffers from an inherent contradiction. To save energy, when the computer is idle it must transit to a low-power sleep mode. However, when in sleep mode, the computer's network presence is lost and the data and resources within the computer are inaccessible from other nodes on the network. We have found that when power management is enabled in desktop computers:

- When certain applications running in the computer are accessing data on a remote shared directory, the computer would refuse to enter low-power mode even when idle for the required period of time.
- If the computer transitions into a low-power mode while remote-access applications are executing, after a time-out period, the loss of TCP connections occur. When the computer wakes-up, these lost TCP connections are not automatically restored and any ongoing sessions are lost. For example, telnet and SSH protocols exhibit this behavior.
- With desktop computers using Microsoft Windows XP, if a computer is in sleep mode when another system is powered up, the former system will not appear in the latter's Network Neighborhood.
- When a computer that is visible in the Microsoft Windows Network Neighborhood of other computers transitions into sleep mode, the shared drives and other resources on that computer will be inaccessible.

Direct packet wake-up (e.g. on an ARP) makes it possible for a computer to wake-up on standard protocol packets such as an ARP or TCP SYN. However, direct packet wake-up will wake-up a computer when the required response might not need the resources of a fully active computer. For example, responding to an incoming ICMP ping packet does not require any data or resources in the desktop computer. We address this problem with the design of a proxying adapter in Section 6.

### 3.4. Energy savings

The range of plausible estimates for the energy that could be saved by fixing the network problem is wide. Some factors are not well known and others are hard to predict for the coming years. Nevertheless, we believe that the following estimate is a reasonable initial estimate—a snapshot for the year 2007—for savings from PCs. Thus, this is a 'technical potential' to indicate the ultimate

magnitude of the savings when the entire stock of PCs has turned over, not a forecast for any particular year.

Hardware and software, such as described in this paper, for network-friendly desktop computers could plausibly be introduced widely into the market beginning in 2006, and be most of the stock of these machines several years after that. Estimating the stock, power levels, and usage patterns of future desktop computers is necessarily speculative, but reasonable assumptions can be made, drawing on data for those parameters for recent years, and known trends in equipment and use. We based our estimate on figures that underlie savings estimates for the Energy Star program. For simplicity and to be conservative, several sources of savings were not included: savings from notebook computers; additional savings from more power management of displays; the likelihood that higher-power PCs will be on more than lower-power ones; and savings from PCs that are on less than 40 h/week. Key parameters are the power savings from shifting from on to sleep mode (55 W) and the total stock of residential and commercial PCs (about 220 million by 2007). For usage we consider a low and high estimate, based on the fraction of machines that are on most or all of the time, and the fraction of those that could be shifted from on to sleep when not actively being used. The two estimates should bound the plausible range of savings that are attainable from fixing the network problem. The savings estimates are about 11 and 36 TWh/year, which at current prices cost about 0.8–2.7 billion US dollars/year. Because the estimate is conservative on a number of accounts, we expect the actual figure to be toward the upper end of this range. In Appendix A, we show how the savings estimates for implementing solutions to the network problem were calculated.

Companies and government programs are beginning to develop products and programs to capture some of the possible savings. The EPA power management program seeks to enable monitor power management on desktop computers [27]. Based on an estimate of 55 million office computers in the US, EPA estimates that about 900 million US dollars/year could be saved through this program. Companies, including Verdiem [32] in the US and 1E [1] in the UK, are beginning to develop products to address power savings in desktop computers. Verdiem [32] was founded in 2001 and has a software product to globally control power management settings in the desktop computers of large organizations. Verdiem's Surveyor product can block individual users from disabling power management in their Microsoft Windows PCs. It also allows a multiplicity of power schemes to be automatically implemented at different times of day. Thus, it becomes possible to have minimal power management during the day (i.e. to cause little interference to users) and more aggressive power management at night when users are not at their computers. With global control—using the WOL capabilities in Ethernet adapters of most modern desktop computers—it is possible to initiate system-wide wake-ups for

management (e.g. virus updates and data back-ups). The Surveyor product also has the ability to estimate energy savings by monitoring the time spent in various power states of the computers that it controls. The Surveyor product cannot directly address applications that require constant TCP connections or other higher-level visibility on the network, but can scan for the presence of these applications on a Surveyor-enabled PC and veto any attempt to drive the machine into a low-power state if such an application is running.

#### 4. Traffic characterization study

To study the potential for energy savings from power-managed networked desktop computers, we characterized the busy and idle periods of computers in the University of South Florida (USF) undergraduate dormitory. Dormitory computers are known to be running cutting-edge file sharing applications and contribute the majority of the traffic by volume (60%) to the USF Internet connection. Thus, these computers are of interest to study. It is during idle periods that a computer could power-down and save energy. Of interest is how much idle time there is and what are the characteristics of idle periods. We collected 1 day of Cisco NetFlow trace data from the top 100 (by traffic volume—bytes sent and received) of about 3000 computers on March 27, 2003. In this 24-h interval, the dormitory computers sent and received 1350 GB. Fig. 2 shows the network configuration with Cisco NetFlow traffic data collection at the egress router to the Internet. The USF dormitories are throttled to 64 Mbps by a traffic shaper and only external traffic flows could be collected (the quantity of internal flows is unknown, but not expected to affect the results). A Cisco NetFlow record defines the start and end time of a flow and contains a count of the number of packets and bytes carried. Fig. 3 shows the definition of a *busy period* and an interflow *idle period*. A busy period is comprised of one, or more, overlapped flows. An interflow idle period is the time during which the computer has no active flows. Fig. 4 shows the ranked volume of data for the top 100 computers in

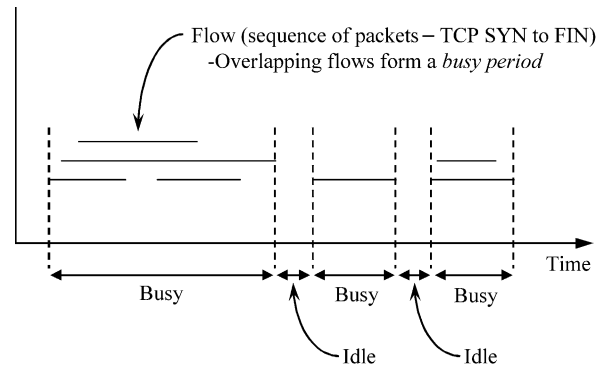


Fig. 3. Definition of idle and busy periods.

the measured 24-h interval. A total of 459 GB were sent and 132 GB received in the collected flows. One computer alone is responsible for 12% of this total volume. About 99% of the overall traffic volume was in TCP flows with UDP about 1% and ICMP negligible. Fig. 5 shows the flow size where it can be seen that most flows are very small in size, but a few flows are very large. Plotting a ranking of flow sizes results in an approximate power law fit as shown in Fig. 6. It is interesting that power law appears to apply to idle times much as it has been show to apply to many other measures of network traffic.

The total interflow idle time of a computer is the sum of its idle periods during the measurement interval. Fig. 7 shows the total idle time per desktop computer ordered by rank (from Fig. 4). As expected, the computers with lower traffic volume have the greatest idle time. Some computers have no idle time at all—they had active flows during the entire measurement period. Fig. 8 shows the distribution of idle periods. The mean idle period is 15 s, the standard deviation is 2.1 min, and the longest idle period is 4.7 h. Fig. 9 plots the accumulated idle time as a histogram of the idle time durations. That is, the value for the  $X$ th second gives the accumulated total value of idle times that are greater than or equal to  $X$  seconds but are less than  $(X+1)$  seconds. Fig. 9 shows that the majority of the total idle time is concentrated for both short and long idle period durations.

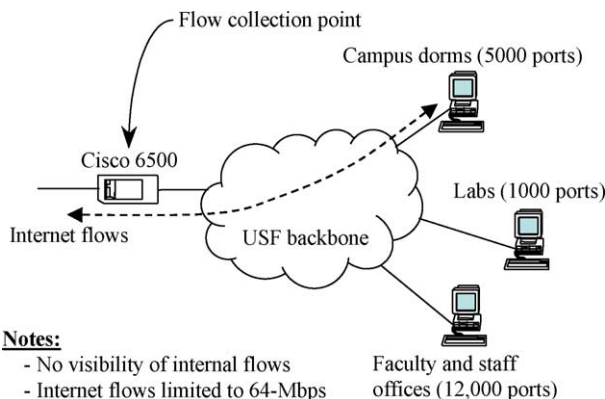


Fig. 2. USF dormitory network configuration.

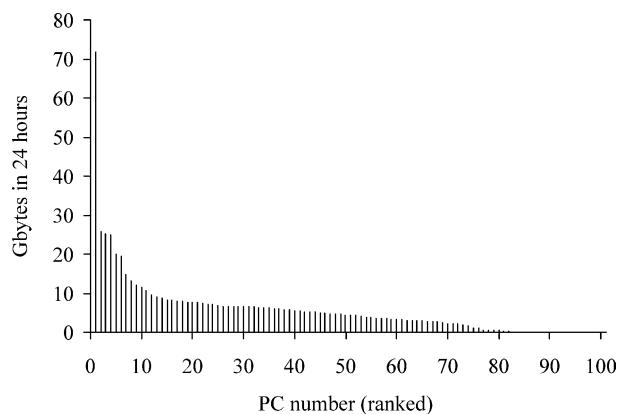


Fig. 4. Volume of data sent and received for top 100 USF dormitory computers.

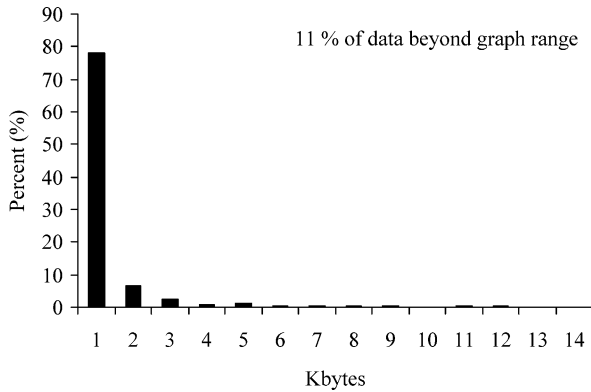


Fig. 5. Flow size for top 100 USF dormitory computers.

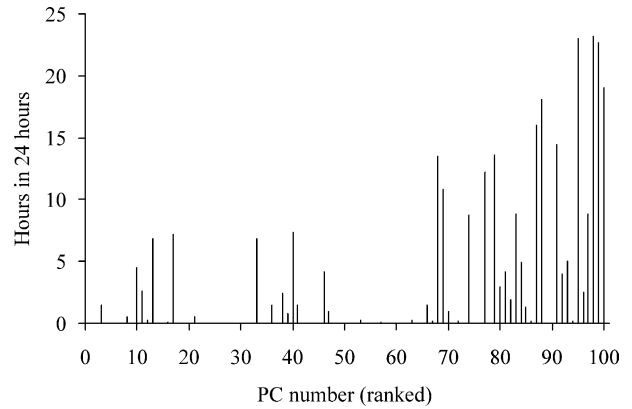


Fig. 7. Total idle time.

The cause of this short-long characteristic needs further investigation. The autocorrelation between idle periods (also between idle and busy periods) is very small. Fig. 10 shows the autocorrelation of idle periods up to a lag of 80. This low autocorrelation makes predictive schemes for power management infeasible.

Idle periods also exist within flows—intraflow idle periods—in the time between packets. To measure and characterize intraflow idle periods the USF dormitory top 100 NetFlow traces were studied. The total number of bytes sent and received was divided by the link speed (100-Mbps) to get the total transmit time. From this we can determine the idle time within flows. The breakdown for idle periods for the top 100 USF dormitory computers was found to be, intraflow idle times were 86% (of the total idle time), interflow idle times were 13%, and actual packet transmission time was 1%. The mean flow time is approximately 103 s and of this time, less than 1 s is spent transmitting and receiving data. The rest of the time is idle periods between packets. These measurements suggest that there is opportunity for small time-scale power management or for methods to reduce the length of flows.

We are also interested in traffic at a semantic level. For the proxying adapter described in Section 6, it is important to know what protocol messages are received by a typical

desktop computer. Some messages can be ignored (e.g. spanning tree protocol packets), others may require a trivial response (e.g. an ICMP ping), and still others may require the full resources of the computer. A packet trace program was installed on an idle Windows 2000 desktop computer in the Department of Computer Science and Engineering at USF. The computer was on a dedicated switched-Ethernet link, so only broadcast and direct-addressed packets were sent on the link. No applications were running at that time in the computer under test and it was configured to be a member of a Windows workgroup. Thus, for example, it would not be expected to see any HTTP packets (as they do not use broadcast). The single HTTP packet that was seen could be from a malicious port scan. In 30 min over 6000 packets were captured. Table 1 gives a breakdown of the packets received. These packets can be classified into five categories:

1. Received packets that can be ignored (i.e. the packets are intended for other computers—the majority)
2. Received packets that require a simple response to be transmitted (e.g. ICMP, ARP)
3. Received packets that require a simple response and a state update (e.g. NetBIOS)
4. Received packets that require an application response (e.g. TCP)

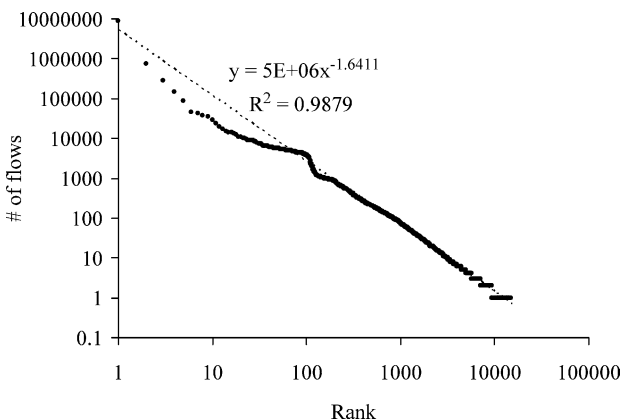


Fig. 6. Flow size rank.

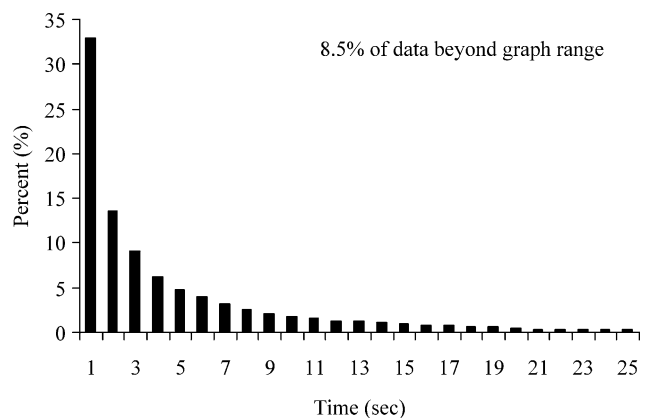


Fig. 8. Distribution of idle periods.

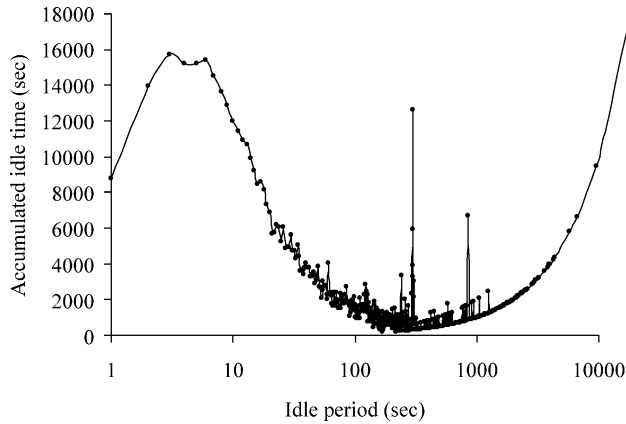


Fig. 9. Accumulated time plot for idle periods.

5. Transmitted packets originated by a protocol or application and not in response to a received packet (e.g. DHCP).

**5. A time-out policy for power management**

Time-out based power management schemes are used for powering off inactive systems and subsystems (such as disk drives in laptop computers). In a time-out based scheme, the system powers down at the expiration of an inactivity timer, which is started when the system is detected to be idle (i.e. there are no active flows). The system powers up on detecting activity that requires the resources of the sleeping system. The time for a system to go to sleep and wake up are non-zero and must be factored into the response time performance. The wake-up time affects the response time of the request that triggers the wake-up. In addition, a wake-up may cause a power consumption spike so the frequency of wake-ups and their effect on energy consumption must be considered.

We used the collected USF dormitory data as input to a simulated time-out based power management scheme. Fig. 11 shows the finite-state machine for a time-out

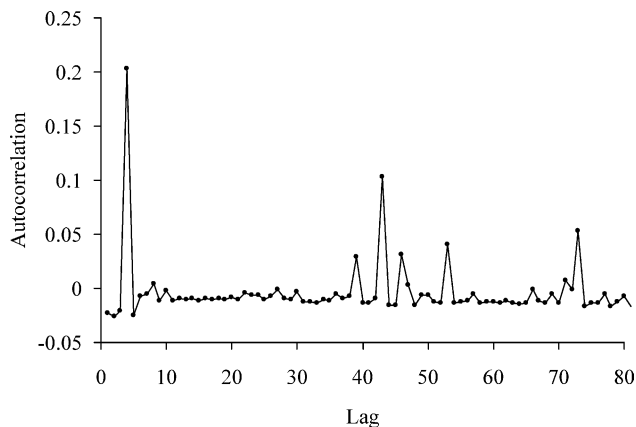
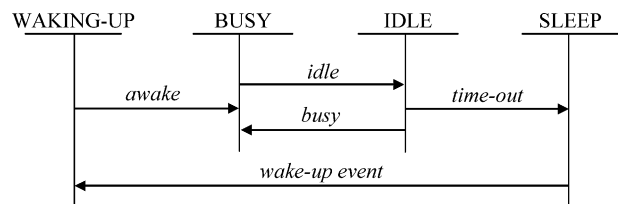


Fig. 10. Autocorrelation of idle periods.

Table 1  
Breakdown of protocol messages for a single observed computer

Packet type	# in trace	% in trace
NetBIOS over IPX	40	0.60
NetBEUI	30	0.45
RARP	79	1.19
SNAP	835	12.55
ARP	2143	32.20
BRIDGE	882	13.25
DEC MOP	3	0.05
DHCP	433	6.51
HTTP	1	0.02
SNMP	1	0.02
NetBIOS Datagram	568	8.53
NetBIOS Name service	483	7.26
NetBIOS Session service	17	0.26
DNS	166	2.49
TCP port 5353	14	0.21
OSPF Hello	177	2.66
RIP	64	0.96
NTP	27	0.41
SSH	2	0.03
SAP	302	4.54
X Display	18	0.27
Sun remote procedure calls	4	0.06
Remote cache service	69	1.04
IGMP	150	2.25
ICMP	3	0.05
ESRO	4	0.06
Groove	88	1.32
Unknown	53	0.80
Total	6656	100.00

power management scheme. In this scheme, a desktop computer goes to sleep after a time-out period based on the expiration of an inactivity timer. Starting from a BUSY state, the computer transitions to the IDLE state at the start of an idle period (and resets/starts the inactivity timer). The idle event transition occurs when there are no active flows. The computer then transitions to a low-power SLEEP state when the inactivity timer expires. It takes non-zero time to achieve low-power operation (i.e. in the diagram of Fig. 11, the SLEEP state includes the time to go from full power to low-power consumption). A wake-up event, which can occur at any time after entering the SLEEP state, transitions the computer to the WAKING-UP state where the computer



**Notes:**

- In IDLE state an inactivity timer is started
- A wake-up event can be a connection request or other activity
- The time needed to wake-up is a performance penalty
- In SLEEP state achieving low power does not occur instantly

Fig. 11. Finite state machine for power management.

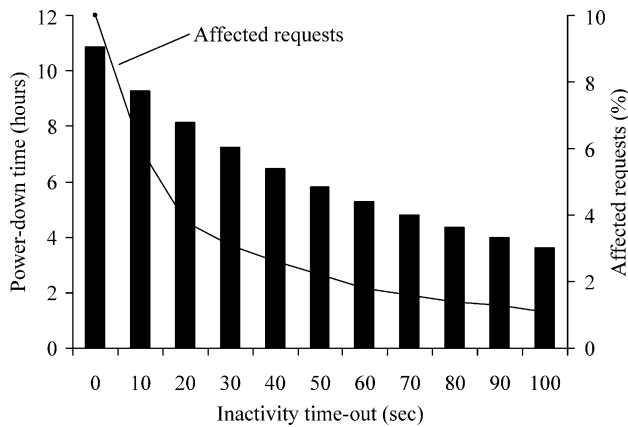


Fig. 12. Inactivity time versus power managed idle time and affected requests for computer #68.

resumes full-power operation. When full-power operation is resumed, the *awake* event transitions the computer back to the BUSY state. The time spent in the WAKING-UP state is the response time performance penalty that a request to trigger a system wake-up will experience. The time spent going from full to low-power operation (i.e. in the SLEEP state) is irrelevant from a response time point of view, which is why it is unnecessary to represent it as a separate state here. The simulated time-out based power management scheme was executed for computer #68 of the top 100 USF dormitory computers described in Section 4. This computer had 11 h of idle time in the 24-h measurement interval. Fig. 12 shows the results of this experiment on a double-axis graph of total power-down time in a 24-h interval (the bars) and the percentage of affected requests (the line) as a function of inactivity time-out. As the inactivity time increases, the total idle time and the percentage of affected requests both decline. This is, there is a trade-off in the amount of possible power management time versus performance penalty to incoming requests. For an inactivity time value of 0 s (i.e. the full idle time of the computer is used for a low-power sleep mode), the percentage of affected requests is only 10% for computer #68. The 10% result occurs because an average busy period contains 10 overlapping flows where only the first flow will be affected (i.e. have a delayed response due wake-up time). These results show that there is potential for energy savings where only a very small percentage of requests are affected; the impact to response time performance can be measured.

## 6. A proxying Ethernet adapter

In this section, we describe the initial design and development efforts for a proxying Ethernet adapter. The proxying adapter adds a small, low-cost microcontroller or processor to an Ethernet adapter. Similar to the WOL adapter of Fig. 1, the proxying adapter can operate on

1. receive a packet
2. if (system is awake) then
3.   pass packet to computer
4. else
5.   if (packet requires no action) then
6.     discard packet
7.   else
8.     if ("simple" request) then
9.       handle locally on the adapter
10.     discard packet
11.   else
12.     trigger computer wake-up
13.     wait for computer to wake-up
14.     pass packet to computer

Fig. 13. Control flow for a received packet in a proxying Ethernet adapter.

auxiliary power and can trigger a wake-up interrupt to the desktop computer. This onboard processor does several things:

1. Proxies for the desktop computer for simple routine protocol events
2. Wakes up the desktop computer when the resources are needed
3. Possibly exchanges state information with the operating system of the computer

Fig. 13 shows the control flow for a received packet. If the computer is awake, the adapter operates as normally (i.e. packets are passed to the computer). If the computer is sleeping, packets are either discarded, handled locally on the adapter, or trigger a wake-up interrupt to the computer. We believe that ARP, ICMP, DHCP, and some NetBIOS functions can easily be proxied on an Ethernet adapter. This corresponds to approximately 2% of the traffic received by one computer (from Table 1). TCP connection requests to listening applications on the computer would trigger a system wake-up. Other protocols such as SNMP GETs may also result in a system wake-up. We believe that ACPI defines the necessary interfaces to exchange state information between an adapter and computer. For example, the adapter needs to know all of the IP addresses used by the computer in order to be able to respond to ARP packets and perform direct wake-ups.

Offloading the entire TCP/IP protocol stack to an Ethernet adapter may become feasible in the near future. However, such an offloaded protocol stack cannot replace the protocol software in the system. The CPU on an Ethernet adapter would be (by cost consideration alone) significantly less powerful than the main system CPU. Thus, raw throughput performance for bulk data transfers of a TCP/IP implementation running on an adapter would be

very poor (compared to the throughput performance of keeping the TCP/IP implementation running on the system CPU). In addition, such a protocol offload would require non-trivial changes to existing operating systems. Off-loading the protocol stack during sleep states could facilitate less frequent wake events. This would, however, require more advanced hardware and software. Thus, the challenge is to determine which functions can be transparently (or nearly transparently) off-loaded to the adapter with its small onboard processor. We plan to implement the proxying adapter in four phases:

- *Phase #1.* Use of a second computer to emulate the proxying Ethernet adapter. This emulation method will allow us to quickly and easily investigate proxying and wake-up.
- *Phase #2.* Use of an internal Ethernet controller (such as a commercial Ethernet embedded controller [10]) to implement proxy functions. This intermediate step will be the basis of a convincing demonstration.
- *Phase #3.* An implementation of a proxying Ethernet adapter using an on-board microcontroller.
- *Phase #4.* An implementation of a proxying Ethernet adapter using multi-level hardware and software to achieve greater energy efficiency of the adapter itself.

A conservative estimate of the number of Ethernet adapters in the USA is 100 million. If each adapter consumes 2 W and we assume that the adapters are powered-on half the time, then the total energy use is 0.9 TW/h per year. This large energy use motivates phase #4, which is to investigate more energy-efficient designs and implementations of Ethernet adapters.

Fig. 14 shows a photograph of our phase #1 proxying adapter test bed with a Test System and Proxy Emulator. The Test System is a power-managed Dell desktop computer with Microsoft Windows XP with an Ethernet adapter that supports WOL. The Proxy Emulator is another

Dell computer, and it emulates the proxy adapter. The Test System and Proxy Emulator are connected via a private 100-Mbps Ethernet link. The Proxy Emulator is connected to the laboratory's 1-Gbps Ethernet switch and then to the USF backbone. The Test System has its power management features enabled and automatically goes into Microsoft Windows XP standby state after a period of inactivity (set to 120 s). When the Test System awakes, the Proxy Emulator checks whether the Test system is powered up or sleeping by sending ARP Request packets at a rate of one every 5 s to the Test System and waiting for replies. If three consecutive ARP Requests do not elicit a reply from the Test System, then it is assumed that the Test System has transitioned to a sleep mode. At this point, the Proxy Emulator broadcasts an ARP Reply packet linking its hardware, or MAC address, to the Test System's IP address. This causes the computers and network devices on the laboratory LAN and USF backbone to update their ARP tables with the new address association. Now, any data packet sent to the Test System's IP address will be received by the Proxy Emulator. The Proxy Emulator replies to any ARP Request broadcasts for the Test System's MAC address with its own MAC address, ensuring that any data packets for the Test System will be redirected to it. As a demonstration, we have encoded the capability to respond to Ping requests. The Proxy Emulator will receive any ping echo packets sent to the Test system's IP address and will reply with an ICMP reply packet that appears to come from the Test System. Thus, the Proxy Emulator is responding on behalf of the sleeping Test System (and the Test System need not wake-up to handle pings). If the redirected packet received by the Proxy Emulator is a TCP SYN packet, the emulator will transmit a WOL packet to the Test System to wake-up the computer. At this point, the proxy emulator will again start probing the Test System to determine whether it is powered up or in sleep mode. Our Windows XP Test System used for testing can be woken-up from the Windows XP standby mode to a fully powered state in approximately 4 s by a WOL packet. We found that this wake-up time was sufficient for an HTTP browser sending a TCP SYN request not to time out. If newer technologies such as Intel's Instantly Available PC [16] and IBM's MRAM [9] continue to evolve we can expect the wake-up time to decrease to tens or hundreds of milliseconds.

## 7. Summary and future work

We have described ongoing work in investigating new power management methods for Internet-connected desktop computers. Simply addressing the current wake-up and message response problems will result in a higher adoption rate of power management in computers and operating systems that are currently shipping. The goal is that *always on need not be always fully powered on*. The first step in achieving this goal is characterizing network traffic at



Fig. 14. Photograph of phase 1 (emulator) proxying Ethernet adapter.

multiple levels. We showed that even the busiest computers on a university campus have many hours of idle time per day during which they could be power managed. We are refining estimates of the possible energy savings of this technology and working to gain the interest of the technology industry and relevant policy makers in bringing technical solutions to new products [26]. The potential impacts of this effort can be significant as measured in energy conservation in the many TWh per year (and the resulting decrease in greenhouse gases). A savings of 1 TWh/year equals \$80 million at 8 cents/kWh. The energy that we expect can be saved—with a value of 0.8–2.7 billion US dollars/year—should be compelling to the technology industry and relevant policy makers. In the long-term view are new devices such as residential set-top boxes. As millions of households acquire these set-top boxes - non-desktop computer devices that will have network connections and multiple power modes - energy efficiency of these devices will be critical.

### Acknowledgements

The authors thank Joe Rogers (USF academic computing) for collecting the USF dormitory NetFlow traces, Mamatha Kumar (USF graduate student) for data reduction of the traces, Bo Gullede (USF graduate student) for suggesting the Ethernet controller approach, and Rich Brown and Carrie Webber (LBNL) for contributions to the energy savings estimate.

### Appendix A. Details of the energy savings estimate

In this section, we show how the savings estimates of 0.8–2.7 billion US dollars/year for implementing solutions to the network problem were calculated. The stock and power level data are derived from the assumptions embedded in energy savings calculations done for the EPA Energy Star program [33,34]. All figures and discussion here, unless otherwise noted, refer to desktop PCs only (and do not include displays). To be conservative and to simplify the analysis, there are several sources of savings that we do not consider. The savings potentials are estimated separately for the residential and commercial sectors. Each sector is divided into two groups: *continuous* use machines, that are on for over 40 h/week, and *occasional* use machines, that are on for 40 h/week or less. We assume no savings for the occasional use PCs. Notebook PCs can be shifted from on to sleep just as with desktops. They have a smaller difference between on and sleep power, and different usage patterns from desktop PCs. We do not include these savings in our estimate. Since a sleeping PC cannot provide a video signal, it is likely that monitor power management rates will increase. We also ignore these potential savings. We assume the following parameters:

- *Stocks and power levels.* Stocks for each sector are taken directly from savings estimates prepared for the Energy Star program, as are the average power levels in on and sleep. Machines left on all the time may be more heavily loaded with power consuming hardware, but we assume that there is no correlation between these two factors and so no extra savings.
- *On-times.* Continuous on machines are assumed to be on 90% of the time (and so to be off or disconnected for the other 10%) to account for those machines on between 40 and 168 h/week.

For the remaining parameters, two sets of possible values are used to show a low and high estimate to bound the range of plausible estimates. The values and results are presented in Tables A1 and A2. In Table A1, the breakdown of commercial and residential PCs is 99.5 and 122.1 million, respectively, for a total of 221.6 million PCs.

Recent field studies of office equipment usage patterns showed about 60% of commercial PCs on at night [29]. We assume little change in this. By contrast, the residential rate for continuous on was 19% in 2001 according to the RECS survey [11] (split evenly between those left on all the time and those between 41 and 167 h/week), but we expect a variety of usage motivations to increase this percentage substantially. Reasons for more on-time and continuous operation include: more networked PCs (especially in homes), more intensive use of applications already in wide use (such as email and web surfing), relatively new applications (instant messaging and voice over IP for telephony), more use of data on PCs elsewhere in buildings, more use of data on PCs in remote locations, and backups in the commercial sector. In 2001 (per the RECS survey [11]), while only 37% of households with a PC have multiple PCs, 56% of residential PCs are in houses with more than one PC (this includes notebooks). By 2007, with the significantly increasing stock of home PCs, the percent of PCs in households with more than one PC should increase substantially. The presence of multiple PCs in a household may be the strongest driver of continuous on-time. Other factors also have this effect, such as persistent connections

Table A1  
Power levels, usage patterns, energy savings, and equipment stocks

Measure (calculated values indented)	Residential and commercial
On power (W)	59
Sleep power (W)	4
Sleep savings (W)	55
Hours per week actual usage for continuous PCs	40
Percent of time continuous PCs on (%)	90
Hours per week continuous PCs on	152
Hours per week shiftable	115
Energy savable (kWh/week)	6.3
Energy savable (kWh/year)	331
2007 stock (million)	221.6

Table A2  
Savings potentials

Measure (calculated values indented)	Low			High		
	Residential	Commercial	Total	Residential	Commercial	Total
Percent in continuous operation (%)	20	50		70	70	
Percent of continuous PCs affected (%)	30	50		70	70	
Percent of total PC stock affected (%)	6	25		49	49	
Total PCs affected (millions)	7.30	24.90	32.20	59.80	48.70	108.60
Energy savable (billion kWh/year)	2.40	8.22	10.60	19.80	16.10	35.90
Dollar value per year (\$ billion; direct)	0.20	0.60	0.80	1.50	1.20	2.70

The electricity prices used are 7.62 and 6.78 cents/kWh, respectively.

to the Internet, and to non-PC networked devices in the home such as media servers.

Table A1 shows the percent of continuous-on machines that might be shiftable from on to sleep for non-use times. Note that this is not the percent of machines that are shiftable, just the percent of the continuous portion that could be shifted. For example, for our commercial sector high savings estimate, we assume that 70% of all PCs will be in continuous operation and 70% of those will be shiftable, so that 49% of the entire stock of commercial PCs is shiftable. In certain corporate environments, machines can be shifted to sleep through use of Magic Packet; these are not included in the percent shiftable figure and so are not part of this savings estimate. There are a variety of reasons that we believe are common that lead to failure of network sleep to work to lead to disabling of power management include software updates from IT administrators, backups, and remote access. There is likely some percentage of PCs that would not be enabled for power management even with the network problem solved; this can be due to software that inhibits sleep, combinations of software and hardware that lead to malfunctions when entering or coming out of sleep. The fact that about 20% of printers have power management disabled suggests that some amount of disabling can be expected for any complex device.

The key results from this initial estimate are savings of 10 to 36 billion kWh/year, which amounts to 0.8–2.7 billion dollars/year. There are a number of simplifying assumptions that understate the savings. In addition, the results are fairly uncertain in general as there is little data to inform the key factors of percent of machines that will be on continuously in future and the percent of these that will be readily shiftable. Sales of PCs are relatively well known, though the split between the residential and commercial sectors is less clear. The lifetime in use of PCs is more uncertain, and usage patterns present the greatest uncertainty.

## References

- [1] IE: Software Products: Power Management: Overview, 2004, URL: <http://www.ie.com/SoftwareProducts/PowerManagement/Index.aspx>.
- [2] Advanced Configuration and Power Interface Specification, Rev. 2.0b, URL: <http://www.acpi.info/>.
- [3] S. Akella, R. Balan, N. Bandal, Protocols for low-power, Project Report, Carnegie Mellon University, Spring 2001, URL: <http://gs129.sp.cs.cmu.edu/papers/15744s01-low-power.ps>.
- [4] L. Benini, A. Bogliolo, G. De Micheli, A survey of design techniques for system level dynamic power management, IEEE Transactions on Very Large Scale Integration (VLSI) Systems 8 (3) (2000).
- [5] R. Bianchini, Research directions in power and energy conservation for clusters, Technical Report DCS-TR-466, Department of Computer Science, Rutgers University, November 2001.
- [6] J. Chase, C. Anderson, P. Thakar, A. Vahdat, Managing energy and server resources in hosting centers, 18th ACM Symposium on Operating System Principles, October, 2001, pp. 103–116.
- [7] J. Chase, R. Doyle, Balance of power: energy management for server clusters, Proceedings of the Eighth Workshop on Hot Topics in Operating Systems, May, 2001.
- [8] K. Christensen, The next frontier for communications networks: power management Proceedings of SPIE—Performance and Control of Next-Generation Communications Networks vol. 5244, 2003, pp. 1–4.
- [9] J. De Boeck, G. Borghs, Magnetoelectronic devices, Proceedings of the International Electron Devices Meeting 1999, pp. 215–218.
- [10] Embedded Ethernet, 2003, URL: <http://www.ethernet.de/en/index.html>.
- [11] Energy Information Administration (EIA), Residential Energy Consumption Surveys: 2001, Housing Characteristics Tables: Home Office Equipment Tables: Table HC7-1a-12a and HC7-1b-12b, Data for 2001, URL: <http://www.eia.doe.gov/emeu/recs/contents.html>.
- [12] EPA Energy Star, 2003, URL: <http://www.energystar.gov/index.cfm?c=home.index>.
- [13] EPA Energy Star: Computers and Monitors, Key Product Criteria, 2003, URL: [http://www.energystar.gov/index.cfm?c=computers.pr\\_crit\\_computers](http://www.energystar.gov/index.cfm?c=computers.pr_crit_computers).
- [14] J. Fuller, US Officials Cite Serious Energy Shortage, US Department of State, International Information Programs, April 2, 2001, URL: <http://usinfo.org/wf-archive/2001/010402/epf111.htm>.
- [15] M. Gupta, S. Singh, Greening of the internet, Proceedings of ACM SIGCOMM, August 2003.
- [16] Intel Instantly Available Technology, 2003, URL: <http://www.intel.com/technology/iapc/>.
- [17] Intel PRO Family of Network Adapters, 2003, URL: <http://support.intel.com/support/network/adapter/onlineguide/PRO1000/DOCS/DESKTOP/index.htm>.
- [18] C. Jones, K. Sivalingam, P. Agrawal, J. Chen, A survey of energy efficient network protocols for wireless networks, Wireless Networks 7 (4) (2001) 343–358.
- [19] K. Kawamoto, J. Koomey, B. Nordman, R. Brown, M. Piette, M. Ting, A. Meier, Electricity used by office equipment and network

- equipment in the US: detailed report and appendices, Technical Report LBNL-45917, Energy Analysis Department, Lawrence Berkeley National Laboratory, February 2001.
- [20] J. Koomey, Sorry, Wrong Number, *IEEE Spectrum* 40 (6) (2003) 11–12.
- [21] R. Latta (survey manager), Residential Energy Consumption Survey, Table 3. Electricity Consumption and Expenditures in US Households by End Uses and Census Region, Department of Energy, 1997, URL: [http://www.eia.doe.gov/emeu/recs/byfuels/table\\_3.pdf](http://www.eia.doe.gov/emeu/recs/byfuels/table_3.pdf).
- [22] C. Lefurgy, K. Rajamani, F. Rawson, M. Kistler, R. Keller, Energy management for commercial servers, *IEEE Computer* 36 (12) (2003) 39–48.
- [23] Magic Packet Technology, 2003, URL: [http://www.amd.com/us-en/ConnectivitySolutions/TechnicalResources/0,50\\_2334\\_2481,00.html](http://www.amd.com/us-en/ConnectivitySolutions/TechnicalResources/0,50_2334_2481,00.html).
- [24] S. Manne, A. Klauser, D. Grunwald, Pipeline gating: speculation control for energy reduction, *Proceedings of 25th International Symposium on Computer Architecture* 1998, pp. 132–141.
- [25] J. Mitchell-Jackson, Energy needs in an internet economy: a closer look at data centers, Master's Thesis, Energy and Resources Group, University of California at Berkeley, July 2001.
- [26] Network-present Sleeping PCs, 2003, URL: <http://eetd.lbl.gov/Controls/network/>.
- [27] Power Management: ENERGY STAR, 2004, URL: [http://www.energystar.gov/index.cfm?c=power\\_mgt.pr\\_power\\_management](http://www.energystar.gov/index.cfm?c=power_mgt.pr_power_management).
- [28] J. Roberson, personal communication, December 2003.
- [29] J. Roberson, C. Webber, M. McWhinney, R. Brown, M. Pinckard, J. Busch, After-hours power status of office equipment and inventory of miscellaneous plug-load equipment Lawrence Berkeley National Laboratory, Berkeley, CA, LBNL-53729, January 2004, URL: <http://enduse.lbl.gov/Projects/OffEqpt.html>.
- [30] A. Safwat, H. Hassanein, H. Moufta, A MAC-based performance study of energy aware routing schemes in wireless ad hoc networks, *Proceedings of IEEE GLOBECOM 2002*; 47–51.
- [31] R. Sasanka, C. Hughes, S. Adve, Joint local and global hardware adaptations for energy, *Proceedings of 10th International Conference on Architectural Support for Programming Languages and Operating Systems 2002*, pp. 144–155.
- [32] VERDIEM—Energy Efficiency for PC Networks, 2003, URL: <http://www.verdiem.com/company.shtml>.
- [33] C. Webber, R. Brown, J. Koomey, Savings estimates for the ENERGY STAR voluntary labeling program, *Energy Policy* 28 (15) (2000) 1137–1150.
- [34] C. Webber, R. Brown, A. Mahajan, J. Koomey. Savings estimates for the ENERGY STAR voluntary labeling program: 2001, Status Report, Lawrence Berkeley National Laboratory, Berkeley, CA, LBNL-48496, February 2002.
- [35] C. Webber, J. Roberson, R. Brown, C. Payne, B. Nordman, J. Koomey, Field Surveys of Office Equipment Operation Patterns, Lawrence Berkeley Laboratory, 2001 LBNL-46930, URL: <http://enduse.lbl.gov/Info/46930-abstract.html>.