EXTENDED ABSRACT FOR KEYNOTE TALK
The next frontier for communications networks: Power management

Kenneth J. Christensen[*]
Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620

ABSTRACT

Performance evaluation is used to gain an understanding of how to make the best use of scarce resources. Storage, memory, processing, and communications bandwidth are all relatively plentiful and inexpensive. What is the next frontier for communications networks and performance evaluation? I will argue that it is power management to achieve cost-effective operation. In the past few years, entirely new network protocols have been developed for battery-hungry sensor networks. But, what about the existing Internet? Estimates place the Internet as consuming from 2% to 8% of the total electricity produced in the USA – much of this power consumption is unnecessary. Do our "always on" desktop computers really need to be fully powered-up all the time? What can be done to achieve power-savings in these computers? The goal is to eliminate unnecessary energy usage by desktop computers in the near future and by networked embedded systems in the longer term. Traffic characterization is the first step towards this goal. Traffic characterization at inter-flow, intra-flow, and protocol levels is being done to investigate power management. The resulting savings achievable from relatively simple power management schemes are measured in TWh per year – or roughly equivalent to the electricity generated by one nuclear power plant. This is cost-effectiveness on a large scale!

Keywords: Power management, traffic characterization, Cisco NetFlow

1. INTRODUCTION

The amount of electricity consumed by devices connected to the Internet is rapidly increasing. Estimates place the Internet as consuming from 2% to 8% of the total electricity produced in the USA [3]. Significant growth in electricity consumption is expected as 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 [2]. However, such extreme projections have been shown to be based on incorrect assumptions [4]. A validated 1999 study [3] from Lawrence Berkeley National Laboratory (LBNL) states:

> We found that total direct power use by office and network equipment is about 74 TWh per year, which is about 2% of total electricity use in the U.S. 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 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.

Most of the expected savings would come from desktop computers and printers. What is preventing much of these savings from being achieved is the need to maintain network connectivity at all times. Desktop computers are 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. Existing power management features are disabled in 75% of PCs often because of network reasons [6]. In short, "always on" has come to mean "always fully powered on". Traffic characterization of networked desktop computers is the first step in addressing this problem. Traffic can be studied from inter-flow, intra-flow, and protocol perspectives. The goal is to derive insights that can be used to develop better power management methods. In this extended abstract, new directions in traffic characterization and development of proxying and wake-up capabilities for Ethernet Network Interface Controllers (NICs) are outlined.

---

[*] christen@csee.usf.edu; phone 1 813 974-4761; fax 1 813 974-5456; www.csee.usf.edu/~christen

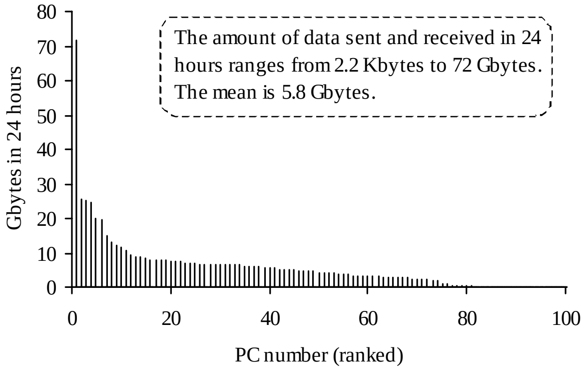

Figure 1. Bandwidth usage of 100 dormitory PCs

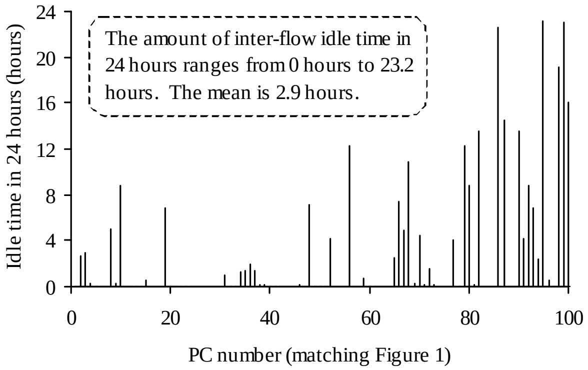

Figure 2. Idle time of 100 dormitory PCs

2. TENENTS OF POWER MANAGEMENT

Power management of a computer is possible whenever it is idle. Idle periods can be defined over multiple time scales. For each time scale, different power management methods are possible. The time scales are:

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

A flow can be a TCP connection bounded by connection establishment and termination. From a network view, inter-flow idle times are a conservative estimate of the available idle time for power management. Even within flows, there may be periods between packet transmissions that can be used for power management. Methods have been explored to bunch packet transmissions to increase idle time [1]. 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 this is 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 arenas including:

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

Historically, link bandwidth, memory, storage, and processor cycles have been the constraints for performance in computer networks and telecommunications systems. To a large extent, Moore's Law has removed these constraints – at least for the time being. Certainly, challenging problems in achieving quality of service still exist. Energy use is the major constraint for many future applications and for further deployment of existing applications.

3. INTER-FLOW CHARACTERIZATION OF PEER-TO-PEER APPLICATIONS

As a preliminary work, the amount of inter-flow idle time in University of South Florida (USF) dormitory desktop computers was measured. These computers contribute the majority of traffic to the university Internet connection; in a 24-hour period the dormitory computers send and receive 1350 Gbytes, which is about 60% of all USF network traffic for the day. These computers are known to be running modern file-sharing programs such as Kaaza, eDonkey, etc. Cisco NetFlow traces were collected from the USF dormitories for 24 hours on March 27, 2003. Figures 1 and 2 show the characterization results for sum of bytes sent and received and inter-flow idle time for the top 100 computers. These results show that there is considerable idle time for power management in what are likely the busiest desktop computers on campus.

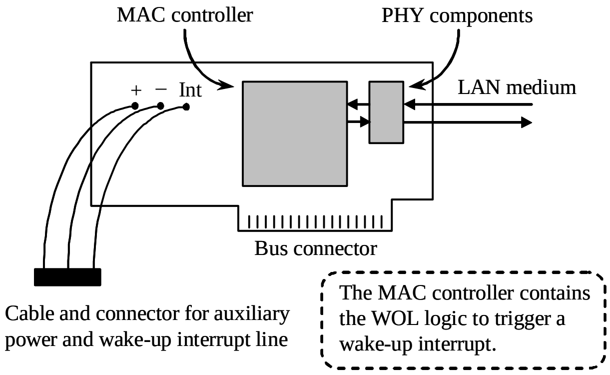

Figure 3. WOL Ethernet NIC with auxiliary power

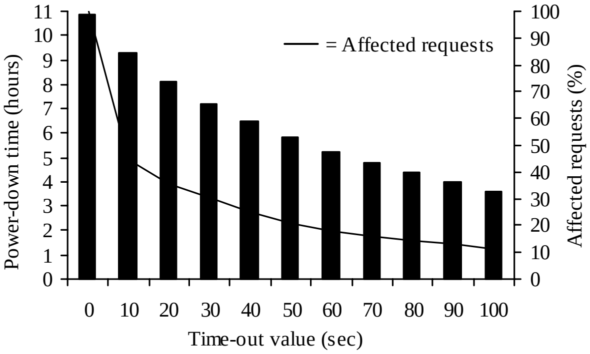

Figure 4. Power down time versus requests affected

The simplest method of dynamic power management is to set a fixed idle time-out value. If a system is idle for the time-out value, it is powered-down. The system is powered-up on an interrupt caused by detected activity or by a timed event. The time to power-up the system is a major consideration as it affects the response time of a request that triggers the wake-up. Existing commodity PCs and the Microsoft Windows operating system support power management. Windows PCs in a sleep state can use Wake on LAN (WOL) [7] Ethernet NICs to trigger a wake-up of the PC. WOL 10/100-Mbps Ethernet NICs are readily available at retail prices of about US $6 in mid-2003. Figure 3 shows a WOL NIC, which has three functions:

1. Connection to auxiliary power to allow partial NIC operation when the system bus and CPU powers-off.
2. Connection to a wake-up interrupt line 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 (called a "Magic Packet" by one vendor). On recognition of a Magic Packet, a wake-up interrupt signal is generated.

A Magic Packet is defined as any packet that in the data field contains the NIC Media Access Control (MAC) address repeated sixteen times. Since the Magic Packet requires use of a MAC address, it is difficult to use it across an IP-routed network. This Magic Packet is also not part of the existing TCP/IP connection packet flows or semantics. WOL has not achieved significant adoption or use. Improving wake-up semantics for Internet-connected computers is a significant open problem that will be addressed.

Figure 4 shows the trade-off in requests affected and amount of power-down time for one of the USF dormitory computers (in this case, computer #68 from Figure 1 with about 11 hours of inter-flow idle time in 24 hours). For example, about 6 hours of power-down time can be achieved with less than 21% of requests affected (i.e., affected as "first requests" triggering a power-up before being served). The next steps will be to 1) understand what is the performance impact of a wake-up, and 2) explore adaptive methods to improve upon a fixed idle time-out. Intra-flow idle periods will also be characterized, but these periods are expected to be much smaller than inter-flow idle periods.

4. POWER MANAGEMENT BY PROTOCOL PROXYING

There are no TCP/IP compatible or transparent mechanisms for triggering the wake-up of a powered-down system. There is no definition of what events should trigger wake-up and limited understanding for how long a system should remain awake. Protocols such as DHCP, ARP, ICMP, and others require periodic responses to messages – responses that cannot be made by a powered-down computer. Existent TCP connections cannot be maintained when a system powers down, even if a connection is to remain idle. This raises a key question: can a small processor on a NIC proxy for the larger system (that can then be powered-down)? The goal is to be able to develop proxying capabilities on a NIC to enable such power-down. This will allow existing desktop computer power management features to remain enabled!

To determine what needs to be proxied, a traffic study of packets arriving to and from an idle computer was done. A packet trace program was installed on an idle Windows 2000 PC in the Department of Computer Science and

Engineering. The PC is on a dedicated switched-Ethernet link, so only broadcast and direct-addressed packets will be seen at the PC. In 30 minutes over 6000 packets were traced. There packets can be classified into five categories:

1. Received packets that can be ignored (i.e., the packets are intended for other computers)
2. Received packets that require a state response
3. Received packets that require a state response and a state update
4. Received packets that require an application response
5. Transmitted packets originated by a protocol or application and not in response to a received packet

The majority of traced packets fall into category (1). Packets such as ICMP ping are in category (2). ARPs are in category (2) or (3). TCP connection requests are in category (4). DHCP lease renewals are in category (5). Which of these protocols (and applications) can be proxied on the NIC? Which require wake-up of the computer itself? How can the computer operating system inform the NIC what applications it has running (e.g., that have an outstanding listen for some port number) when it goes into a sleep state? In general, how is state information exchanged between the NIC and the computer? What size processor is needed on the NIC to filter and respond to incoming packets? These are all open questions.

Of the 6000 packets traced in 30 minutes, over 30% were bridge or routing related (e.g., spanning tree, RIP, IGMP, etc.) and can be ignored by a desktop computer. About 2% of the packets were ARPs (but, most were not directed to the traced computer). Much more work needs to be done in determining what traffic an idle computer receives and how this traffic should be handled to make proxying and wake-up possible. Packet interarrival time characteristics need to be understood so that the proxy processor can handle packet arrival bursts.

5. SUMMARY AND FUTURE WORK

This extended abstract has described ongoing work in investigating new power management methods for Internet-connected desktop computers. Simply “fixing” the wake-up and message response problems will result in higher adoption of power management in PCs 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 multiple levels. This work is a subject of an NSF proposal (jointly with Alan George at the University of Florida) currently in review. In conjunction with Bruce Nordman at LBNL, we are developing estimates of the potential energy savings of this technology and working to gain the interest of the technology industry and relevant policy makers in bringing technical solutions (as they are developed) to new products [5]. The potential impacts of this effort can be significant as measured in energy conservation in the many TWh/yr (and the resulting decrease in greenhouse gases). A savings of 1 TWh/yr equals $80 million at 8 cents per kWh. Achieving full night-time power down of existing office desktop computers will save a yearly amount of energy equivalent to that a typical nuclear power plant produces. With the assistance of LBNL, expected energy savings from the advanced wake-up and proxying implementations will be quantified. In the longer term are new devices such as set-top boxes. As millions of households acquire set-top boxes – non-PC devices that will have network connections and multiple power modes – energy efficiency of these devices will be critical.

REFERENCES

[1] S. Akella, R. Balan, and 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.

[2] J. Fuller, “U.S. Officials Cite Serious Energy Shortage,” U.S. Department of State, International Information Programs, April 2, 2001. URL: http://usinfo.state.gov/topical/global/climate/01040201.htm.

[3] K. Kawamoto, J. Koomey, B. Nordman, R. Brown, M. Piette, M. Ting, and A. Meier, “Electricity Used by Office Equipment and Network Equipment in the U.S.: Detailed Report and Appendices,” *Technical Report LBNL-45917*, Energy Analysis Department, Lawrence Berkeley National Laboratory, February 2001.

[4] J. Koomey, “Sorry, Wrong Number,” *IEEE Spectrum*, Vol. 40, No. 6, pp. 11-12, June 2003.

[5] Network-present Sleeping PCs, 2003. URL: http://eetd.lbl.gov/Controls/network/.

[6] B. Nordman, A. Meier, and M. Piette, “PC and Monitor Night Status: Power Management Enabling and Manual Turn-off.” *Proceedings of the 2000 ACEEE Summer Study on Energy Efficiency in Buildings*, pp. 89-99, 2000.

[7] “Wake up to Wake-on-LAN,” IBM Corporation, 1996. URL: http://www.networking.ibm.com/eji/ejiwake.html.