Ethernet Adaptive Link Rate (ALR): Analysis of a Buffer Threshold Policy

Chamara Gunaratne and Ken Christensen
Department of Computer Science and Engineering
University of South Florida
Tampa, Florida 33620
{pgunarat, christen}@cse.usf.edu

Stephen W. Suen
Department of Mathematics
University of South Florida
Tampa, Florida 33620
suen@math.usf.edu

Abstract— **Rapidly increasing energy use by computing and communications equipment is a significant problem that needs to be addressed. Ethernet network interface controllers (NICs) consume hundreds of millions of US$ in electricity per year. Most Ethernet links are underutilized and link power consumption can be reduced by operating at lower data rates. An output buffer threshold policy to change link data rate in response to utilization is investigated. Analytical and simulation models are developed to evaluate the performance of Adaptive Link Rate (ALR) with respect to mean packet delay and time spent in low data rate with Poisson traffic and 100 Mb/s network traces as inputs. A Markov model of a state-dependent service rate queue with rate transitions only at service completion is developed. For the traffic traces, it is found that a link can operate at 10 Mb/s for over 99% of the time yielding energy savings with no user-perceivable increase in packet delay.**

Keywords— **Ethernet; energy efficiency; adaptive link rate (ALR); performance analysis**

I. INTRODUCTION

Global warming and energy dependence are major concerns today. The Internet has great potential for reducing energy use by enabling telecommuting, teleconferencing, ecommerce, and other new modalities. However, the Internet and all the hosts that connect to it are themselves becoming significant energy consumers. The hubs, switches, and routers that comprise the Internet consumed an estimated 6.15 TWh/yr in 1999 [14]. One TWh/yr corresponds to $80 million at $0.08 per kW/hr. Projections of growth are less clear, but it appears that electronic equipment (most of it connected to the Internet) is the fastest growing consumer of electricity. Today 9% of the energy used by commercial buildings is due to electronic office equipment [14]. The large energy use of the Internet and possible directions for reducing it were first examined by Gupta and Singh [11] and Christensen et al. [3]. The possibility of adaptively varying Ethernet link data rate in response to utilization to reduce energy use was first proposed (but not modeled) in Christensen et al. [9]. Adaptive Link Rate (ALR) for Ethernet was presented by Nordman and Christensen to the IEEE 802.3 in July 2005 [12]. A formal design and simulation study of ALR was done by Gunaratne and Christensen [8]. ALR proposes to vary Ethernet link data rate in multiples of 10 using existing data rates (i.e., from 10 Mb/s to 10 Gb/s). ADSL2 (ITU-T Recommendation G.992.3 and G.992.5) already supports multiple link rates and power states [6]. ADSL2 is the recommended technology for home broadband access in the EU. The next generation of VDSL2 has less power management than ADSL2. The EU Stand-by Initiative has the goal to "trigger action on energy efficiency within standardization" for VDSL2 [4].

One attribute of a "perfect network" must be energy use proportional to utilization. We have measured a difference of about 4 W on a PC to first-level LAN switch Ethernet link for 1 Gb/s versus 100 Mb/s [9]. These measurements were taken at the PC and LAN switch with a Watt meter at the wall socket. For 10 Gb/s versus 1 Gb/s, the difference may be much greater, possibly 10 to 20 W per link [9]. Measurements of Ethernet links suggest that utilization is generally less than 5% [3, 13]. Thus, there may be opportunity for significant energy savings with only minimal impact to packet delay by operating links at a lower data rate during low utilization periods.

Varying Ethernet link data rate as a function of link utilization is a trade-off between energy use and packet delay. In this paper, we develop performance models for a buffer threshold policy for ALR. Section 2 describes the ALR mechanism and buffer threshold policy. Section 3 develops a Markov model. Section 4 extends the model using simulation and also presents results for actual Ethernet link traffic based on traffic traces. Section 5 is an estimate of possible energy savings if ALR were widely adopted. Finally, Section 6 summarizes and describes future work.

II. ETHERNET ADAPTIVE LINK RATE (ALR)

Key challenges in realizing ALR are to define a mechanism for quickly switching link data rate and create policies to change the data rate without adversely affecting packet delay.

A. ALR MAC Frame Handshake Mechanism

A fast mechanism for initiating and agreeing upon a link rate change is necessary. Either end of a link (e.g., the NIC in a desktop PC or switch port in a LAN switch) must be able to initiate a request to change the rate. It is also necessary at link establishment to negotiate existing capabilities including support of ALR at both ends and the possible data rates that are in common between both ends. Existing 802.3 Auto-Negotiation can be used at link establishment for capability negotiation (e.g., using unformatted code word pages). Auto-Negotiation could also be used to switch data rates during link operation, but this would require a minimum of 256 ms for a

This material is based on work supported by the National Science Foundation under Grant No. 0520081.

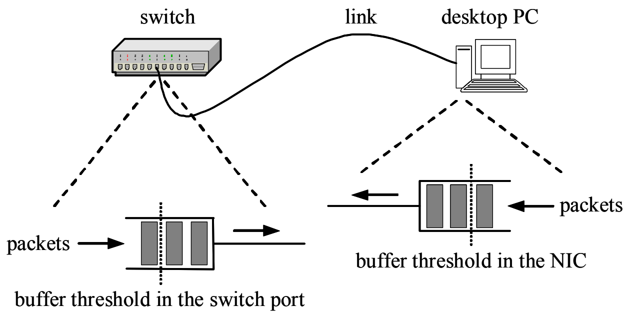

Figure 1. System view of an ALR capable Ethernet link

1000BASE-T implementation. A faster handshaking could be implemented using Ethernet MAC frames. A two-way handshake could be implemented as:

- One link partner requests a data rate change using an ALR REQUEST MAC frame.
- The receiving link partner acknowledges the data rate change request with either an ALR ACK or ALR NACK MAC frame.

An ALR ACK response would trigger a link rate switch (rate transition) and link resynchronization. The total time for handshake plus resynchronization could be less than 100 µs for 1 Gb/s Ethernet. This is based on 10,000 clock cycles needed for link resynchronization. It is likely that link resynchronization can be achieved in much less than 10,000 clock cycles. For 10 Gb/s more work needs to be done to understand if link retraining needs to be done every time a link is resynchronized, or only at link establishment. Fig. 1 is a system view of an ALR capable Ethernet link showing output buffers with thresholds to be used in the ALR policy.

B. *ALR Buffer Threshold Policy*

When and under what conditions the link data rate is switched is determined by the policies that can be implemented in an Ethernet NIC and switch port. Additional cost and power consumption has to be minimized and therefore simplicity is important. We define that packets arrive at a rate λ and are serviced at a low service rate μ_1 or high service rate μ (i.e., $\mu_1 < \mu$). We define respective utilizations as $\rho_1 = \lambda/\mu_1$ and $\rho = \lambda/\mu$ where $\rho < 1$ for stability.

In a single threshold policy, the output buffer occupancy level equaling or exceeding a threshold value k causes the service rate to switch from μ_1 to μ. An output buffer occupancy level dropping below k causes the service rate to switch from μ to μ_1. In a dual threshold policy, two buffer occupancy thresholds k_1 and k_2 are defined where $k_1 < k_2$. An output buffer occupancy level equaling or exceeding threshold value k_2 causes the service rate to switch from μ_1 to μ if the present rate is μ_1. An output buffer occupancy level dropping below threshold value k_1 causes the data rate to switch from μ to μ_1 if the present rate is μ.

Of interest are the packet delay and time spent in low service rate as a function of threshold values, arrival and services rates, and rate switching time. We seek to develop performance models to gain insight into the effects of these parameters on packet delay and energy saved. Energy is saved when the link is in low service rate (low data rate).

III. Markov Models for ALR

Given standard Poisson arrivals and exponential service time assumptions, ALR can be modeled as a state-dependent service-rate, single-server queue. Markovian traffic assumptions do not capture burstiness and self-similarity properties often seen in real network traffic. However, a Markov model can yield significant insights into fundamental behaviors of a protocol – insights that apply to other traffic types as well. In this section, we model single and dual threshold systems with rate transition possible 1) during a service and 2) only on service completion. In the case of an Ethernet link, rate transition clearly cannot occur during a packet transmission; it can only occur after the completion of packet transmission (i.e., on service completion). We derive expressions for the steady state probability of n customers (packets) in the system, P_n, and the mean number of customers in the system, L (in the derivations we use π_n to denote the steady state probability of state n and unless explicitly stated, P_n does not equal π_n). Relative time in low rate is the sum of steady state probabilities of states with service rate μ_1.

A. *Single Threshold, Rate Transtion During Serivce*

The simplest system is that of a single threshold queue with state-dependent service rate where a service rate transition can occur during a service period. Fig. 2 is the well known Markov model for this system [7]. For this system, closed-form expressions are readily available:

$$P_n = \begin{cases} P_0 \rho_1^n & (0 \le n < k) \\ P_0 \rho_1^{k-1} \rho^{n-k+1} & (n \ge k) \end{cases} \tag{1}$$

$$P_0 = \left(\frac{1-\rho_1^k}{1-\rho_1} + \frac{\rho \rho_1^{k-1}}{1-\rho} \right)^{-1} \tag{2}$$

$$L = P_0 \left(\frac{\rho_1 \left(1 + (k-1)\rho_1^k - k\rho_1^{k-1}\right)}{(1-\rho_1)^2} + \frac{\rho \rho_1^{k-1} \left(k - (k-1)\rho\right)}{(1-\rho)^2} \right) \tag{3}$$

B. *Dual Threshold, Rate Transtion During Serivce*

Extending the single threshold model to two thresholds was first done by Gebhard [5]. The Markov chain for this system is shown in Fig. 3. The two rows of states for k_1 to $k_2 - 1$ customers in the system model the cases for the system being in rate μ_1 or μ depending on the previous threshold crossing and rate. The number of customers in the system is shown below the Markov chain. The following closed-form expressions are from [5] and are converted to the notation used in this paper:

$$P_n = \begin{cases} P_0 \rho_1^n & (0 \le n < k_1) \\ \dfrac{P_0 \rho_1^{k_1-1}(1-\rho_1)}{\left(1-\rho_1^{k_2-k_1+1}\right)} \left(\dfrac{\rho_1^{n-k_1+1} - \rho_1^{k_2-k_1+1}}{1-\rho_1} + \dfrac{\rho_1^{k_2-k_1} \rho \left(1-\rho^{n-k_1+1}\right)}{1-\rho} \right) & (k_1 \le n < k_2) \\ \dfrac{P_0 \rho_1^{k_2-1} \rho^{n-k_2+1} (1-\rho_1)\left(1-\rho^{k_2-k_1+1}\right)}{(1-\rho)\left(1-\rho_1^{k_2-k_1+1}\right)} & (n \ge k_2) \end{cases} \tag{4}$$

$$P_0 = \left(\frac{1}{1-\rho_1} - \frac{(1+k_2-k_1)\rho_1^{k_2-1}(\rho_1-\rho)}{\left(1-\rho_1^{k_2-k_1+1}\right)(1-\rho)} \right)^{-1} \tag{5}$$

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

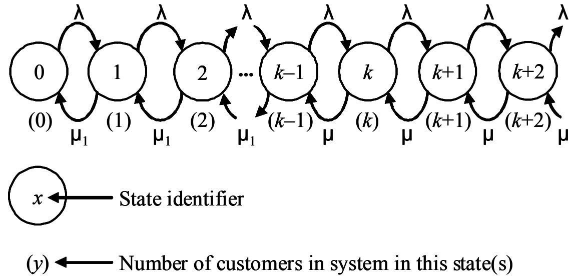

Figure 2. Single threshold, rate transition during service

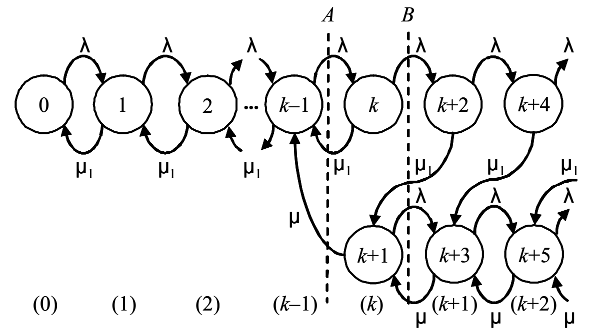

Figure 4. Single threshold, rate transition at service completion

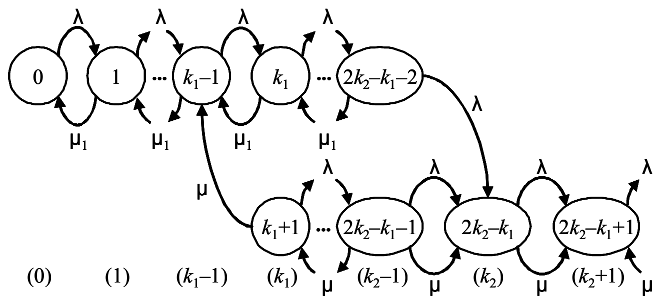

Figure 3. Dual threshold, rate transition during service

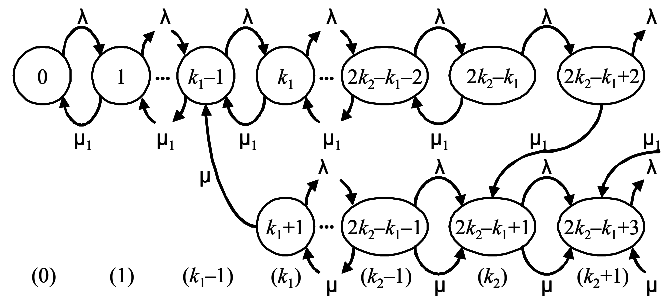

Figure 5. Dual threshold, rate transition at service completion

C. Single Threshold, Rate Transtion at Service Completion

To correctly model ALR for Ethernet, rate change can only occur at a service completion (i.e., at the completion of sending a packet). Fig. 4 shows the Markov chain for this system, which is similar to the Markov chain presented by Chong and Zhao [2]. Chong and Zhao solve the Markov chain using transforms, we solve it directly in a novel manner. The states k, $k+2$, $k+4$, … $k+2n$ ($n \geq 0$) model the case where arrivals causing a threshold crossing occur during a service period (and the rate transition occurs first on the completion of the current service). The steady state probabilities of these states are the probability of receiving n arrivals during a service interval. The cumulative probability of receiving an arrival before time t is

$$F(t) = \int_0^t \lambda e^{-\lambda y} dy = 1 - e^{-\lambda t}. \quad (6)$$

Due to the memoryless property of exponentially distributed events, the probability of receiving an arrival before the current service (with rate μ_1) is completed is

$$\int_0^\infty \left(1 - e^{-\lambda t}\right) \mu_1 e^{-\mu_1 t} dt = \frac{\lambda}{\lambda + \mu_1}. \quad (7)$$

In view of (7), the steady state probabilities for states k, $k+2$, $k+4$, … are related by the following equations:

$$\pi_k = \left(\frac{\lambda}{\lambda + \mu_1}\right) \pi_{k-1}, \quad (8)$$

and for $n \geq 1$,

$$\pi_{k+2n} = \left(\frac{\lambda}{\lambda + \mu_1}\right) \pi_{k+2n-2}, \quad (9)$$

giving that for $n \geq 1$,

$$\pi_{k+2n} = \left(\frac{\lambda}{\lambda + \mu_1}\right)^{n+1} \pi_{k-1}. \quad (10)$$

By partitioning the Markov chain into two subsets with no shared states and writing the balance equations for the state transitions between the subsets, we are able to derive the steady state probabilities for the number of customers in the system. In Fig. 4, we see that up to state $k-1$ the model is similar to an M/M/1 queue and therefore

$$P_{k-1} = \pi_{k-1} = P_0 \rho_1^{k-1}. \quad (11)$$

Partitioning the chain on cut (A) yields

$$\lambda \pi_{k-1} = \mu_1 \pi_k + \mu \pi_{k+1}. \quad (12)$$

Partitioning the chain on cut (B) yields

$$\lambda \pi_k + \lambda \pi_{k+1} = \mu_1 \pi_{k+2} + \mu \pi_{k+3}. \quad (13)$$

From Fig. 4, we see that

$$P_k = \pi_k + \pi_{k+1}. \quad (14)$$

By substituting from (10) and (11), (14) can be written as

$$\pi_{k+1} = P_k - P_{k-1} \left(\frac{\lambda}{\lambda + \mu_1}\right). \quad (15)$$

Similarly, we can write

$$\pi_{k+3} = P_{k+1} - P_{k-1} \left(\frac{\lambda}{\lambda + \mu_1}\right)^2. \quad (16)$$

By substituting from (10) and (15) in (12), we get

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

$$\lambda P_{k-1} = \mu_1 P_{k-1}\left(\frac{\lambda}{\lambda+\mu_1}\right) + \mu\left(P_k - P_{k-1}\left(\frac{\lambda}{\lambda+\mu_1}\right)\right). \quad (17)$$

By substituting from (10), (15), and (16) in (13), we get

$$\lambda P_k = \mu_1 P_{k-1}\left(\frac{\lambda}{\lambda+\mu_1}\right)^2 + \mu\left(P_{k+1} - P_{k-1}\left(\frac{\lambda}{\lambda+\mu_1}\right)^2\right). \quad (18)$$

Finally, from (17) and (18) we can generate the following recursive equation for the steady state probability of having n packets in the system where $n \ge k$

$$P_n = \rho P_{n-1} + \left(1 - \frac{\rho}{\rho_1}\right) P_{k-1}\left(\frac{\rho_1}{1+\rho_1}\right)^{n-k+1}. \quad (19)$$

Equation (19) can be iteratively solved and the final closed-form equations for the steady state probabilities and P_0 are:

$$P_n = \begin{cases} P_0 \rho_1^n & (0 \le n < k) \\ \dfrac{P_0 \rho_1^{k-1}}{\rho + \rho/\rho_1 - 1}\left(\rho^{n-k+2} - \left(1 - \dfrac{\rho}{\rho_1}\right)\left(\dfrac{\rho_1}{1+\rho_1}\right)^{n-k+1}\right) & (n \ge k) \end{cases} \quad (20)$$

$$P_0 = \left(\frac{1-\rho_1^k}{1-\rho_1} + \frac{\rho_1^k}{1-\rho}\right)^{-1} \quad (21)$$

Note that the equation (2) for P_0 in Section III.A can be rewritten as

$$P_0 = \left(\frac{1-\rho_1^{k-1}}{1-\rho_1} + \frac{\rho_1^{k-1}}{1-\rho}\right)^{-1} \quad (22)$$

and is therefore very similar to (21).

D. Dual Threshold, Rate Transtion at Service Completion

The Markov model for the dual threshold queue with rate transition at service completion is shown in Fig. 5. Though more complex than the Markov model for the single threshold system, similar direct techniques can be used to derive the steady state probabilities. The steady state probabilities are given by (23) shown below. The probability of zero customers in the system, P_0, is given by

$$P_0 = \left(\frac{1-\rho_1^{k_1}}{1-\rho_1} + \frac{1}{1-\rho_1^{k_2-k_1+2}}\left(\frac{\rho_1^{k_2}-\rho_1^{k_1}}{\rho_1 - 1} + \frac{\rho_1^{k_2}\left((k_2-k_1)(\rho_1-\rho)+\rho_1^2-1\right)}{\rho-1}\right)\right)^{-1}. \quad (24)$$

IV. Performance Evaluation of ALR

With the Markov models developed in the previous section the performance of ALR can be evaluated for Poisson arrivals, exponential services times, and a rate switching time of zero seconds. To extend the Markov models, a simulation model of a single-server queue with state-dependent service time was developed using CSIM19 [15] allowing the investigation of performance for the cases of 1) non-zero rate switching times and 2) actual (traced) Ethernet traffic. The simulation model is freely available from the authors and is described in [8]. Two general assumptions are:

- An Ethernet link supports two data rates (e.g., 100 Mb/s as low rate and 1 Gb/s as high rate).
- Time spent changing from low rate to high rate is considered (for power use) as time spent in high rate, and time spent switching from high to low rate is considered as time spent in low rate.

The performance metrics of interest are mean packet delay and percentage of time spent in the (low power) low data rate.

A. Performance Evaluation with Markov Assumptions

The fraction of time in low rate and the mean packet delay are measured using the dual threshold, rate transition at service completion Markov model (Fig. 5). The low service rate μ_1 was set to 0.1, and the high service rate μ was set to 1.0 (Ethernet data rates increase in multiples of 10). The threshold values k_1 and k_2 were set to 15 and 30 customers, respectively (NIC buffers are generally larger than 16 KBytes [1]). The rate of arrivals λ was varied from 1% to 25% of μ. The numerical results for this case (or experiment) are shown in Fig. 6 where the bars shows the percentage of time in low rate and the line shows mean packet delay. The time units for measuring delay are units of mean service interval at high service rate (μ). The CSIM19 simulation model was validated by exactly reproducing the numerical results of Fig. 6. The simulation model was used to repeat the previous experiment with non-zero switching (rate transition) times of 10 and 100 mean service intervals at the high service rate. The simulation results are shown in Fig. 7. The number of service rate transitions per 1000 time units for switching times of 0, 10 and 100 time units as the rate of arrivals is increased is shown in Fig. 8.

In Fig. 6 and Fig. 7 we observe that for up to 9% offered load (offered load is relative to the high service rate) the percentage of time in low rate is approximately 100% and constant. This is not surprising since in this region the system is an M/M/1 with low rate service time at 90% offered load (with mean queue length of 9 customers). With an increase in the rate of arrivals, the fraction of time in low rate decreases

$$P_n = \begin{cases} P_0\rho_1^n & (0 \le n < k_1) \\ \dfrac{P_0 \rho_1^{k_1-1}(1-1/\rho_1)}{1-1/\rho_1^{k_2-k_1+2}}\left(\dfrac{1-1/\rho_1^{k_2-n+1}}{1-1/\rho_1} + \dfrac{\rho\left(1-\rho^{n-k_1+1}\right)}{1-\rho}\right) & (k_1 \le n < k_2) \\ \dfrac{P_0\rho_1^{k_1-1}}{1-1/\rho_1^{k_2-k_1+2}}\left(\rho^{n-k_2+1}\left(\dfrac{(1-1/\rho_1)\left(1-\rho^{k_2-k_1}\right)}{1-\rho} - \dfrac{1-1/\rho_1^2}{1-\rho-\rho/\rho_1}\right) + \dfrac{\left(1-1/\rho_1^2\right)(1-\rho/\rho_1)}{1-\rho-\rho/\rho_1}\left(\dfrac{\rho_1}{1+\rho_1}\right)^{n-k_2+1}\right) & (n \ge k_2) \end{cases} \quad (23)$$

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

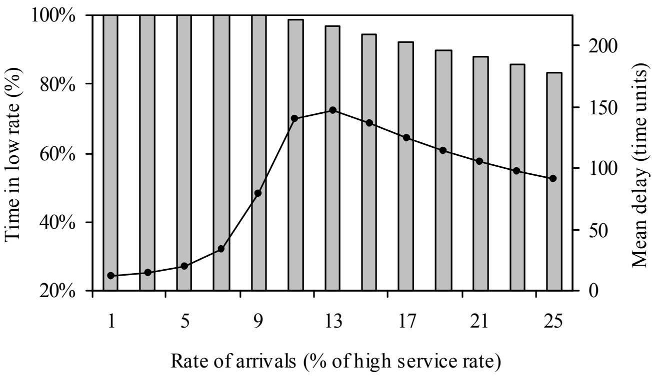

Figure 6. Numerical results for Markov model

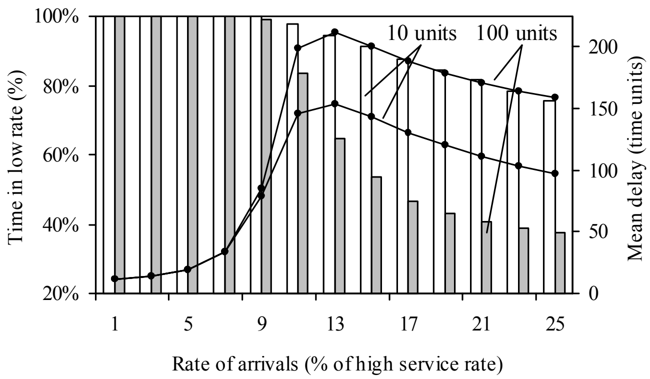

Figure 7. Simulation results with non-zero switching time

with the biggest decrease shown (not surprisingly) for the largest switching time. For zero switching time with offered load at 25%, over 80% of the time is in low data rate. However, for a switching time of 100 time units the percentage of time in low rate is less than 40% and the majority of that time is spent in rate switching and not servicing arrivals at the low rate. For switching times of 0 and 10 time units, the number of switches thereafter increase with the increasing rate of arrivals as the modeled link oscillates between high and low service rates. For a switching time of 100 time units, it can be seen in Fig. 8 that the number of rate switches increases initially and then starts to decrease. This is due to the greater number of arrivals queued during the longer rate switching time, which reduces the time spent servicing arrivals at the low service rate (μ_1) as the upper threshold k_2 is exceeded quicker than in the case of 0 or 10 time unit switching times. As the switching time increases, the time to drain the queue increases causing less switching activity. With the switching time of 100 time units, the time spent servicing arrivals at low service rate (μ_1) becomes insignificant with increasing arrival rate as by the time the service rate change is complete the queue length is again greater than k_2.

B. Performance Evaluation with Trace Traffic

Six traffic traces were collected in mid-2004 from six 100 Mb/s Ethernet links. USF #1 to USF #3 are from the University of South Florida network. The USF "busy" traces are from links to student dormitory rooms and consist mainly of file-sharing traffic. The USF "average" trace is from the link to an office workstation and consists mainly of Windows workgroup and web traffic. Three traces are from the Portland State University network (PSU #1 to PSU #3) [10]. Table 1 summarizes the trace characteristics. Average link utilization is less than 5% in all cases.

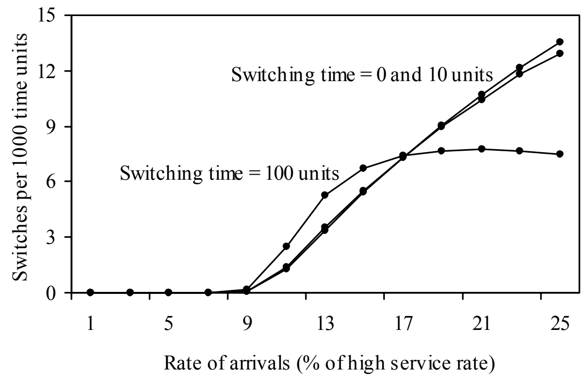

Figure 8. Rate switches per 1000 time units

TABLE 1
SUMMARY OF TRACES

Trace	Duration	Description	Avg util.[1]
USF #1	0.5 hours	Link to "busiest" user in USF	4.11 %
USF #2	0.5	Link to 10th busiest user	2.63
USF #3	0.5	Link to an average user	0.03
PSU #1	2.0	Link to a desktop PC	0.13
PSU #2	2.0	Link connecting two switches	1.01
PSU #3	2.0	Link connecting switch to router	1.03

[1] Calculated at 100 Mb/s with 1518 byte packets assumed for PSU traces.

TABLE 2
MEAN PACKET DELAY OF TRACES

Trace	10 Mb/s	100 Mb/s	ALR[1]	
USF #1	7.60 ms	0.09 ms	2.79 ms	99.42 %
USF #2	3.95	0.08	1.81	99.81
USF #3	196.29	0.05	1.48	99.99
PSU #1[2]	33.51	0.18	5.63	99.98
PSU #2[2]	2321.31	0.12	9.55	99.12
PSU #3[2]	1147.83	0.51	4.07	99.83

[1] First column is mean packet delay and second column is time in low rate.
[2] Packet size is assumed to be 1518 bytes for PSU traces.

The mean packet delay at fixed 10 Mb/s and 100 Mb/s data rates was measured using the simulation model. The results for packet delay are shown in Table 2. Clearly, operation at 10 Mb/s only results in excessive packet delays. For the ALR simulation, the rate switching time was set to 1 ms and the threshold values k_1 and k_2 were set to 15 and 30 packets, respectively. The results for both packet delay and percentage of time in 10 Mb/s are shown in Table 2. Fig. 9 shows the simulation results for the USF #1 trace as the rate switching time is increased from 0 ms to 50 ms. The simulation parameters were the same as the above experiment with results in Table 2 except for the rate switching time.

We observe that for all the traces, the percentage of time in low rate exceeds 99% with ALR and that the increase in mean packet delay, though greater than when operating at 100 Mb/s, is an average of 4 ms greater and this is most likely not perceivable by the end-users. This indicates that significant energy savings potential exists with ALR. Further, in Fig. 9 we observe that for the trace USF #1 the mean packet delay remains stable until the rate switching time increases beyond 15 ms (at which time the increase in packet delay is approximately linear with the increase in rate switching time).

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

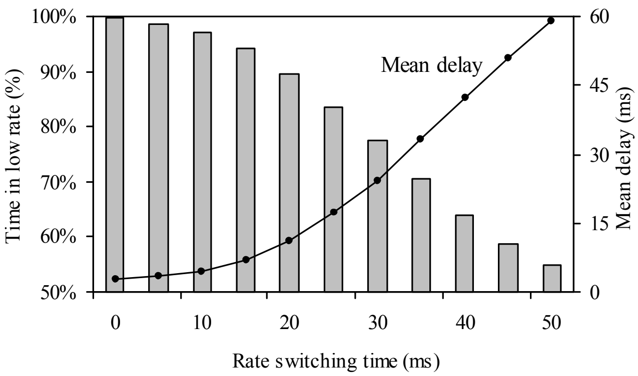

Figure 9. Results for USF #1 trace

As described in Section II.A, the rate switching time for ALR is expected to be less than 1 ms and at these values significant energy savings are possible.

V. ESTIMATE OF ENERGY SAVINGS

If the estimated 160 million Ethernet-connected PCs in the US could operate their link at a lower data rate for 80% of the time that they are actively used or in sleep mode (about 14 hours per day on average), a savings of $240 million per year would be achieved [9]. The results for the ALR policy investigated in this work show that assuming 80% of the time is spent in low data rate is conservative and that time in low rate can be close to 100%, thus enabling greater savings. If ALR is standardized by the IEEE 802.3, products supporting ALR will over time replace and supplement existing NICs and switch ports, and energy savings of hundreds of millions of US$ per year can be expected.

VI. SUMMARY AND FUTURE WORK

Adaptive Link Rate (ALR) is a method to reduce energy use of Ethernet links by adapting the link data rate to link utilization. The utilization of Ethernet links is, on average, extremely low. This suggests that there is ample opportunity for energy savings by operating Ethernet links at a low data rate for most of the time, and with no perceivable performance impact to the user. We showed that a simple output buffer dual threshold policy could be used to trigger link rate changes and to do so with negligible increase in mean packet delay. Energy savings of hundreds of millions of US$ per year in the US alone are achievable if ALR is widely adopted.

A significant contribution of this paper is the Markov model for a state-dependent service rate, single server queue where service rate transitions can only occur at the completion of a service (this exactly models the behavior of a packet switched network). Rate transition at service completion increases the mean number of arrivals in the system and the increase appears to be bounded by λ/μ_1. Proving this for both the single and dual threshold cases remains as future work.

For smooth traffic the dual threshold policy can result in link data rate oscillation. For example, for 10 and 100 Mb/s data rates, 15% offered load, and a rate switching time of 1 ms, about 50 rate switches per second would occur consuming 5% of the time for rate switching. Future work will address, 1) a stabilizing policy suitable for smooth and bursty traffic, 2) context-specific clues for initiating link data rate switches, 3) realistic modeling of 1 Gb/s and future 10 Gb/s traffic in order to evaluate ALR performance for future very high bandwidth (and very high energy use) desktop to LAN switch links, and 4) investigation of transient effects on packet delay and what effect they may have on user-perceived application performance. Future work will also explore deeper energy savings possible in LAN switches when large parts of a switch may be able to operate at a reduced data rate.

ACKNOWLEDGMENT

We thank Dr. Suresh Singh at Portland State University for making the network traces used in [10] available to us and Mr. Joe Rogers at the University of South Florida for collecting traces from the USF network. We also thank the anonymous referees for their helpful comments.

REFERENCES

[1] Broadcom BCM5701 10/100/1000BASE-T controller product brief, 2005. URL: http://www.broadcom.com/collateral/pb/5701-PB10-R.pdf.

[2] E. Chong and W. Zhao, "Performance evaluation of scheduling algorithms for imprecise computer systems," *Journal of Systems and Software*, Vol. 15, No. 3, pp. 261-277, July 1991.

[3] K. Christensen, C. Gunaratne, B. Nordman, and A. George, "The next frontier for communications networks: power management," *Computer Communications*, Vol. 27, No. 18, pp. 1758-1770, December 2004.

[4] EU stand-by initiative, Minutes of the third meeting on energy consumption of broadband communication equipment and networks, European Commission DG JRC, Ispra, November 29, 2005. URL: http://energyefficiency.jrc.cec.eu.int/html/standby_initiative.htm.

[5] R. Gebhard, "A Queueing process with bilevel hysteretic service-rate control," *Naval Research Logistics Quarterly*, Vol. 14, pp. 55-68, 1967.

[6] G. Ginis, "Low power modes for ADSL2 and ADSL2+," Broadband Communications Group, Texas Instruments, SPAA021, January 2005.

[7] D. Gross and C. Harris, *Fundamentals of Queueing Theory*, 3rd ed., New York: John Wiley & Sons, 1998, p. 89.

[8] C. Gunaratne and K. Christensen, "Ethernet adaptive link rate: system design and performance evaluation," *Proceedings of the IEEE Conference on Local Computer Networks*, November 2006.

[9] C. Gunaratne, K. Christensen, and B. Nordman, "Managing energy consumption costs in desktop PCs and LAN switches with proxying, split TCP connections, and scaling of link speed," *International Journal of Network Management*, Vol. 15, No. 5, pp. 297-310, September/October 2005.

[10] M. Gupta, S. Grover, and S. Singh, "A feasibility study for power management in LAN switches," *Proceedings of the 12th IEEE International Conference on Network Protocols*, pp. 361-371, October 2004.

[11] M. Gupta and S. Singh, "Greening of the Internet," *Proceedings of ACM SIGCOMM*, pp. 19-26, August 2003.

[12] B. Nordman and K. Christensen, "Reducing the energy consumption of network devices," Tutorial presented at the July 2005 IEEE 802 LAN/MAN Standards Committee Plenary Session, July 2005. URL: http://www.csee.usf.edu/~christen/energy/pubs.html.

[13] A. Odlyzko, "Data networks are lightly utilized, and will stay that way," *Review of Network Economics*, Vol. 2, No. 3, pp. 210-237, September 2003.

[14] K. Roth, F. Goldstein, and J. Kleinman, "Energy consumption by office and telecommunications equipment in commercial buildings, Volume I: energy consumption baseline," Arthur D. Little Reference No. 72895-00, January 2002.

[15] H. Schwetman, "CSIM19: a powerful tool for building system models," *Proceedings of the 33rd Winter Simulation Conference*, pp. 250-255, December 2001.

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.