

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

Chamara Gunaratne

Ken Christensen

Computer Science and Engineering
University of South Florida
Tampa, FL 33620
{pgunarat, christen}@cse.usf.edu

Stephen Suen

Mathematics
University of South Florida
Tampa, FL 33620
suen@math.usf.edu

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

Agenda

- **Introduction**
- **ALR threshold policy**
- **Analysis**
- **Summary and future work**

Introduction

- **Internet is a major consumer of electricity**
 - Over 6 TWh/yr in 2000 (not including end nodes) [14]
- **One TWh/yr = \$80 million**
 - 0.75 million tons of CO₂
- **Ethernet links operate at very low utilization**
 - Typical link utilization is less than 5%
- **Adaptive Link Rate (ALR) proposed in 2005**
 - Matches link rate to utilization
 - Energy saved by operating at a lower link rate most of time
 - Potential to save a large amount of electricity

[14] K. Roth, F. Goldstein, and J. Kleinman, "Energy consumption by office and telecommunications equipment in commercial buildings, Volume 1," Arthur D. Little Reference No. 72895-00, January 2002.

Agenda

- Introduction
- **ALR threshold policy**
- Analysis
- Summary and future work

Overview of ALR

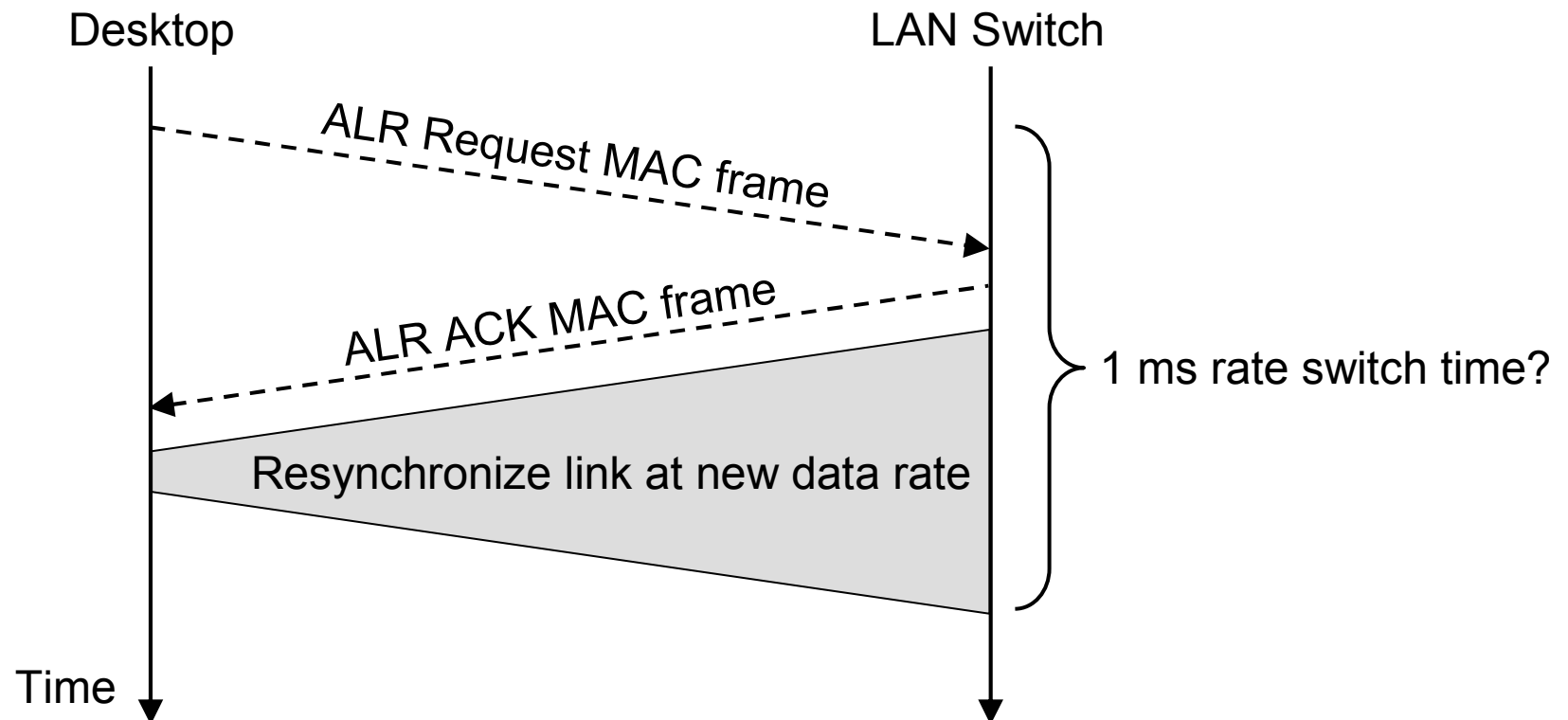
- **Adaptive Link Rate (ALR)**
 - For full-duplex Ethernet
 - Changes link rate in response to utilization
 - Consists of a *mechanism* and *control policy*
- **Mechanism**
 - MAC frame handshake
- **Control Policy**
 - Use queue length thresholds

The subject
of this talk

- **Queue length threshold policy needs analysis**

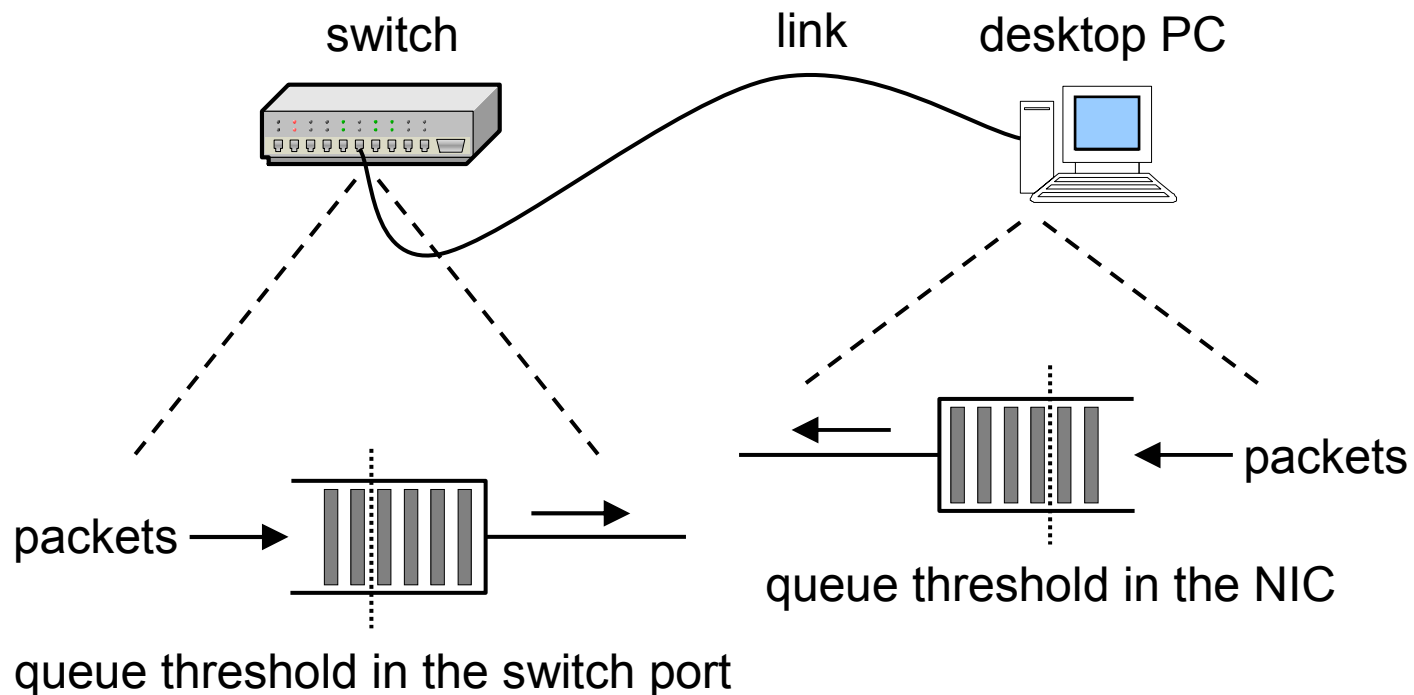
ALR mechanism

- **Mechanism implements link rate switching**
 - ALR must be implemented in both ends of an Ethernet link



ALR system view

- **ALR must be supported at both ends of a link**
 - Shown here is a single threshold policy
 - If queue is above threshold then switch to high rate
 - If queue is below threshold then switch to low rate



Agenda

- Introduction
- ALR threshold policy
- Analysis
- Summary and future work

Analysis approach

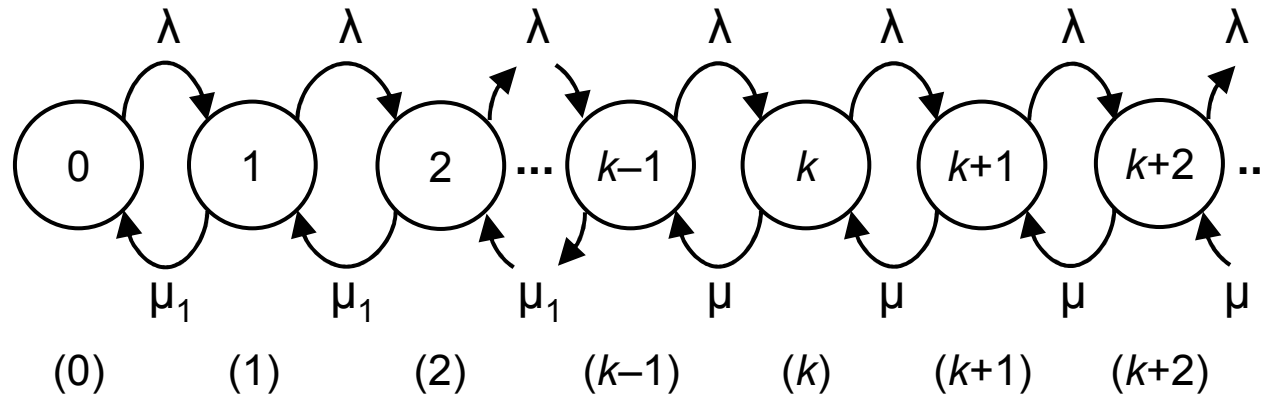
- **First, we built Markov models**
 - We assumed Poisson arrivals and exponential service times
- **Next, we used simulation to evaluate ALR**
 - Used actual traced Ethernet traffic
- **Control variables**
 - Threshold values
 - Utilization
 - Time to switch between link rates
- **Response variables**
 - Mean response time (mean packet delay)
 - Percentage of time in low data rate (energy saving)
 - Number of rate switches

Markov models

- **We model the following...**
 - One or two thresholds
 - Rate switch *during service* or at *completion of service*
 - Rate switching time is zero
- **Switch at completion of service models packets**
 - Rate cannot switch until current packet is fully transmitted
- **Four Markov models**
 - #1: Single threshold, rate switch during service
 - #2: Dual threshold, rate switch during service
 - #3: Single threshold, rate switch at completion of service
 - #4: Dual threshold, rate switch at completion of service

Markov model #1

- Single threshold, rate switch during service



(y) ← Number of customers in system in this state(s)

Markov model #1 continued

- **Solution is well known (e.g., Gross and Harris)**
 - We solve for mean number in system (L)
 - Percentage time in low rate is sum of probabilities of states below threshold (k)

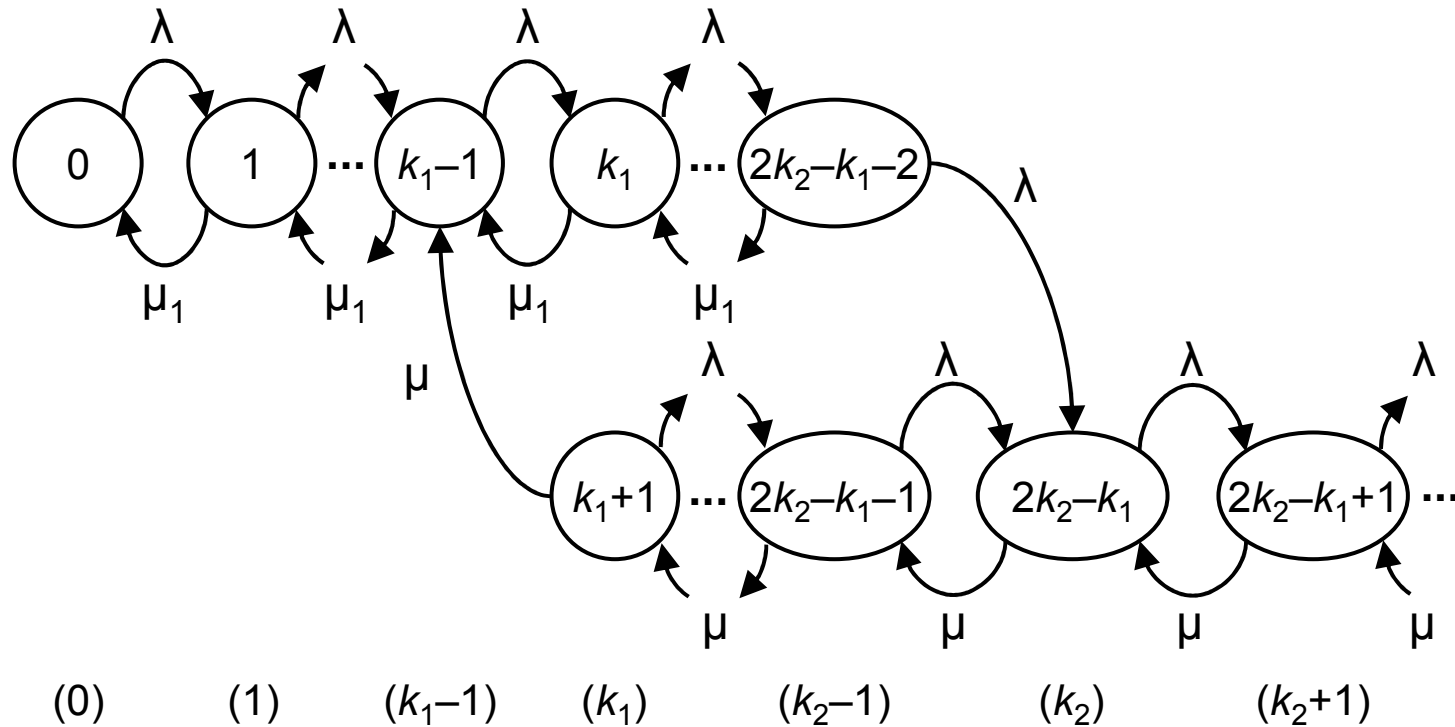
$$P_n = \begin{cases} P_0 \rho_1^n & (0 \leq n < k) \\ P_0 \rho_1^{k-1} \rho^{n-k+1} & (n \geq k) \end{cases}$$

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

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

Markov model #2

- Dual threshold, rate switch during service



Markov model #2 continued

- **Solution is straight-forward (done in [5])**
 - Mean number in system (L) is solved numerically
 - Closed form expression is very long

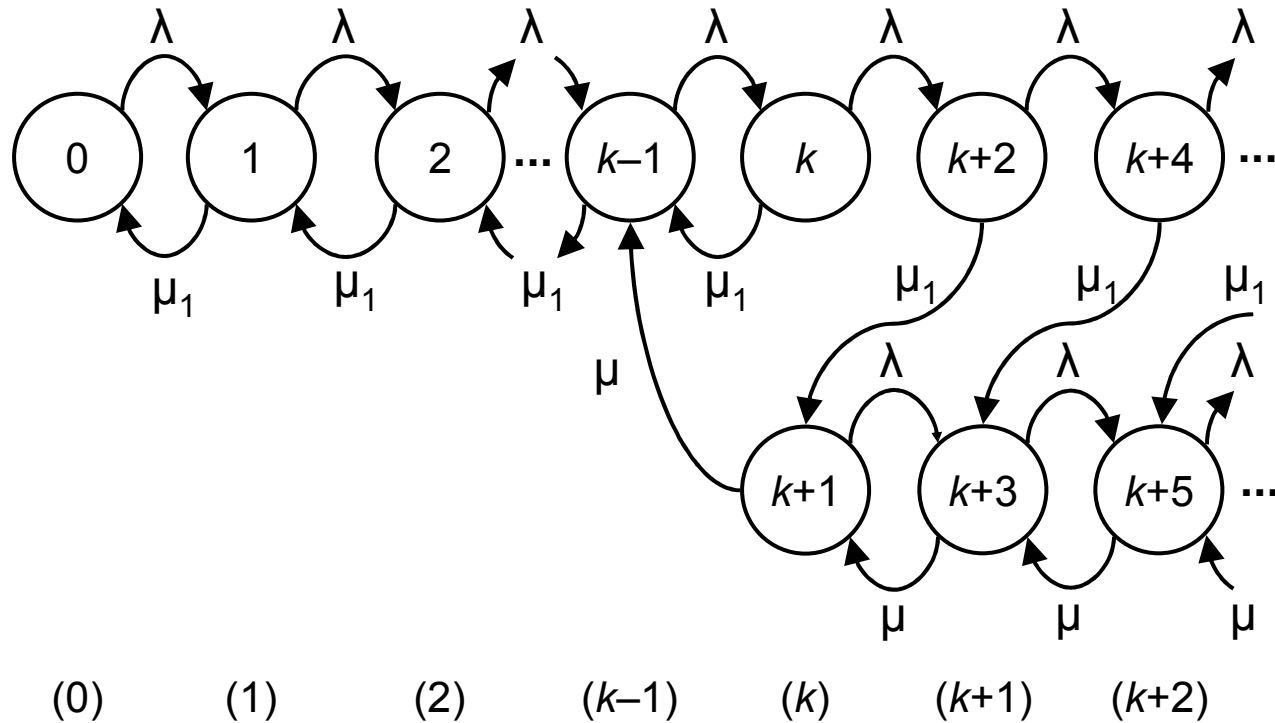
$$P_n = \begin{cases} P_0 \rho_1^n & (0 \leq n < k_1) \\ \frac{P_0 \rho_1^{k_1-1} (1 - \rho_1) \left(\frac{\rho_1^{n-k_1+1} - \rho_1^{k_2-k_1+1}}{1 - \rho_1} + \frac{\rho_1^{k_2-k_1} \rho (1 - \rho^{n-k_1+1})}{1 - \rho} \right)}{(1 - \rho_1^{k_2-k_1+1})} & (k_1 \leq n < k_2) \\ \frac{P_0 \rho_1^{k_2-1} \rho^{n-k_2+1} (1 - \rho_1) (1 - \rho^{k_2-k_1+1})}{(1 - \rho) (1 - \rho_1^{k_2-k_1+1})} & (n \geq k_2) \end{cases}$$

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

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

Markov model #3

- Single threshold, rate switch at completion of service



Markov model #3 continued

- **Direct solution is novel (in [2] done via transforms)**
 - Mean number in system (L) is solved numerically
 - Closed form expression is very long

$$P_n = \begin{cases} P_0 \rho_1^n & (0 \leq n < k) \\ \frac{P_0 \rho_1^{k-1}}{\rho + \rho/\rho_1 - 1} \left(\rho^{n-k+2} - \left(1 - \frac{\rho}{\rho_1}\right) \left(\frac{\rho_1}{1 + \rho_1}\right)^{n-k+1} \right) & (n \geq k) \end{cases}$$

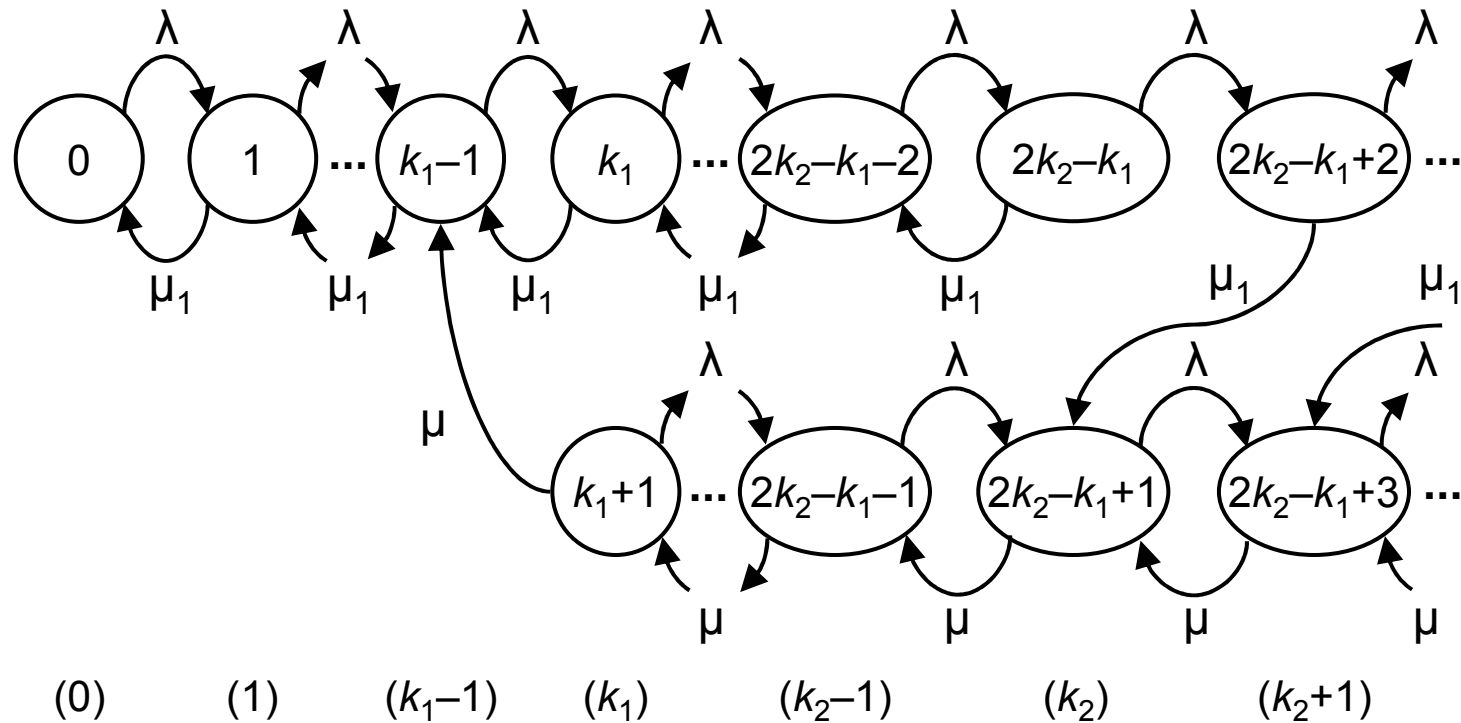
$$P_0 = \left(\frac{1 - \rho_1^k}{1 - \rho_1} + \frac{\rho_1^k}{1 - \rho} \right)^{-1}$$

Derivation
is in paper

[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.

Markov model #4

- Dual threshold, rate switch at completion of service



Markov model #4 continued

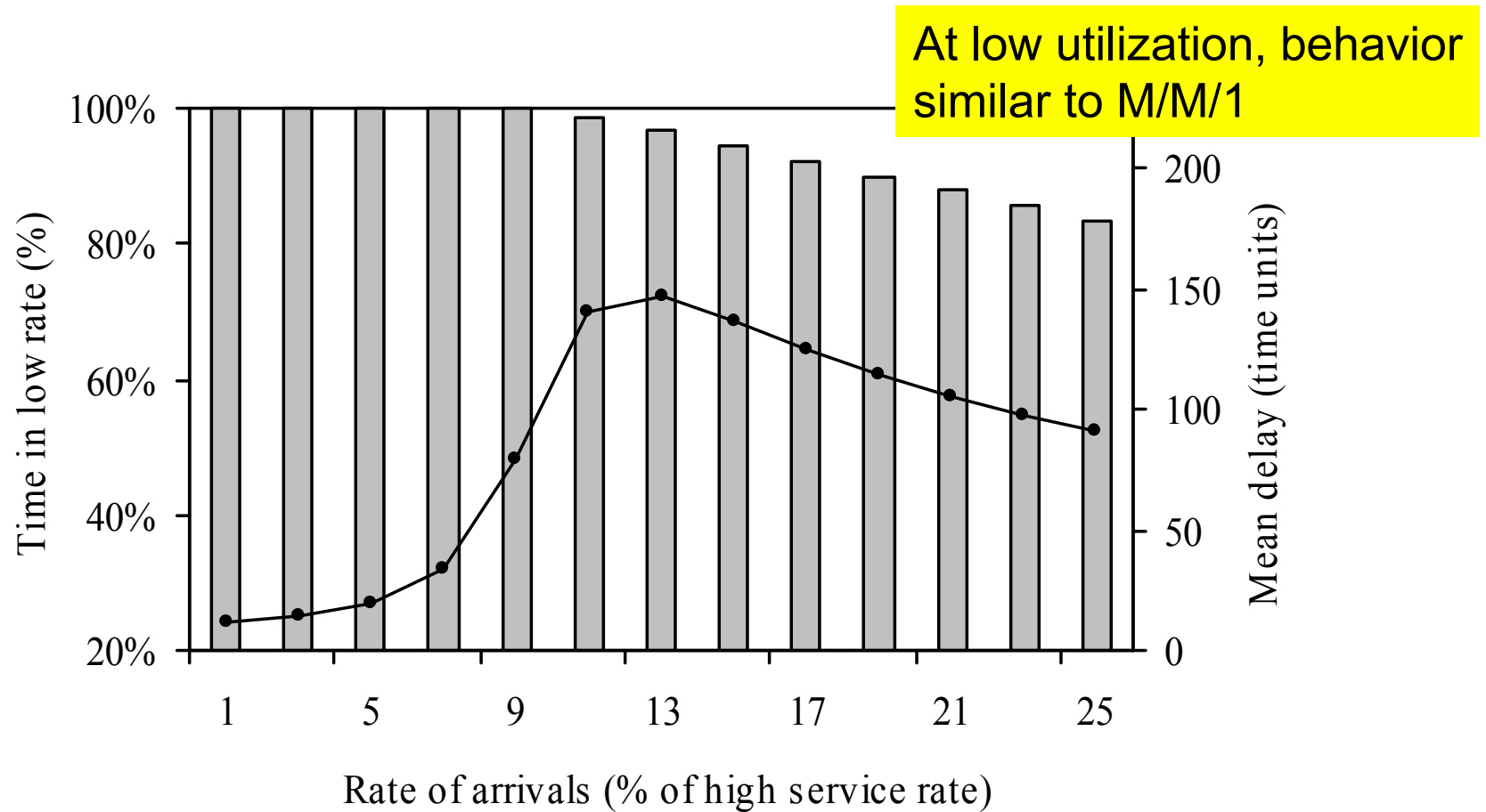
- **Direct solution is novel**
 - Mean number in system (L) is solved numerically
 - Closed form expression is very long

$$P_n = \begin{cases} P_0 \rho_1^n & (0 \leq n < k_1) \\ \frac{P_0 \rho_1^{k_1-1} (1-1/\rho_1) \left(\frac{1-1/\rho_1^{k_2-n+1}}{1-1/\rho_1} + \frac{\rho(1-\rho^{n-k_1+1})}{1-\rho} \right)}{1-1/\rho_1^{k_2-k_1+2}} & (k_1 \leq n < k_2) \\ \frac{P_0 \rho_1^{k_1-1}}{1-1/\rho_1^{k_2-k_1+2}} \left(\rho^{n-k_2+1} \left(\frac{(1-1/\rho_1)(1-\rho^{k_2-k_1})}{1-\rho} - \frac{1-1/\rho_1^2}{1-\rho-\rho/\rho_1} \right) + \frac{(1-1/\rho_1^2)(1-\rho/\rho_1)}{1-\rho-\rho/\rho_1} \left(\frac{\rho_1}{1+\rho_1} \right)^{n-k_2+1} \right) & (n \geq k_2) \end{cases}$$

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

Results from Markov model #4

- Results for switching time of zero

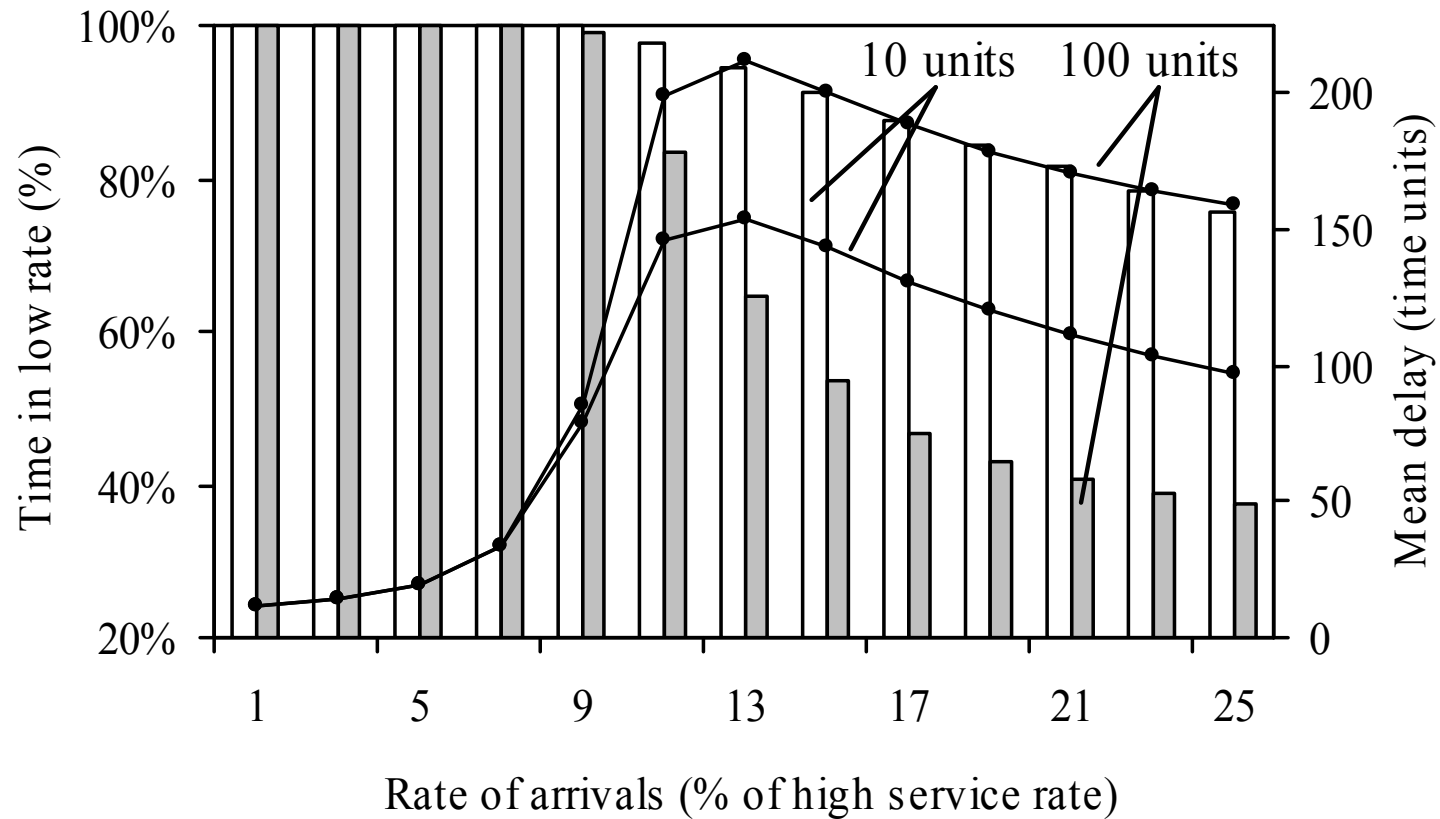


Performance evaluation

- **Evaluate ALR with a simulation model**
 - Built a queueing model using CSIM19
 - Can use Poisson and trace traffic inputs
 - **Non-zero rate switching time**
- **Evaluation with Poisson traffic**
- **Evaluation with trace traffic**

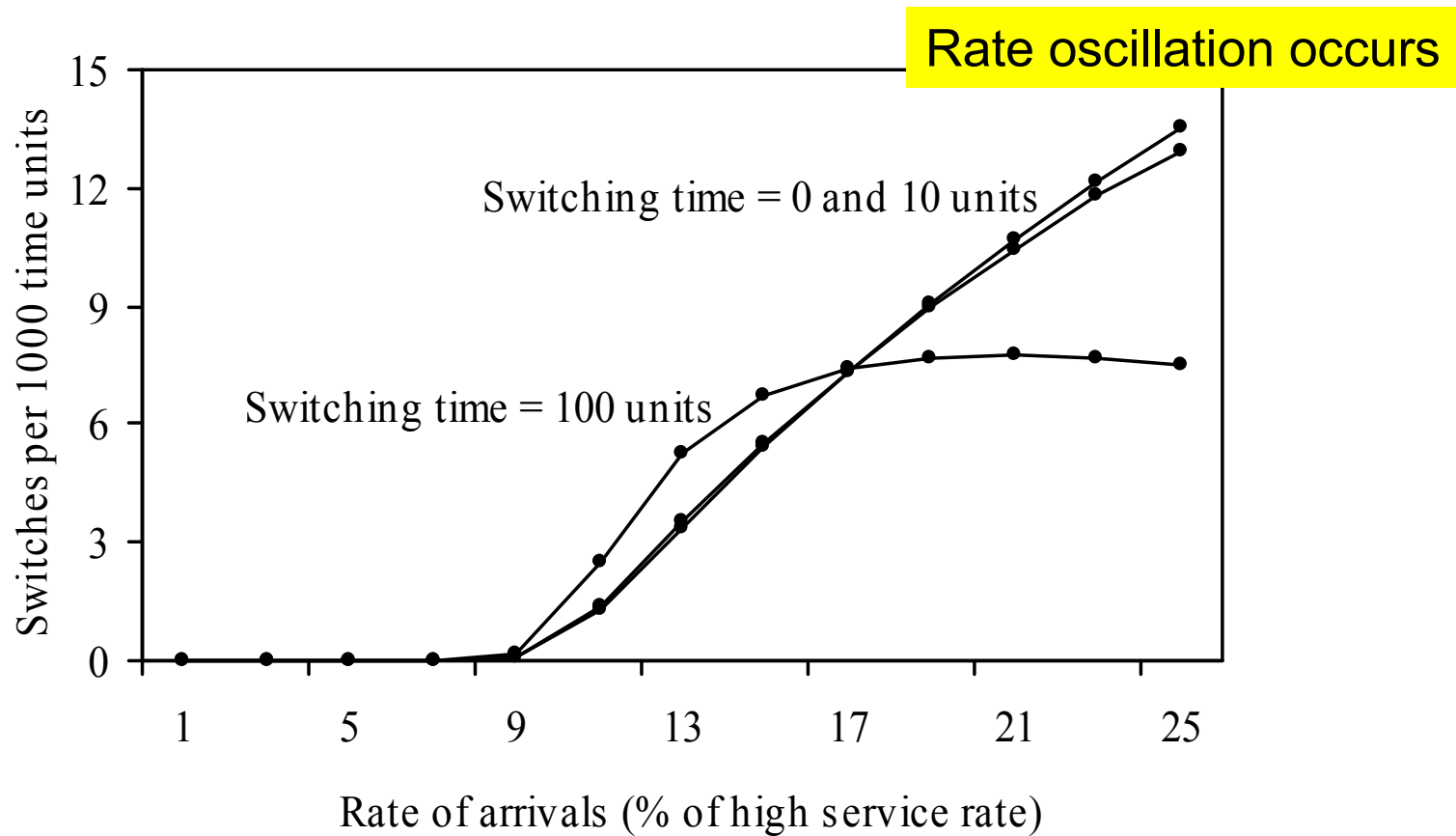
Results using Poisson traffic

- **Results for non-zero switching time**
 - “time unit” = mean service time at high rate



Results using Poisson traffic continued

- Results for non-zero switch time



Trace traffic characteristics

- **Summary of traces**
 - Note very low average utilization

Trace	Duration	Description	Avg util
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

Results using trace traffic

- **Results for all traces without and with ALR**
 - Showing mean packet delay

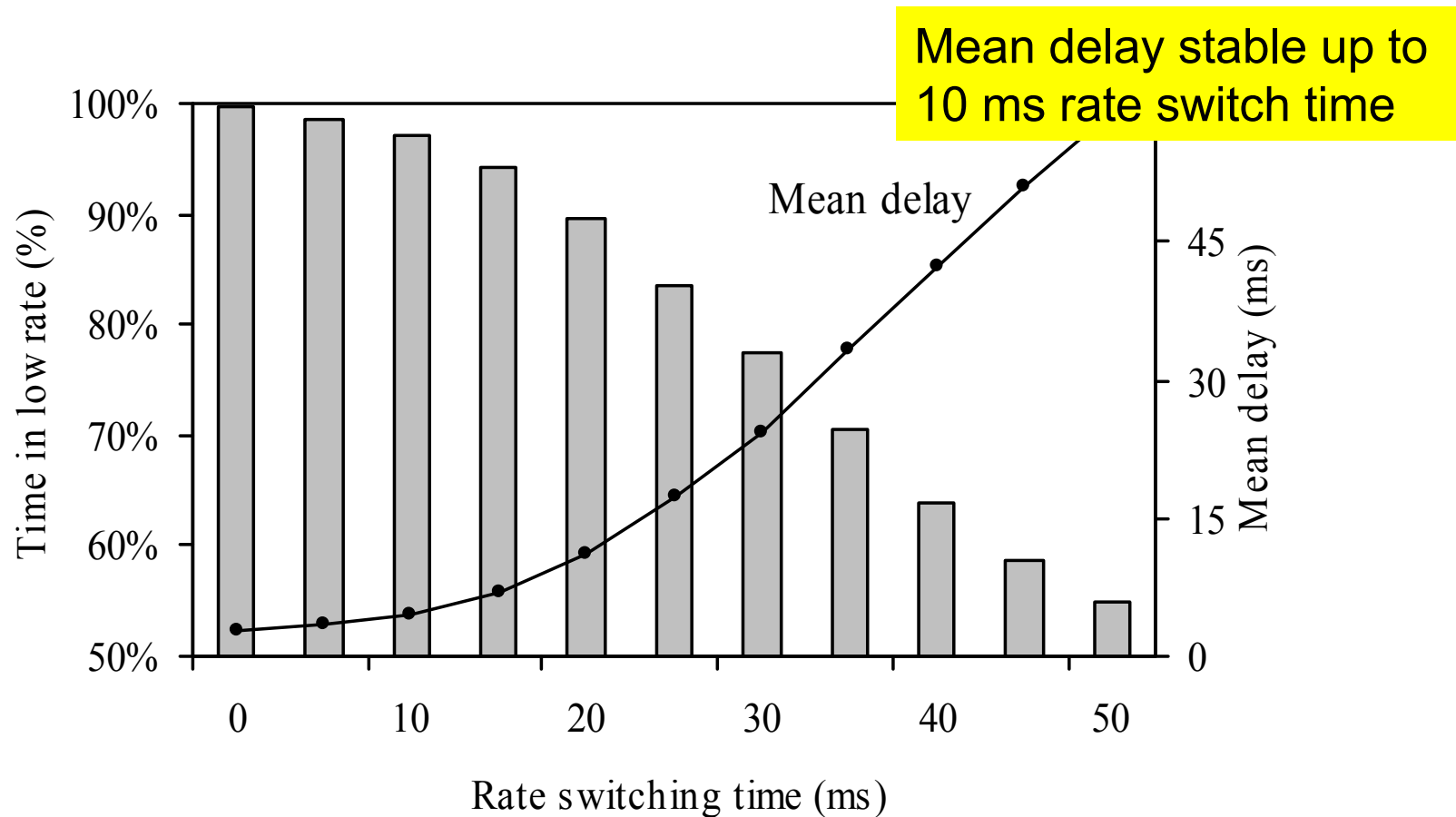
Trace	10 Mb/s	100 Mb/s	ALR	
			Mean Packet Delay	ALR %
USF #1	7.60 ms	0.09 ms	2.79 ms	99.42 %
USF #2	3.95	0.08	1.81	99.81
USF #3	196.29	0.05	1.48	99.99
PSU #1 ²	33.51	0.18	5.63	99.98
PSU #2 ²	2321.31	0.12	9.55	99.12
PSU #3 ²	1147.83	0.51	4.07	99.83

Time in 10 Mb/s



Results using trace traffic continued

- Time in low rate and mean delay for USF #1



Agenda

- Introduction
- ALR threshold policy
- Analysis
- Summary and future work

Summary

- **Queue threshold policy for ALR investigated**
 - At 5% utilization, almost all the time is in low rate
- **Policy tested with real 100 Mb/s trace traffic**
 - 99% of the time in 10 Mb/s low data rate (energy saving)
- **Energy savings of hundred of \$millions possible**
 - See paper for discussion

Future work

- **Need to investigate link rate oscillations**
 - Policy must be stable under all conditions
- **Explore context specific clues for switching rate**
- **Improved traffic models for 10 Gb/s**
- **ALR is moving towards standardization in 802.3**
 - CFI approved at November 14th IEEE 802.3 meeting in Dallas
 - 55 Yes, 2 No, 7 Abstain

Thank you!

I'll be happy to answer any questions 😊

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

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