

PAPER

Providing Consistent Delay Differentiation

Ikjun YEOM[†], *Member* and A.L.Narasimha REDDY^{††}, *Nonmember*

SUMMARY *Class-based delay differentiation* model has been recently proposed as a part of relative differentiated services frameworks, and it is shown that the model can provide delay differentiation without admission control and end-to-end resource reservation. In this paper, however, we observe that there can be inconsistent delay differentiation caused by different size of packets. We propose *packet size-based delay differentiation* model and show that packet size-based queueing is effective to achieve equal delay within a class and provide consistent delay differentiation between classes through simulations. Simulation results also show that the proposed model improves jitter characteristics of CBR flows.

key words: *Differentiated service, Relative Delay Differentiation, Quality of service*

1. Introduction

Differentiated Services (diff-serv, DS) architecture has been proposed to provide different levels of services to different users [1], [2]. It differs from the integrated services [3], [4] in the sense that core routers in DS domain do not need to maintain per-flow state to provide service differentiation. Currently, two types of service are defined. Assured Forwarding service was motivated from [5] to provide *better service* than best-effort [6]. When a packet arrives at AF service domain, the packet is classified and marked according to its service profile. Inside of the domain, each router treats the packet differently by its marking. Expedited Forwarding (EF) service, also called *Premium Service*, has been developed to provide guaranteed service (e.g., low loss rate, low latency and low jitter) [7].

Recently, a new service model, called *relative differentiated services*, has been proposed. In this service model, each packet is classified and marked by its subscribed class of service. Core routers provide better treatment to higher classes in terms of per-hop delay and/or loss rate. The main advantage of this model is to provide no worse guarantee (i.e., higher classes are guaranteed to get better service or at least no worse than lower classes.) without admission control and end-to-end resource reservation.

In [8], a proportional delay differentiation model has been proposed. In this model, a packet subscribed to a higher class of service experiences less queueing delay than a packet subscribed to a lower class of service. It has been shown that the model can adjust level of delay differentiation between classes independent of class loads. This model is suitable to network applications which is sensitive to delay rather than loss rate such as Internet telephony and video/audio streaming. Two packet schedulers, Backlog Proportional Rate (BPR) and Waiting Time Priority (WTP) schedulers, have been used to evaluate the model. In the evaluation, bursty traffic with identical packet size distribution in each class has been used. The results show that queueing delay observed by each class of traffic can be controlled in short-term as well as long-term timescales by the model. We call this model *class-based delay differentiation (cbd²)*.

In this paper, we study the end-to-end delay characteristics (average queueing delay, jitter and buffer requirement) of individual CBR (constant bit rate) flows through relative differentiated services network. Our study shows that flows within the same class may get different delays by their packet size in a class-based delay differentiation network. In a low utilized network, no worse guarantee between individual flows can be broken even though the guarantee between averages of the classes still holds.

Our study in Section 2 shows that packet sizes need to be taken into account in order to provide consistent delay differentiation. We propose to do this through *packet size-based delay differentiation (psd²)*. This model classifies packets by their size, and gives higher priority to smaller packets. Consequently, smaller packets get less delay than larger packets. Advantages of the psd² model are (1) removing inconsistent delay differentiation between classes caused by different packet size; and (2) useful to deliver real-time traffic since most real-time applications tend to send small packets, and also flows sending larger packets may be able to tolerate more delay and jitter.

The proposed model requires that flows requiring smaller delays will have to utilize smaller packets, and flows subscribing for larger delays will have to send data in larger packets. For a given sending rate, this tradeoff is quite natural since a flow with smaller packets has smaller inter packet spacing. Applications can tailor

Manuscript received , 2001.

Manuscript revised , 2001.

[†]The author is with the Department of Computer Science at Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea.

^{††}The author is with the Department of Electrical Engineering at Texas A&M University, USA.

their packet sizes based on the level of delay required. On the other hand, edge routers can be desinged to fragment or aggregate packets to conform to the size requirements of the subscribed delay class.

To implement our model, we use the WTP scheduler with an estimated waiting time. Our simulation study shows that the proposed model can provide consistent delay differentiation between individual flows as well as between averages of the classes in various networks.

The rest of this paper is organized as follows: In Section 2, we study per-flow end-to-end delay characteristics in a cbd^2 network with CBR flows and address the motivation of this paper. In Section 3, we propose the psd^2 model and describe implementation details. In Section 4, we present a number of simulations to evaluate our service model. We summarize our study in Section 5.

2. Class-based delay differentiation

In this section, we study the impact of heterogeneity of packet size in the cbd^2 network. It has already been shown that the cbd^2 is effective to provide delay differentiation among a set of classes when packet size distribution is identical to each class and each flow [8], [9].

Does this delay differentiation between classes still hold when packet size of each individual flow is different? Consider a simple network: There are two flows, F_1 and F_2 , with arrival rates, r_1 and r_2 , within the same class. Packet sizes of F_1 and F_2 are k_1 and k_2 respectively. We assume that there is no other traffic for simplicity, and FCFS (first-come first-serve) queue is used for each class. Then, from [10], the worst-case queueing delay of F_1 is bounded by \bar{D}_1 , where

$$\bar{D}_1 \leq \frac{1}{C_{out}} \max_{u \geq 0} [b_1(u) + b_2(u) - C_{out}u + k_2] \quad (1)$$

where $b_n(u) = \int_0^u r_n(t)dt$, and C_{out} is the departure rate. k_2 is for considering the situation in which a packet from F_2 arrived just before a packet from F_1 and has not yet queued completely. Since $b_2(u)$ includes only packets from F_2 received completely, k_2 is not included the term $b_2(u)$. In a work conserving server, C_{out} is given by $\min[r_1 + r_2, C]$, where C is the maximum service rate. (1) shows that a packet in F_1 exits the queue after waiting for transmission of remaining packets ($b_1(u) + b_2(u) - C_{out}u$) and also after waiting for a packet in F_2 which arrived just before the packet. Similarly, \bar{D}_2 , the upper bound of delay of F_2 is given by

$$\bar{D}_2 \leq \frac{1}{C_{out}} \max_{u \geq 0} [b_1(u) + b_2(u) - C_{out}u + k_1] \quad (2)$$

In a highly utilized network ($\bar{r}_1 + \bar{r}_2 > C$, \bar{r}_n is the mean of r_n), the queue length keeps increasing, and

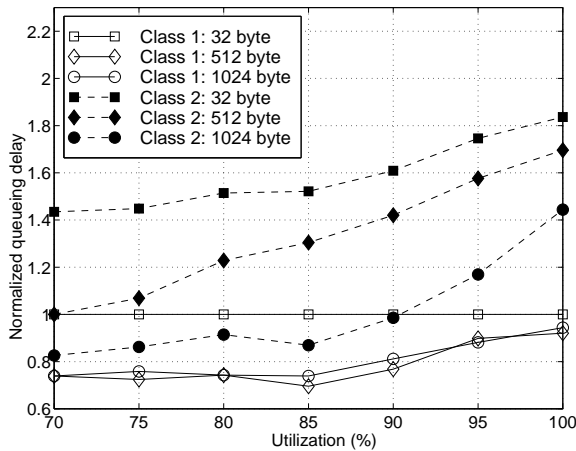
thus \bar{D}_1 and \bar{D}_2 also increase. Consequently, the queue builds up and drops packets. Drop policy is beyond the scope of this paper, and we only consider cases in which $\bar{r}_1 + \bar{r}_2 \leq C$. Similar assumption is also made in [8], [9]. When $\bar{r}_1 + \bar{r}_2 \ll C$, $(b_1(u) + b_2(u) - C_{out}u) \approx 0$ even if we take bursty traffic into consideration. Then, \bar{D}_1 and \bar{D}_2 are determined by k_2/C_{out} and k_1/C_{out} , respectively. If $k_1 \gg k_2$, $\bar{D}_1 < \bar{D}_2$, and vice versa. As $\bar{r}_1 + \bar{r}_2$ increases, $(b_1(u) + b_2(u) - C_{out}u)$ may also increase. Then, the impact of k_1 and k_2 is reduced.

Of course, (1) and (2) are too simple to reflect the whole network operation. However, it is obvious that the queueing delay in a low utilized network is highly impacted by packet sizes. This may result in inconsistent delay differentiation between individual flows. That is, a flow with small packets within class i gets larger delay than a flow with large packets within class $i + 1$ while the average delay of whole class i is still less than the average delay of whole class $i + 1$. To observe this impact in detail, we conduct a set of simulations with ns-2 [11]. We implemented two types of delay-sensitive priority schedulers, WTP and MDP (Mean Delay Priority) [9] for the cbd^2 .[†]

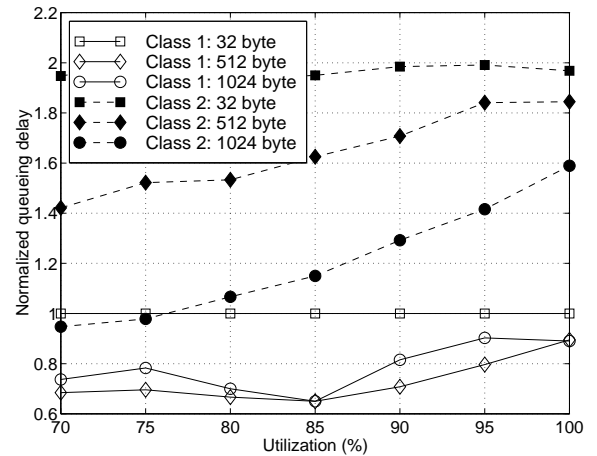
Fig. 1 shows the simulation topology. This multiph topology has been used to study delay characteristics [8],[12],[13]. We simulate two classes of delay differentiation. We set weights for each class so that delay of class 2 is expected to be two times the delay of class 1. Tagged traffic consists of three CBR flows within each class. The packet sizes of the three flows are 32, 512 and 1024 bytes, and the sending rates are the same as 1 Mbps. 90 exponential on/off sources are used as cross traffic. 45 sources subscribe to class 1, and the rest subscribe to class 2. 15 sources within each class send 32, 512 and 1024 byte packets. The average sending rate of each source is 1 Mbps, and busy and idle periods are randomly selected. Hence, load of each class is the same as 48 Mbps, and the packet size distribution of each class is also the same. In each simulation, link capacity between R_n and R_{n+1} , where $n = 0, \dots, 18$, is changed from 137 to 96 Mbps so that utilization is changed from 70% to 100%.

Fig. 2 shows the results. We take the average of queueing delays of each flow over 20 hops and normalize the average delays by the average delay of the flow with 32 byte packet within class 1. Ideally, it is expected that the flows within class 1 stay at 1, and the flows within class 2 stay at 2. However, simulation results show that (1) Within a class, flows with larger size packets get

[†]To implement our schedulers, we follow [8], [9], [14]. In the WTP scheduler, the priority, q , at time t of a packet (within class i) which arrives at time τ is calculated from $q(t) = (t - \tau)b_i$ where b_i is the weight of class i . In the MDP scheduler, average delay of each class is used to calculate q instead of individual delay $(t - \tau)$. Whenever the server is ready for a new packet, it selects the packet with the highest priority among packets which are at the head of each queue.



(a) WTP scheduler



(b) MDP scheduler

Fig. 2 Normalized queuing delay with different utilizations

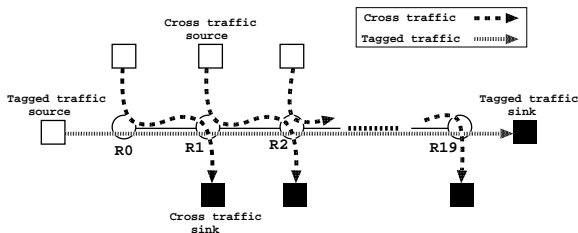


Fig. 1 A multihop topology

lower delays in both WTP and MDP schedulers. (2) Delay differentiation caused by different packet sizes increases as link utilization decreases. (3) No worse guarantee between individual flows is broken when link utilization is less than 90% in the WTP scheduler and less than 75% in the MDP scheduler. That is, the flow with 1024 byte packets within class 2 gets lower delay than the flow with 32 byte packets within class 1.

3. Packet size-based delay differentiation

In the previous section, we have studied that the cbd^2 may fail to provide no worse guarantee in a low utilized network due to heterogeneity of packet sizes. To prevent out-of-order service differentiation and provide consistent delay differentiation, we propose a new service model, *packet size-based delay differentiation* (psd^2).

The basic idea of the psd^2 model is to classify packets by their size and provide queuing delay to each packet proportional to its size. This is a simple variant of the cbd^2 but has the following significant attributes: (1) It provides fairly equal delay to individual flows within a class since packet size distribution of individual flows within a class is similar to each other. (2) No

worse guarantee can be protected between individual flows as well as between classes.

It also has merit in scheduling CBR flows regarding to delay jitter. In the psd^2 model, delay jitter is expected to be proportional to the packet size since queuing delay is proportional to the packet size. We claim that a CBR flow with large packets can tolerate large jitter from two metrics: (1) Normalized jitter[†]: Consider two CBR flows with the same sending rate and different packet sizes. Normalized jitter is calculated by $r\delta/k$ where r is the sending rate, δ is absolute jitter^{††}, and k is the packet size. Thus, a flow with large packets can tolerate large δ . (2) Buildout buffer: Currently, most multimedia data is compressed and decompressed in units of packets to be delivered and played back. Thus, we estimate that buildout buffer requirement (B) should be larger than at least one packet size and has to be a multiple of packet size.

$$B = \lceil \frac{c_{max}(d_{max} - d_{min})}{k} \rceil \cdot k \quad (3)$$

where c_{max} is the peak arrival rate at the receiver. It is clear that an application sending large packets can tolerate large δ without increasing B .

To implement the psd^2 model, we consider WTP and MDP schedulers as described in the previous section. The difference between the two schedulers is to measure waiting time (or queuing delay). In the WTP scheduler, the waiting time of a packet is the difference between its arrival time and the current time. The waiting time reflects the instantaneous queuing delay. Thus, it is possible to provide relative delay differenti-

[†]In this paper, normalized jitter means jitter normalized by inter-packet spacing.

^{††}Delay-jitter is usually defined as $(d_{max} - d_{min})$, where d_{max} is the maximum delay, and d_{min} is the minimum delay.

ation in short timescales while the delay differentiation may fail in long timescales. In the MDP scheduler, the waiting time of a packet is the aggregated delay observed by all the previous and current packets in a queue (or class). In [9], it has been shown that the MDP scheduler provides more accurate delay differentiation of long-term average delay. However, it may not be effective in adapting quickly changing network dynamics.

We do not attempt to argue that short-term or long-term delay differentiation should be achieved in the relative differentiated services network. It can be discussed and decided between users and network providers. Instead, in this paper, we provide a more flexible way to measure the waiting time using a TCP-like method. The waiting time (\bar{d}) of a packet which arrives at τ is calculated by

$$\bar{d} = \alpha \bar{d} + (1 - \alpha)(t - \tau) \quad (4)$$

where t is the current time, and $0 \leq \alpha < 1$. When $\alpha = 0$, \bar{d} reflects individual queueing delay of a packet. With large α , large number of delays are accumulated to \bar{d} . Therefore, we can control timescale of delay differentiation by adjusting α .

There are two possible ways to subscribe to a certain class of service. First, an application itself can adjust its packet size so that the packets observe the class (or queue) which it wants. When the application cannot adjust its packet size, edge devices can fragment or aggregate packets so that they can observe the subscribed class of service. In either case, note that load at core routers does not increase.

4. Simulation study

In this section, we present simulations to evaluate the psd² model using ns-2 [11]. The objectives of simulations are to show that the psd² provides consistent delay differentiation and improves jitter characteristics.

4.1 Simulation setup

We implemented a WTP scheduler with 6 queues (or classes). Queue i is associated with k_i which is the upper limit of packet size and b_i which is the weight of that queue. In [8], it is shown that a WTP scheduler effectively provides delay differentiation based on weight, $\{b_i\}$, to each class. In this paper we set $\{b_i\}$ to $\{2^0, 2^{-1}, 2^{-2}, 2^{-3}, 2^{-4}, 2^{-5}\}$ for both the psd² and cbd² simulations. $\{k_i\}$ can be determined based on packet size distribution or requested class distribution by customers. Here we set to $\{2^6, 2^7, 2^8, 2^9, 2^{10}, \text{MTU}\}$ in order to observe performance of the psd² over wide range of packet size. The priority q_i of a packet in Queue i is calculated from $q_i = \bar{d}_i b_i$, where \bar{d}_i is given by (4). We set α to 0.99.[†]

[†]We have simulated with various α from 0 to 0.99.

4.2 Providing consistent delay differentiation

In this simulation, we evaluate the psd² model for providing (1) equal service (or delay) within a class and (2) delay differentiation between classes. We also present the results of the cbd² model for comparison. In the simulation with the cbd² model, the same scheduler as described in Section 4.1 is used except that packets are queued by their subscribed classes instead of their sizes.

Simulation topology is the same as in Fig. 1. For the tagged traffic, we use 36 CBR flows with 1 Mbps sending rate each. Each 6 flows subscribes each class. For the cross traffic, we use 90 exponential on/off sources with 1 Mbps sending rate each. 15 flows subscribe to each class. Therefore, traffic load of each class is the same as 21 Mbps. In the simulation for the cbd² model, the packet size is equally distributed among $\{32, 96, 192, 384, 768, 1536\}$ bytes for each class. In the psd² model, note that packets are classified by their sizes. As consequence, the simulation setups including topology, tagged/cross traffic, traffic load of each class, and the number of classes for the psd² and cbd² models are the exactly same except that traffic in psd² simulation is classified by packet size while traffic in cbd² is classified by its subscribed class. We set a link capacity so that the utilization is 70, 85 and 90% for each simulation.

Fig. 3 and 4 show queueing delays of the tagged traffic normalized by the flow with 32 bytes in class 1. It is observed that (1) As the utilization increases, the delay differentiation becomes clear in both models. (2) In the cbd² model, flows with large packets get less delay. In 70 and 85% utilized networks, inconsistent delay differentiations are observed. (3) The psd² model provides consistent delay differentiation even in 70% utilized network.

4.3 Improving jitter characteristics

We compare the psd² with FIFO scheduling using the following metrics: (1) Buildout buffer requirement calculated from (3). (2) 99% delay-jitter normalized by inter-packet spacing. (3) 99% rate-jitter^{††} normalized by inter-packer spacing. All those metrics are widely used to measure performance of real-time applications or networks delivering real-time traffic. For example, buildout buffer requirement is important performance index for video streaming service, and rate and delay jitter bound is important to provide seamless conversations over Internet telephony. A simulation for FIFO

Results show that our model can provide either long or short timescale delay differentiation. In this paper, we only present long-term delay differentiation due to the space limitation.

^{††}In this paper, we define rate-jitter as the difference of inter-arrival times of two consecutive packets at a receiver.

scheduling has been done with the same simulation setup as in the previous section except using FIFO instead of the psd².

Fig. 5 and 6 show the results. It is observed that (1) For flows that cross 10 or more hops in the network, the psd² requires the same or less buffer than FIFO. For flows going over smaller number of hops, the psd² may require more buffer than FIFO. (2) Both delay and rate jitters of 32 byte packet are drastically reduced by the psd² with small increases for large packets. Overall, the delay and rate jitters have better characteristics than in FIFO. This results support that the psd² model is effective to deliver real-time traffic. Our results (not presented here due to space constraints) show that the psd² provides consistent delay differentiation across the classes even when the load across the classes is varying.

5. Summary

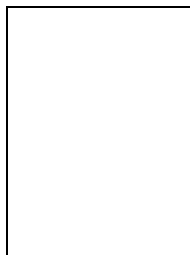
In this paper, we have observed that the class-based delay differentiation is not enough to provide consistent delay differentiation due to heterogeneity of packet sizes. The impact of heterogeneity of packet size increases as a network utilization decreases. To provide consistent delay differentiation, we have proposed the packet size-based delay differentiation model and evaluated it through simulations. The simulation results clearly show that the proposed model has the following attributes: (1) achieving equal delay within a class, (2) providing consistent delay differentiations between classes, and (3) improving delay/rate jitter in scheduling CBR flows.

References

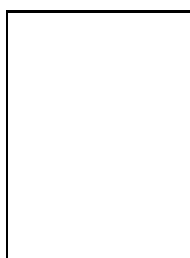
- [1] K. Nichols, S. Blake, F. Baker and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," RFC2474, Network Working Group, December, 1998.
- [2] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, "An Architecture for Differentiated Services," RFC2475, Network Working Group, December, 1998.
- [3] B. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview," RFC1633, Network Working Group, June, 1994.
- [4] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, "RSVP: A New Resource ReSerVation Protocol," *IEEE Transaction on Networking*, September, 1993.
- [5] D. D. Clark and W. Fang, "Explicit Allocation of Best-Effort Packet Delivery Service," *IEEE/ACM Transactions on Networking*, Vol. 6, August, 1998.
- [6] J. Heinanen, F. Baker, W. Weiss and J. Wroclawski, "Assured Forwarding PHB Group," June, 1999.
- [7] V. Jacobson, K. Nichols and K. Poduri, "An Expedited Forwarding PHB," RFC2598, Network Working Group, June, 1999.
- [8] C. Dovrolis, D. Stiliadis and P. Ramanathan, "Proportional Differentiated Services: Delay Differentiation and Packet Scheduling," *Proceedings of SIGCOMM*, August, 1999.
- [9] T. Nandagopal, N. Venkitaraman, R. Sivakumar and V. Bharghavan, "Delay Differentiation and Adaptation

in Core Stateless Networks," *Proceedings of INFOCOM*, March, 2000.

- [10] R. L. Cruz, "A Calculus for Network Delay, Part I: Network Elements in Isolation," *IEEE Transactions on Information Theory*, vol. 37, no. 1, January 1991.
- [11] Network simulator v.2 (ns-2), University of California at Berkeley, CA, 1997. Available via <http://www-nrg.ee.lbl.gov/ns-2>.
- [12] M. Grossglauser and S. Keshav, "On CBR Service," *Proceedings of INFOCOM*, March, 1996.
- [13] J. Sahni, P. Goyal and H. M. Vin, "Scheduling CBR Flows: FIFO or Fair Queuing," *Proceedings of NOSSDAV*, June, 1999.
- [14] L. Kleinrock, *Queueing Systems, Vol. II*, John Wiley and Sons, 1976.



Ikjun Yeom received the B.S. degree in Electronic Engineering from Yonsei University, Seoul, South Korea, in 1995, and the M.S. and Ph.D degree in Computer Engineering at Texas A&M University, College Station, TX in August 1998 and May 2001 respectively. He is currently an Assistant Professor in the department of Computer Science at Korea Advanced Institute of Science and Technology in Daejeon, South Korea. His research interests are in Internet QoS provisioning, the differentiated services network and Internet congestion control.



A. L. Narasimha Reddy received his B.Tech. degree in Electronics and Electrical Engineering from the Indian Institute of Technology, Kharagpur, India, in August 1985, and the M.S. and Ph.D degrees in Computer Engineering from the University of Illinois at Urbana-Champaign in May 1987 and August 1990 respectively. At Illinois, he was supported by an IBM Fellowship. He is currently an Associate Professor in the department of

Electrical Engineering at Texas A&M University. He was a Research Staff Member at IBM Almaden Research Center in San Jose from Aug. 1990- Aug. 1995. Reddy's research interests are in Multimedia, I/O systems, Network QOS and Computer Architecture. Reddy is a member of ACM SIGARCH and is a senior member of IEEE Computer Society. He has received an NSF CAREER award in 1996. He received an Outstanding Professor award at Texas A&M during 1997-98.

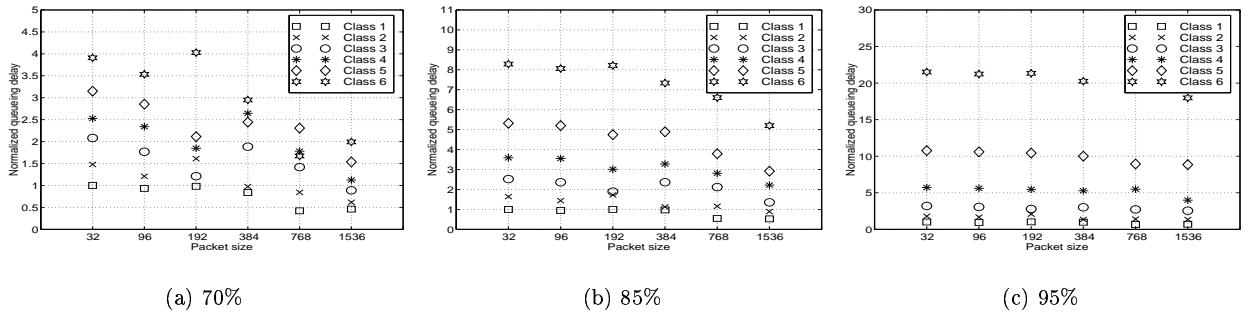


Fig. 3 Normalized queuing delay in class-based delay differentiation

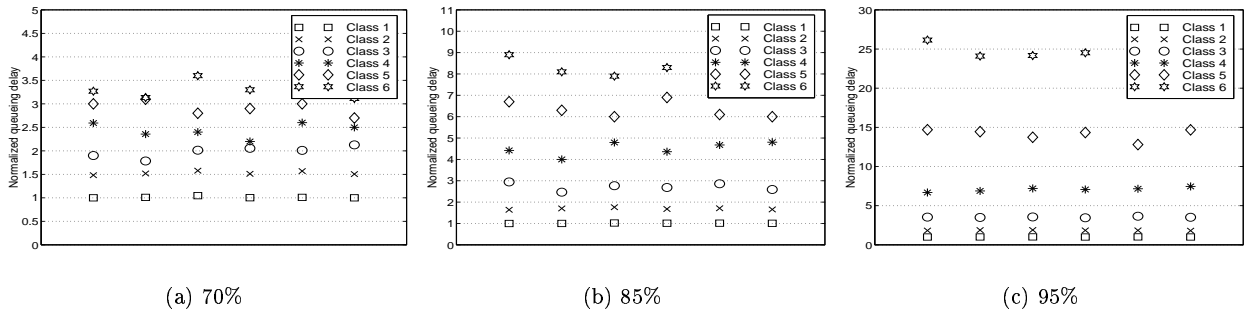


Fig. 4 Normalized queuing delay in packet size-based delay differentiation

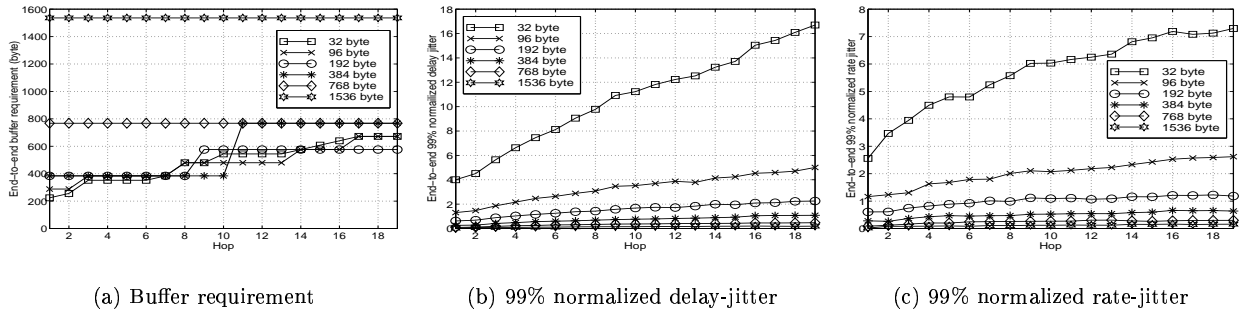


Fig. 5 Measurement of FIFO

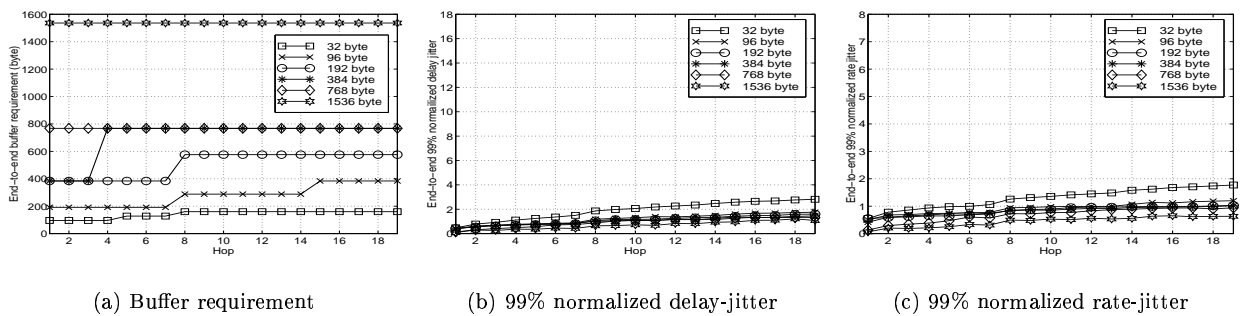


Fig. 6 Measurement of packet size-based delay differentiation