Adaptive Sampling for Network Performance Measurement Under Voice Traffic

Wenhong Ma and Changcheng Huang
Optical Networks Laboratory
Department of Systems and Computer Engineering,
Carleton University
1125 Colonel By Dr., Ottawa, ON K1S 5B6, Canada
{wenhongm, huang}@sce.carleton.ca

James Yan
Nortel Networks
PO Box 3511 Station C, Ottawa, ON K1Y 4H7, Canada
jimyan@nortelnetworks.com

***Abstract*—As the Internet grows in scale and complexity, the benefits of network performance measurements and monitoring are significantly increasing. Sampling-based measurement methods provide adequate techniques for reducing the quantity of control data and attract growing interests. The research reported in this paper, addresses the issue of how to carry out the sampling in an adaptive fashion, so that the accuracy for measuring the quality of service parameters (delay, loss, jitter, throughput) is better if we know something about the traffic type and traffic parameters. Our study proposes and investigates a mechanism that can be set up to adaptively adjust the parameters of the sampling technique. Two realistic network topologies based on MPLS networks are setup to evaluate the proposed adaptive sampling scheme for monitoring and measuring network performance metrics. Compared with conventional sampling techniques (systematic and stratified sampling), simulation results are presented to illustrate that adaptive sampling provides the potential for better monitoring, control, and management of high-performance networks with higher accuracy.**

Keywords-Adaptive sampling; Network performance metrics; MPLS; QoS.

I. INTRODUCTION

As we enter the 21st century, the Internet continues to experience tremendous growth. Providing Quality of Service (QoS) and traffic engineering capabilities in the Internet is essential, especially in supporting the requirements of real-time, as well as mission-critical applications, such as video and voice. Differentiated services and Multi-protocol Label Switching (MPLS) are emerging technologies, which play a key role in IP networks by delivering QoS and traffic engineering features. One of the key questions remaining to be answered is how applications (especially those sensitive to delay and loss) behave when these new protocols and network architectures are used. The proper monitoring and performance evaluation techniques have therefore to be developed in order to make network management effective and timely [1]. Network measurements and monitoring are needed for network design, capacity planning and forecasting, operations and management, and customer-driven activities. The framework for IP performance measurement has been proposed in [2]. There have been several research efforts on network performance measurements and analysis [3, 4, 5, 6].

The methods for measuring and monitoring network performance parameters usually fall into two categories: passive methods and active methods. Active measurement generates the controlled probe traffic and injects it into an ingress port of the network, and measures the received traffic at a receiving node. The method is becoming increasingly important due to its great flexibility, intrinsically end-to-end nature, and freedom from the need to access core network-switching elements. Thus active measurement methods are typically used to obtain end-to-end statistics such as latency, loss and route availability. Passive measurements are based on the actual payload traffic in the network. Unlike active methods, passive methods do not add extra traffic load to the network.

Besides this non-intrusive character, they provide a statement about the treatment of the current traffic in the observed network section. However, recording and logging of packet traces in high-speed networks often require special collection, storage and processing of very large amount of data. The method presented in this paper focuses on the application of adaptive sampling with active measurement.

Sampling-based measurement methods provide adequate techniques for reducing the quantity of measurement data and attract growing interests. A new working group on this topic PSAMP [7] was just formed in IETF, the Internet standards forum. Sampling techniques are used to study the behavior of a population of elements based on a representative subset. Cochran [8] and Krishnaiah and Rao [9] introduce the basic concepts and process of sampling algorithms. In packet networks, performance parameters are computed by way of choosing some particular packets among those crossing the network and can be obtained from them to satisfy policy or evaluation functions. There are three conventional sampling methods (systematic, random and stratified sampling) employed by network management systems and measurements. Sampling has been applied to measurements for different purposes [10, 11, 12, 13]. Systematic sampling is based on the deterministic functions, and the sampling decision for a packet can either use time-based or count-based sampling methods. Time-based sampling is particularly bad for assessing the network performance metrics such as delay and delay variation, since one tends to miss bursty periods, which contain many packets with relatively small inter-arrival times if using a larger timer, and to inject unnecessary monitoring packets during silent periods, which contain less packets if using a smaller timer. In count-based sampling method, there is a high risk that the estimation will be biased for reasons of synchronization if the metric being measured exhibits periodic behavior. Stratified sampling divides the trace into subgroups based on count-based or time-based systematic sampling. A simple random sample is then selected from each subgroup to get the actual distribution of the characteristic in the parent population. Stratified sampling may produce a gain in precision in the estimates of characteristics of the whole population than systematic sampling, but the variance of the estimated value may be still large if the population characteristic is independent or weakly correlated such as voice traffic. In this paper we propose an adaptive sampling method depending on the traffic rate. Our simulation results show that the adaptive sampling scheme is effective, provided that the appropriate sampling interval and rate can be identified and employed.

The remainder of this paper is organized as follows: Section 2 focuses on investigating the mechanisms that can be set up to adaptively adjust the parameters of the sampling technique based on the estimated traffic rate. Section 3 describes our simulation model. Section 4 introduces the evaluation criteria. Our adaptive sampling method is then evaluated in comparisons with conventional sampling methods under different types of traffic models. Finally Section 5 concludes our work and discusses some relevant issues.

II. ADAPTIVE SAMPLING APPROACH

We know that none of current sampling techniques for measuring the quality of service parameters (delay, loss, jitter, throughput) address the issue of how to carry out the sampling in an adaptive fashion so that the accuracy is better if we know something about the traffic type (voice, video and Internet data) and traffic parameters (average rate, burst size, packet length distribution). To address these issues, we introduce the concept of adaptive sampling into active measurement, and investigate how adaptive techniques can be used to adjust the sampling rate for each parameter monitored according to the availability of information about source statistics. One advantage of this method is that it can eliminate the bias caused by synchronization. It achieves this by injecting a monitoring packet randomly during every time period. Another advantage is that it can reduce the variance of the estimated values by employing TSW rate estimator for smoothing instantaneous rates. The rate estimator also reports changes of the smoothed rates to the sampler allowing sampling to be adaptive based on the measured rates. Since the number of monitoring packets over a fixed time interval is adaptively adjusted with the estimated traffic rate, the sampling rate may be reduced.

A key element in adaptive sampling is the prediction of future behavior based on observed behavior. We use the characteristics of the previous block to estimate the characteristics of the current block, and derive from those the sampling rates for the current block. If the prediction is accurate, the sampling rate can be reduced.

In summary our adaptive sampling mechanism has three distinct parts: a rate estimator, sample size estimation algorithm and sampling scheme. In the follows we will discuss each part in detail.

A. *Time Sliding Window Rate Estimator*

Time Sliding Window (TSW), a probabilistic tagging algorithm for marking packets, is proposed as part of a DiffServ mechanism in [14]. The design of TSW is very simple. TSW is employed in our scheme to estimate the traffic rate. The estimated rates are updated upon each packet arrival and then decay over time. This allows it to smooth away the noise of instantaneous rates.

B. *Sample Size Estimation Algorithm*

Time is divided into (non-overlapping) equal observation periods (referred to as time intervals). The sample size (the number of monitoring packets inserted into the traffic) for the current time interval can be adaptively adjusted based on the rate of the previous time interval reported by the rate estimator.

The number of monitoring packets to be inserted in the current time interval is adaptively adjusted with the estimated traffic rate. Upon the arrival of each time interval, the sample size is updated as: *Sample_ size = Avg_rate * Time_interval / (Block_ size * Pkt_size)*. Whereas *Pkt_size* is the average packet size, *Block_size* is the average number of data packets between two monitoring packets. This parameter can be chosen based on the tolerable overhead (monitoring packets are considered as overhead for user traffic). The *Sample_size* is the number of monitoring packets for the current *Time_interval.*

C. *Embedded Monitoring Packets*

We employ and improve the embedded monitoring method [15] by adaptively generating monitoring packets into the user traffic instead of periodically. Based on the above mentioned calculated *Sample_size*, the current time interval is divided into the timer periods following *Timer = Time_interval / Sample_size*. One monitoring packet is then time-wise randomly inserted per timer period.

Each monitoring packet is stamped with a timestamp. Receiving monitoring systems detect the monitoring packets through a unique protocol number in the header, and keep track of the number of received packets sent from the entry node.

III. MPLS-Based IP Network Model

Two relatively simple but representative network topologies based on an MPLS-based IP network are used in our simulations. The first one is the so-called simple network topology used to verify the operation of the adaptive sampling method under self-induced congestion. The second topology with competing traffic streams is used to demonstrate the effects of competing traffic on the performance of the sampling methods.

A. *Experimental Simple Network Model*

The basic setup of the simple MPLS network model is illustrated in Figure 1, and also almost the same topology is implemented with OPNET 8.0 simulation tool for our experimental analysis of adaptive sampling techniques in Section 4.

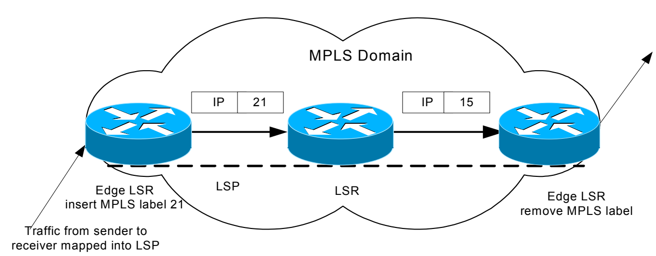

Figure 1 MPLS simple networks

In our simulation model, a static LSP is used. The edge routers are connected via a logical LSP. The LSP is in congestion as its capacity is less than the sum of the bandwidths required by all the sources of the LSP. Each end host group (gateway) consists of multiple users. Therefore the traffic stream traversing the simulation network is an aggregation of multiple individual flows. The goal is to measure and estimate the packet loss, delay and delay variation for the aggregate IP traffic generated by different source models.

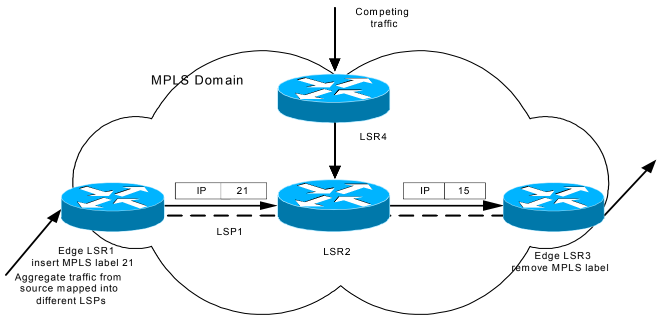

Figure 2 MPLS with competing traffic

B. *Experimental Network Model with Competing Traffic*

The network model shown in Figure 2 is designed to evaluate the performance of the adaptive sampling method when there is competing traffic on the observed path. Choosing different volumes of the competing traffic, the LSP 1 may become congested as it competes with the competing traffic for enough bandwidth to carry its traffic. We can investigate the LSP1 measurement entities under impact by the competing traffic and compare the performance of the adaptive sampling with the conventional sampling methods under different traffic models.

IV. ANALYSIS OF SIMULATION RESULTS

Simulations are carried out to evaluate the performance of different sampling algorithms under the two MPLS-IP based network models. The proposed adaptive sampling and two conventional sampling methods (systematic and stratified sampling) have been evaluated. The comparisons of different sampling techniques are conduced with equal block size so that the results are comparable. The variances of the estimates can be used to calculate confidence intervals. A higher variance corresponds to a larger confidence interval and a lower accuracy. We will show that the improvement in precision is achieved by adaptively adjusting the sampling rate based on actual traffic rate. This makes it possible to meet a given accuracy requirement by sampling much fewer packets than other sampling techniques would require.

A. *Simple Topology with Voice Traffic*

In these experiments, we use voice traffic as the network input source. The voice-sampling rate is set to 64 Kbps. Each packet contains 172 speech bytes. An active voice user generates one packet per 16ms. There are 50 voice users generating about 1.5 Mbps aggregated traffic load from source to destination. The length of simulation runs is setup to be 10 minutes. The number of replications is set to 20. For each experiment the mean value and the standard deviation for the estimation results from the 20 rounds are calculated and compared with the actual value measured with the user traffic.

In Figure 3 delay values and their respective 95% confidence intervals are plotted against block sizes. We observe

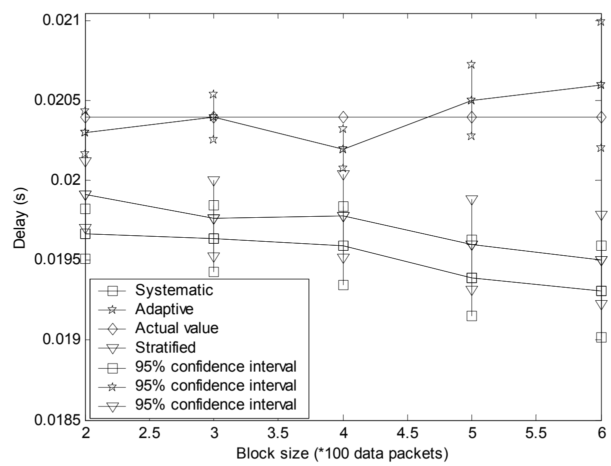

Figure 3 Delay vs. Block size – voice traffic

in this experiment that the true value of the estimated quantity lies within or very close to the confidence interval for the adaptive sampling without exception. The confidence interval is increased as the block size increases. Nevertheless, the adaptive sampling clearly gives more accurate estimated results than conventional samplings, but this better accuracy becomes less significant when block size becomes large.

In Figure 4 delay jitters, along with their respective 95% confidence intervals, are plotted against different block sizes,. With small block sizes, the values obtained by adaptive sampling and the stratified sampling methods provide a good fit for the estimated points. The confidence interval for the adaptive sampling is smaller than that for the stratified sampling in all cases, but this improvement starts getting less

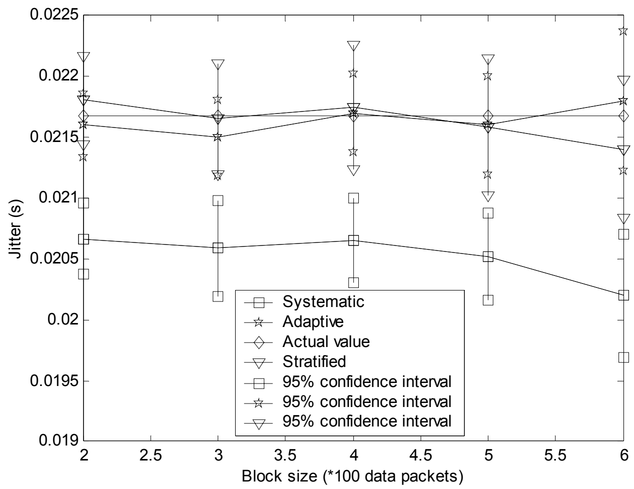

Figure 4 Delay jitter vs. Block size – voice traffic

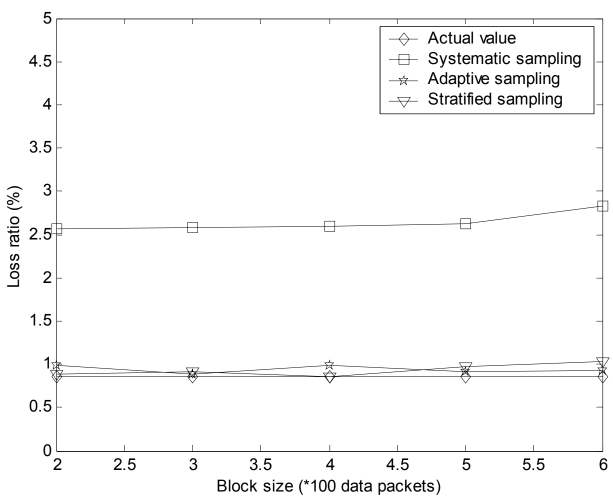

Figure 5 Loss ratio vs. Block size – voice traffic

when the block size exceeds 600 packets.

Figure 5 shows the loss ratios computed by the monitoring packets with these three different sampling methods. Compared with the true value (which is 0.8% in this test), the loss ratio of the systematic sampling is quite a bit higher, while adaptive and stratified sampling achieves values almost equal to the true value. Thus, the loss ratio can be estimated quite closely by simply counting the adaptive and stratified samples instead of all the packets.

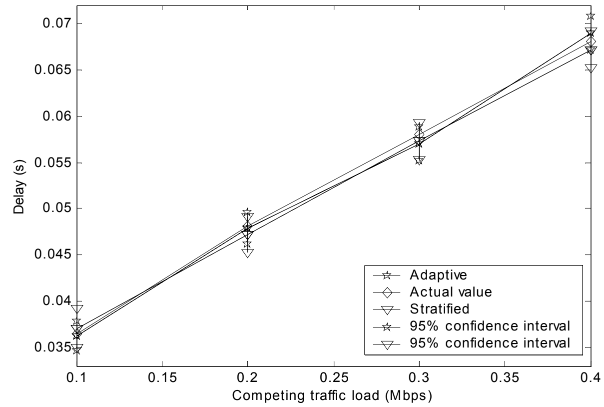

Figure 6 Delay vs. Competing traffic load (block size = 300)

B. *Competing Traffic for Voice*

The competing traffic model is developed to demonstrate the effects of competing traffic on the performance of the sampling methods. The traffic stream from LSR1 to LSR3 is called the monitored traffic, since the monitoring packets are inserted based on this traffic stream. The competing traffic is sent from LSR4 to LSR3, passing through LSR2. The monitored traffic and the competing traffic are merged at LSR2. With the volume of competing traffic increasing, the network performance under the adaptive and stratified sampling methods are simulated and evaluated with different block sizes. In this simulation, the simulation time is set as 10 minutes, and the number of replications is set to 20.

Figures 6 and 7 show the delay and the delay jitter with a 95% confidence interval estimated by the adaptive and the stratified sampling. As the competing traffic increases, the mean values of the delay are increased. We observe that in this experiment, the true value of the estimated quantity lies within the confidence interval for the adaptive and stratified sampling without exception. The confidence interval for adaptive sampling is smaller than that for stratified sampling. Thus, adaptive sampling can provide more precise estimations on delay and delay variation under the effect of the competing traffic.

Figure 8 shows the loss ratio computed by the monitoring packets with these two different sampling methods. As the competing traffic increases, the loss ratio of the monitored user traffic is changed from 0.8% to 5.8%. Compared with the true

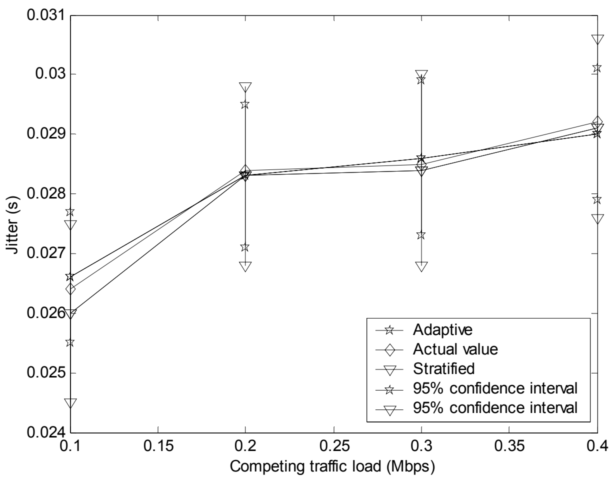

Figure 7 Delay jitter vs. Competing traffic load (block size = 300)

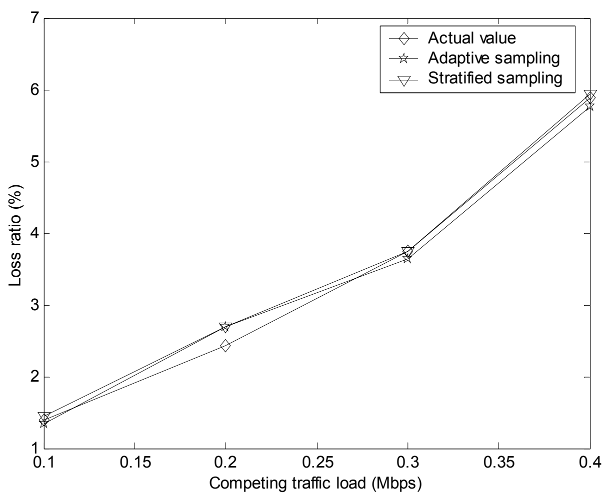

Figure 8 Loss ratio vs. Competing traffic load (block size = 300)

values, the loss ratios of the stratified and the adaptive sampling methods maintain values almost the same as the true value. Thus, the loss ratio can still be estimated by simply counting the adaptive and stratified samples instead of all the packets under the effect of the competing traffic.

V. CONCLUSIONS

In this paper, we have proposed and evaluated the performance of an adaptive sampling method for measuring the network performance of MPLS-enabled networks supporting voice traffic. Through comparisons with the conventional sampling methods, the advantages of the adaptive sampling are presented through a set of numerical results. In brief, the adaptive sampling performs better than all other methods on voice traffic. Furthermore, we demonstrated that the competing traffic has less limited impact on the sampling results when evaluating network the performance of the observed traffic stream.

REFERENCES

[1] S. Floyd, V. Paxson, "An Overview of Internet Engineering, Measurements and Modeling", ITC15 Tutorial, Washington DC, 22 June 1997

[2] V. Paxson, G. Almes, J. Mahdavi, and M. Mathis, "Framework for IP Performance Metrics," Request for Comments 2330, May 1998

[3] Advanced Network & Services, Inc., http://www. Advanced.org

[4] CAIDA Measurement Tool Taxonomy, http://www.caida.org

[5] Internet2 Working Groups, www.internet2.edu/html/working_groups.html

[6] National Laboratory for Applied Network Research (NLANR), http://moat.nlanr.net

[7] Packet Sampling Working Group https://ops.ietf.org/lists/psamp

[8] W. G. Cochran, "Sampling Techniques", Wiley, New York, 1977

[9] P.R. Krishnaiah, C.R. Rao, "Handbook of Statistics", Vol. 6: Sampling, North-Holland, and Amsterdam, 1988

[10] K.C. Claffy, G.C. Polyzos, H. Braun, "Application of Sampling Methodologies to Network Traffic Characterization", Proceedings of ACM SIGCOMM'93, San Francisco, CA, USA, September 13 - 17, 1993

[11] N. Duffield, M. Grossglauser, "Trajectory Sampling for Direct Traffic Observation", Proceedings of ACM SIGCOMM 2000, Stockholm, Sweden, August 28 - September 1, 2000

[12] I. Cozzani, S. Giordano, "Traffic Sampling Methods for End-to-end QoS Evaluation in Large Heterogeneous Networks", Computer Networks and ISDN Systems, 30(16-18), September 1998

[13] T. Zseby, "Deployment of Sampling Methods for SLA Validation with Non-Intrusive Measurements ", Proceedings of Passive and Active Measurement Workshop (PAM 2002)

[14] D. D. Clark, W. J. Fang, "Explicit Allocation of Best-Effort Packet Delivery Service", IEEE/ACM Transactions on networking, August 1998

[15] T. Lindh, "An Architecture for Embedded Monitoring of QoS Parameters in IP Based Virtual Private Networks", Proceedings of Passive and Active Measurement Workshop (PAM 2001)