PAPER

Virtual Queue Occupancy and Its Applications on Periodic Bandwidth On Demand Schemes for IP/SONET

Jingdi ZENG[†], *Nonmember and* **Nirwan ANSARI**[†], *Regular Member*

SUMMARY Carrying IP traffic over connection-oriented networks requires the use of bandwidth on demand schemes at gateways or network interfaces. A new virtual queue occupancy, which is more accurate than the classical one, is being proposed for IP/SONET bandwidth on demand. Based on the virtual queue occupancy, two enhanced periodic approaches for lossless services, LAVQ and LAVQL, are simulated and evaluated. Simulations show that LAVQ outperforms its counterpart LAQ in terms of bandwidth utilization. By curbing the queue occupancy fluctuation, LAVQL further promotes bandwidth utilization and conceals the influence of the system latency on delay jitter as well.

key words: *bandwidth on demand, virtual queue occupancy, periodic bandwidth allocation, IP/SONET*

1. Introduction

With the explosive growth of Internet users, carrying IP over connection-oriented networks such as ATM has been intensively studied [1]-[5]. Moreover, the industry is envisioning the evolution of network structures by combining two technically disparate segments [6]-[8]: the optical domain and the electrical domain. Integrating the management matured electrical domain and the bandwidth abundant optical domain is becoming a hot issue. Among a number of proposed solutions, such as IP/DWDM, IP/SONET, IP/ATM/DWDM, and IP/ATM/SONET, IP/SONET benefits from the integration simplicity and standard maturity.

Synchronous Optical Network (SONET) was originally designed to carry voice services. Once a network allocates a circuit to a user or link, any excess bandwidth within the circuit cannot be reused by others. When transferring bursty IP traffic, network engineers usually over-subscribe the bandwidth to leave a safe margin that is actually a waste of bandwidth. The resulted average circuit utilization rate of SONET is 5 to 10 percent in the access network, and 20 to 30 percent in the core [9]. Therefore, reacting to the real time traffic on the network edge and adjusting bandwidth on a link or a path, referred to as bandwidth on demand, is an issue in IP/SONET as well as in other inter-networking solutions.

IP/SONET, moreover, is characterized by two other attributes: 1. IP datagrams with different lengths make the queue behavior in IP/SONET more complicated than other solutions dealing with fixed-length cells/packets. 2. SONET allocates the bandwidth in a quantized manner. In a SONET port with the finest 1.5 Mbps granularity, for example, a 13.6 Mbps and a 14.4 Mbps bandwidth requirements both receive a 15.0 Mbps allocation. These features distinguish IP/SONET bandwidth on demand schemes from legacy ones such as those for IP/ATM/SONET and IP/ATM.

In the literature, there are three categories of bandwidth on demand strategies [2]: static algorithms keep the original allocation once the bandwidth is set up; periodic algorithms update the bandwidth at equal time intervals; adaptive algorithms allocate the bandwidth whenever the preset conditions are met. In addition, traffic monitoring algorithms can be classified into two broad categories: one follows the real time traffic intuitively [2][3]; the other one learns about and predicts the incoming traffic by analyzing statistical data [4][5].

This article focuses on and introduces periodic bandwidth on demand schemes tailored for IP/SONET. These algorithms are applicable to a single user or link, where the bandwidth can be adjusted based on the queue status change owing to the traffic fluctuation. A new virtual queue occupancy is proposed and applied to two periodic bandwidth on demand schemes.

Several periodic bandwidth on demand policies are reviewed in Sect. 2. The system model is described in Sect. 3. In Sect. 4, a new virtual queue occupancy for bandwidth on demand is proposed. Simulation results and performance evaluation for two IP/SONET bandwidth on demand schemes are presented in Sect. 5. The article concludes in Sect. 6.

2. Previous Work

While static and adaptive bandwidth on demand schemes have the best complexity and flexibility performance, respectively, periodic schemes make a better compromise because they allow acceptable complexity and flexibility at the same time. A periodic scheme calculates and updates the demanded bandwidth for one period by using parameters derived from the last period, such as the queue length and the arrival rate.

Using the average arrival rate, a basic policy is

Manuscript received August 3, 2001.
Manuscript revised January 18, 2002.

[†] The authors are with the Center for Communications and Signal Processing, Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, New Jersey 07102, USA.

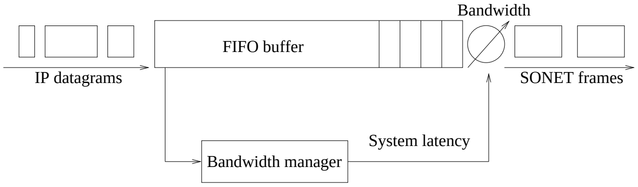


Fig. 1 A bandwidth on demand system model.

to allocate for the next period the amount of bandwidth required to transmit the same amount of traffic that has arrived in the previous period [2]. This passive allocation strategy does not work well. When the traffic increases, the allocated bandwidth falls short of the requirement; when the traffic decreases, this scheme wastes resources by still allocating the same amount of bandwidth as in the preceding period.

An alternative strategy [3] multiplies the assigned bandwidth by a constant factor C that is greater than 1. Allocating the bandwidth more aggressively, this strategy has shorter queue and smaller loss than those of the first one. The dumb parameter, C, however, introduces the performance uncertainty as well as system resource wastage.

Instead of measuring the average arrival rate, other schemes [2][3] employ the queue occupancy up to the beginning of the current period. They allocate for this period the amount of bandwidth needed to drain what has accumulated in the queue during the previous period. The queue status preciseness, as the only parameter, however, may not be robust enough. One comprehensive policy [2], referred to as Last Arrival plus Queue occupancy (LAQ) in this article, takes both the arrival rate and queue status into consideration. It allocates for the next period the amount of bandwidth necessary to transmit the current content of the queue plus the same amount of traffic that has arrived in the previous period. This is a simple but effective method, without requiring any knowledge about the next period. Nevertheless, for an IP/SONET environment, this bandwidth on demand scheme has its drawbacks and thus needs to be tailored, as will be discussed in Sect. 4.

3. System Model

It is assumed that the traffic from a user has already been monitored by an appropriate policing mechanism, and is therefore conforming to its traffic contract with the Internet Service Provider (ISP). A simple tail-drop policy is used for the packet dropping. Owing to signalings and other system operations, there is a system latency, a certain time lag between the time the required bandwidth is calculated and the time the update is reflected in the actual allocation. Accordingly, the system is modeled as a single server queue utilizing a bandwidth manager with a certain system latency, as shown in Fig. 1.

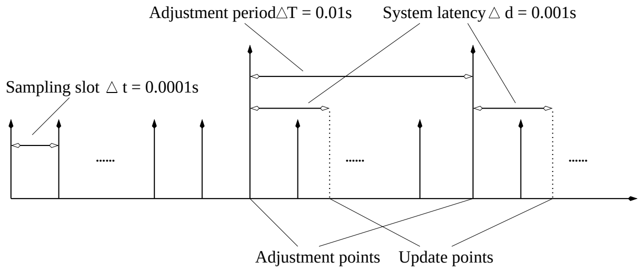


Fig. 2 Periodic measurement.

The following notation is used throughout this article:

ΔT - the adjustment period, which is the adjustment frequency of bandwidth allocation.

Δt - the sampling slot, which is the sampling frequency of arrival rates.

Δd - the system latency, which is the length of time necessary for the update to be effective.

ΔO - the buffer idle timer, which records how long the buffer has been empty in an adjustment period ΔT.

ΔG - the finest SONET granularity, which is assumed to be 1.5 Mbps.

λ_i - the average arrival rate in the adjustment period i.

R_p - the peak arrival rate.

q_i - the queue occupancy at the end of period i.

Q_i - the virtual queue occupancy at the end of period i, which is yielded from q_i and will be described in depth in Sect. 4.

c_i - the required bandwidth estimated for period i.

C_i - the actually allocated bandwidth for period i due to the SONET granularity, i.e., $\lceil \frac{c_i}{\Delta G} \rceil \times \Delta G$.

$\widehat{C_i}$ - the actually required bandwidth, which is an unknown parameter, for period i.

ρ - bandwidth utilization, i.e., $\frac{\sum \lambda_i}{\sum C_i}$, which is affected by the SONET granularity, and is thus called the granular utilization in the rest of the article.

B - the preset buffer size.

As shown in Fig. 2, the bandwidth manager samples designated parameters every Δt seconds and adjusts bandwidth allocation every ΔT seconds. The updated bandwidth allocation takes place after the system latency Δd. To make the periodic calculation meaningful, the bandwidth manager should finish updating the allocation before the next round of calculation, i.e., $\Delta d \leq \Delta T$. According to the industry implementation [9], the dynamic allocation process on a SONET node can be achieved in one millisecond, and therefore, a system latency $\Delta d = 0.001$ seconds is assumed in this article. Moreover, with the speed of light in glass of about 0.01 seconds per 1,000 miles, a conservative but realistic adjustment period $\Delta T = 0.01$ is assumed. Obviously, the adjustment period is ten times as much as the system latency, providing enough time for the

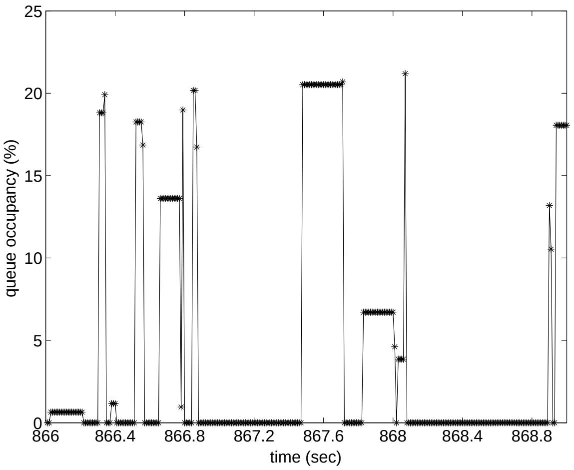


Fig. 3 LAQ queue occupancy vs. time.

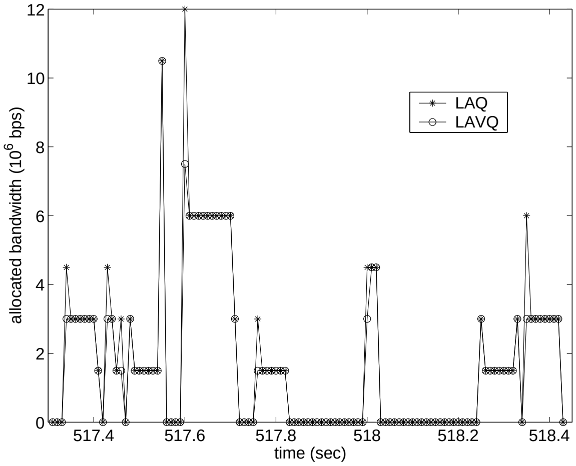


Fig. 4 Bandwidth allocation (LAQ, LAVQ) vs. time.

bandwidth to be updated. The newly adjusted bandwidth is conveyed by the control signaling. Please note that routing is not considered because a route or path, which has already been set up, will not be changed until the next route update which is done sporadically. As far as measurement parameters are concerned, it is assumed that the bandwidth manager measures arrival rates every $\Delta t = 0.0001$ seconds to capture the traffic arrival characteristic as precisely as possible. In addition, the average arrival rate λ_i and the granular utilization ρ are collected every 100 sampling slots and 100 adjustment periods, respectively. Simulations in subsequent sections are based on these system and parameter assumptions. However, other conditions, such as various adjustment periods and system latencies, should be considered in future study.

4. Virtual Queue Occupancy

LAQ [2][3], one of the few periodic bandwidth on demand schemes with better performance, measures the average arrival rate λ_{i-1} in period $i-1$, determines the queue occupancy q_{i-1} at the end of period $i-1$, and then allocates the bandwidth for period i as follows:

$$c_i = \lambda_{i-1} + \frac{q_{i-1}}{\Delta T}. \tag{1}$$

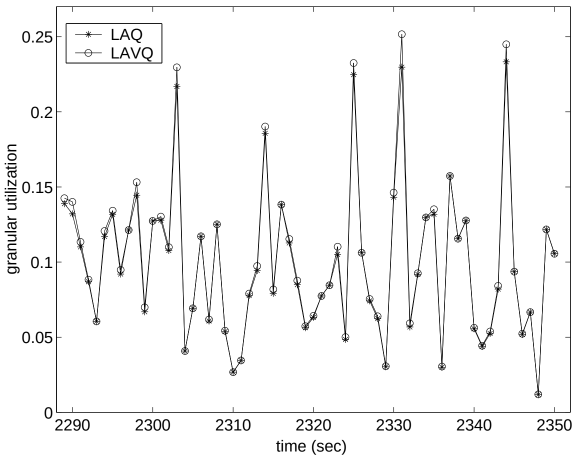


Fig. 5 Granular utilization ρ (LAQ, LAVQ) vs. time.

Since LAQ does not have any prediction ability, the performance of LAQ depends on how accurate the average arrival rate and the queue occupancy are. The smaller the sampling slot Δt is, the better the average arrival rate λ_{i-1} follows the traffic trend. Likewise, this relationship holds between the queue occupancy q_{i-1} and the adjustment period ΔT.

The queue occupancy, however, needs more detailed investigation. Obviously, the average data arrival on a link may not be equal to the bandwidth assigned to this link. Buffer overflows or underflows, consequently, will likely happen if the bandwidth is not adjusted properly. In general, buffer overflows occur when the assigned bandwidth is less than the average data arrival, and buffer underflows occur when the allocated bandwidth is more than the average data arrival. Note that an empty buffer is the extreme case of the underflow. Since an empty buffer definitely infers a slow arrival rate, the bandwidth manager should reduce the bandwidth allocation in the upcoming period. In LAQ, as it turns out, when the bandwidth manager sees the buffer is empty at the end of period $i-1$, the queue occupancy q_{i-1} is 0, and the allocated bandwidth for period i does reduce to $c_i = \lambda_{i-1} + \frac{q_{i-1}}{\Delta T} = \lambda_{i-1}$. However, this only works for a buffer which is recently empty because more bandwidth should have been taken back if the buffer has been empty for a while. As a result, the bandwidth manager, due to its periodic measurement pattern, does not know whether the buffer status is newly changed or has lasted for a while. From another point of view, the queue occupancy q_{i-1} does not always reflect the real buffer status because of possible buffer underflows.

As shown in Fig. 3, owing to the SONET band-

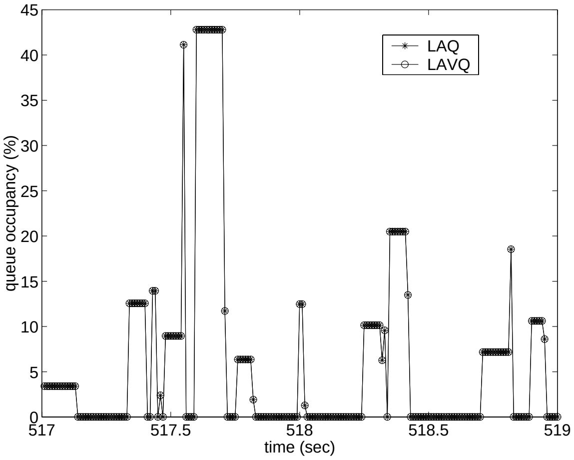


Fig. 6 Queue occupancy (LAQ, LAVQ) vs. time.

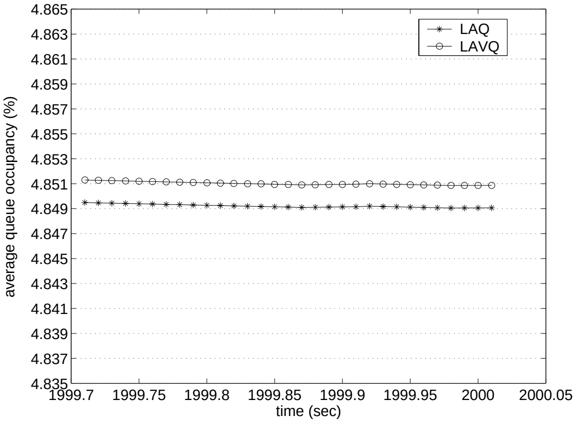


Fig. 7 Average queue occupancy (LAQ, LAVQ) vs. time.

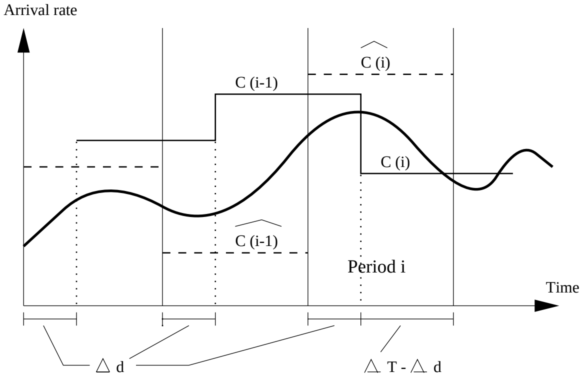


Fig. 8 The buffer size bound.

width granularity, the buffer is extremely underflowed or empty quite often when using LAQ for bandwidth allocation. This phenomenon impacts the queue occupancy q_i and thus the bandwidth allocation preciseness as discussed in the previous paragraph. One way to counteract this effect is to extend the queue occupancy q_i beyond its range $[0\,,\,B]$ so that it reflects the buffer status more precisely. This extended value, called virtual queue occupancy Q_i, can be calculated based on a buffer idle timer ΔO. It intents to capture the time duration in which the buffer has been empty during an adjustment period ΔT. Moreover, for the sake of the information opportuneness, only the latest duration where the buffer remains empty up to the end of the adjustment period is considered. As a result, the proposed bandwidth allocation for period i is determined from

$$c_i = \lambda_{i-1} + \frac{Q_{i-1}}{\Delta T}, \tag{2}$$

where

$$Q_{i-1} = q_{i-1} - \Delta O \times C_{i-1}. \tag{3}$$

When the buffer is empty, $q_{i-1} = 0$, and thus the virtual queue occupancy becomes

$$Q_{i-1} = -\Delta O \times C_{i-1}. \tag{4}$$

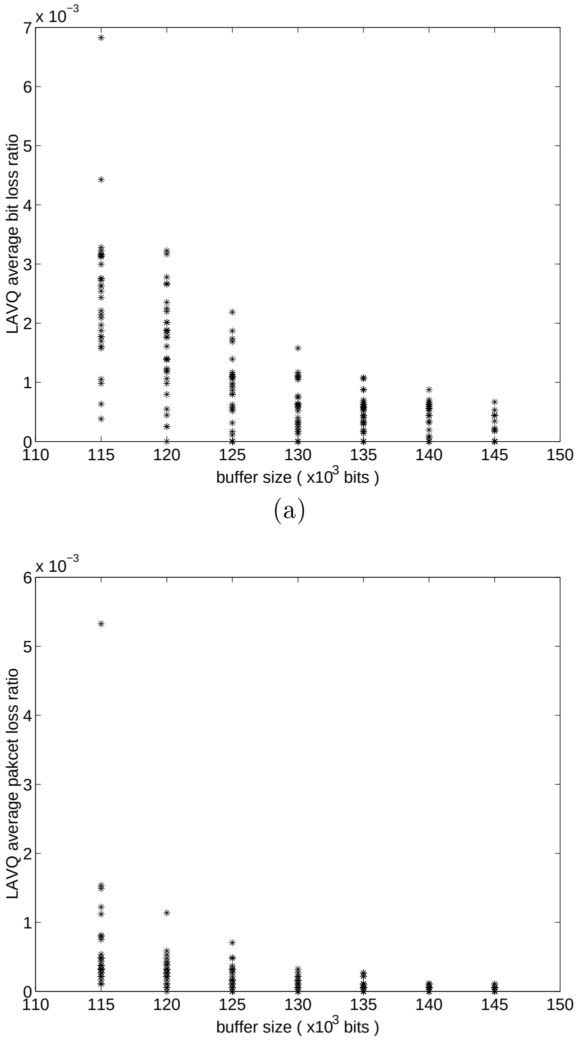
With the extended range $[-\Delta O \times C_{i}, B]$, the virtual queue occupancy Q_i minimizes the inaccurancy due to a prolongedly empty buffer. Numerally, it reduces the bandwidth requirement by $\frac{\Delta O}{\Delta T} \times C_{i-1}$ when the buffer has been empty for the duration of ΔO, which is rather frequent in an IP/SONET environment.

5. Simulations and Discussion

The performance of two bandwidth on demand schemes proposed in the following subsections is simulated in a First In First Out (FIFO) buffer ($B = 130K\,bits$) which is expected to provide lossless services as will be discussed later. Although the average session holding time and the inter-arrival time at IP access networks are at the minute and second level, respectively [10], for the sake of the simulation convenience, the session holding time of two sources is assumed exponentially distributed with the mean of 20 ms and 40 ms, and the inter-arrival time follows the same distribution with the mean of 4 ms and 7 ms, respectively. The packet inter-arrival time of these two sources is Pareto distributed with shape parameters of 1.6 and 1.1, while the size of the packets is generated following an exponential distribution with a mean of 1500 bytes. Each simulation lasts for a 40-minute period, corresponding to 240 thousand adjustment periods and 24 million sampling slots. Furthermore, to better approach the real trend of unknown parameters, two of the simulations have 30 repetitions of trials with different random seeds.

5.1 LAVQ and the Buffer Size

By replacing the queue occupancy q_i with the virtual queue occupancy Q_i, Last Arrival plus Virtual Queue occupancy (LAVQ) scheme allocates less bandwidth

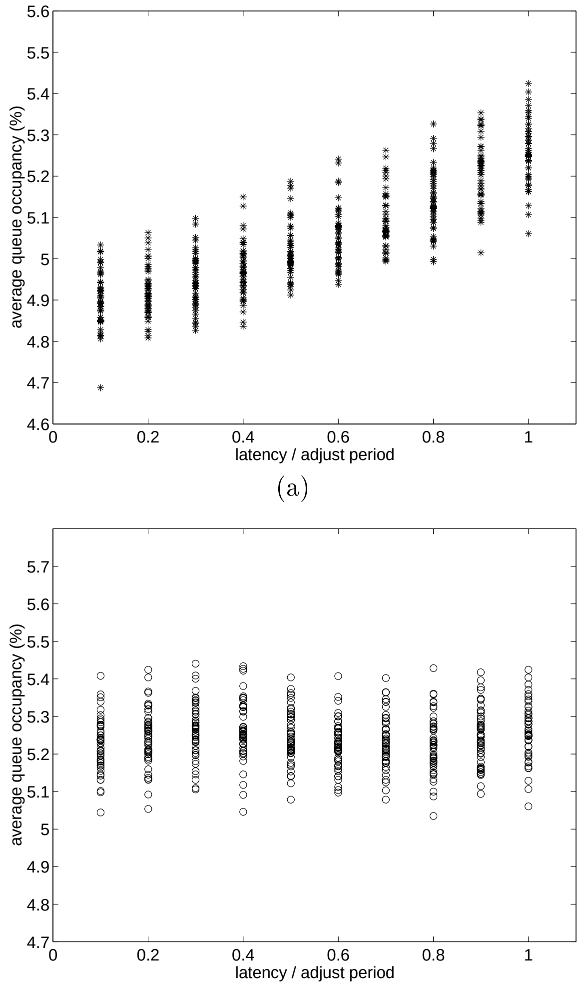


(a)

(b)

Fig. 9 LAVQ average loss ratio vs. buffer size (each point represents one run of simulation): (a) bit loss ratio; (b) packet loss ratio.

than what LAQ does, as shown in Fig. 4 in a small time scale. This conforms to the earlier expectation: in addition to maintaining at least the same performance of LAQ, LAVQ regains over-assigned bandwidth when the buffer underflows. The accumulated effect in a larger time scale is illustrated in Fig. 5, where LAVQ achieves bigger granular utilization ρ and thus better bandwidth utilization. To assure that the performance improvement of bandwidth utilization is not traded with the buffer resource, the queue occupancy of LAQ and LAVQ is shown in Fig. 6 in a small time scale and in Fig. 7 with an average view, respectively. Simulations show that the improved bandwidth utilization can be achieved with a negligible 0.002 percent queue occupancy difference between LAQ and LAVQ. This is exactly the advantage of LAVQ.



(a)

(b)

Fig. 10 Average queue occupancy vs. $\frac{\Delta d}{\Delta T}$ (each point represents one run of simulation): (a) LAVQ; (b) LAVQL.

Next, the lower bound of the buffer size to warrant lossless services shall be derived. As shown in Fig. 8, in the first Δd duration of period i, a user or link still uses bandwidth C_{i-1} realized in the last period $i-1$ because of the system latency Δd; during the rest of the period, the updated bandwidth C_i is used.

Lemma: To warrant lossless services, the buffer size B should meet the requirement of

$$B > \left\lceil \frac{R_p}{\Delta G} \right\rceil \times \Delta G \times \Delta T.$$

Proof: With granularity ΔG, system latency Δd, and the actually required bandwidth $\widehat{C_{i-1}}$ and $\widehat{C_i}$ for period i and $i-1$, respectively, the queue space needed to hold possibly delayed and underestimated traffic is

$$q_i = (\widehat{C_i} - C_{i-1}) \times \Delta d + (\widehat{C_i} - C_i) \times (\Delta T - \Delta d). \tag{5}$$

In the worst case, for any integer i, it is true that $\widehat{C_i} = \left\lceil \frac{R_p}{\Delta G} \right\rceil \times \Delta G$ and $C_i = 0$, or vice versa. Thus, both $\widehat{C_i} - C_i$ and $\widehat{C_i} - C_{i-1}$ lie in the range of

$$\left[-\left\lceil \frac{R_p}{\Delta G} \right\rceil \times \Delta G \,,\; \left\lceil \frac{R_p}{\Delta G} \right\rceil \times \Delta G \right].$$

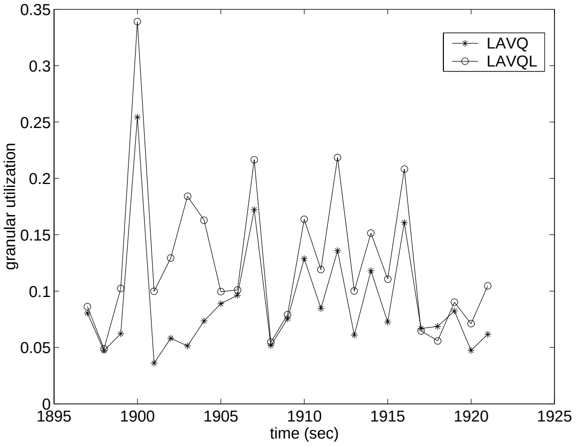


Fig. 11 Granular utilization ρ (LAVQ, LAVQL) vs. time.

The required queue space turns out as

$$q_i \in \left[-\left\lceil \frac{R_p}{\Delta G}\right\rceil \times \Delta G \times \Delta T, \left\lceil \frac{R_p}{\Delta G}\right\rceil \times \Delta G \times \Delta T\right]. \tag{6}$$

For lossless services, therefore, a buffer size of

$$B > \left\lceil \frac{R_p}{\Delta G}\right\rceil \times \Delta G \times \Delta T. \tag{7}$$

is needed. □

Accordingly, the lower bound of the buffer size is 120K bits in this article, where $R_p = 10.6\,Mbps$, $\Delta T = 0.01$ seconds, and $\Delta G = 1.5\,Mbps$. Even in an OC192 (9953.28 Mbps) port, the buffer size for lossless services is only 12.45M bytes.

The average loss ratios of LAVQ do show the trend of approaching zero when the buffer size increases, as shown in Fig. 9. However, since the tail-drop method drops a whole packet even when 99 percent of it has already been in the buffer, the bit loss ratio of IP datagrams with different lengths suffers more than the packet loss ratio. This explains the symptom exhibited in Fig. 9, where the bit loss ratio shows slower convergence process than the packet loss ratio does. The effect of different IP datagram lengths, also as shown in Fig. 9, is that both ratios do not reach zero right at the calculated buffer size bound. The time lag between the calculated bound and the simulation results implies that a more proper model, rather than the fluid flow model, is desired to derive the buffer size bound.

5.2 LAVQL and the System Latency

To provide lossless services, LAVQ needs a buffer to hold the underestimated traffic and the delayed traffic resulted from the system latency Δd. Because of the burstiness of the traffic, network engineers have to set up surplus buffer spaces according to the previously derived buffer size bound. Furthermore, it is observed that the ratio of the system latency to the adjustment period, $\frac{\Delta d}{\Delta T}$, has a big impact on the queue status of LAVQ. As presented in Fig. 10(a), when the ratio increases, its queue occupancy grows in a nonlinear pattern. The system latency Δd, moreover, changes unpredictably in the range of 0 to ΔT because it is related to many uncertain factors, such as the traffic load and available network resources. The former problem, the redundant buffer size, causes an inefficient resource usage and possibly long delay. And the latter one, the fluctuational queue occupancy resulted from the uncertain system latency, causes extra delay jitter which is defined as the absolute delay difference experienced by any two packets in a connection.

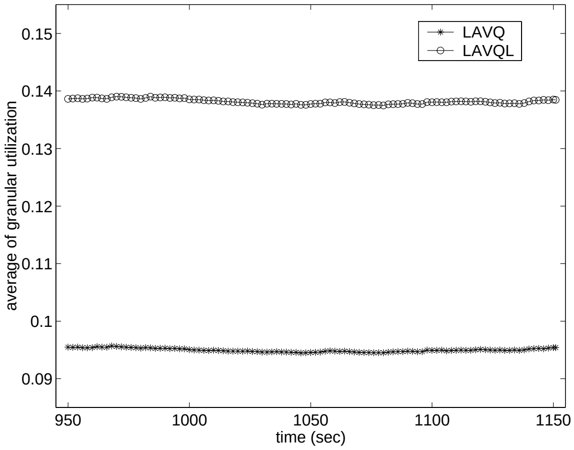


Fig. 12 Average granular utilization (LAVQ, LAVQL) vs. time.

Although delay and delay jitter are literally fine-tuned by further scheduling operations, it makes sense for bandwidth allocation discussed here to provide gross delay performance guarantee. Since providing lossless services is the main concern of LAVQ, the possible long delay cannot be avoided. However, for a redundant buffer, tightening the bandwidth allocation is a straightforward way to increase the buffer usage efficiency. Meanwhile, if the queue occupancy fluctuation caused by the system latency can be controlled, the unexpected delay jitter will be minimized. Aiming to decrease the allocated bandwidth in a manner that curbs the queue occupancy fluctuation, based largely on intuition, a factor which is smaller than 1 and compromises the nonlinear trend of the queue occupancy, is needed. Therefore, by introducing a quadratic factor $(\frac{\Delta d}{\Delta T})^2$, an enhanced scheme, referred to as Last Arrival plus Virtual Queue occupancy plus system Latency (LAVQL), allocates bandwidth according to

$$c_i = \left(\frac{\Delta d}{\Delta T}\right)^2 \times (\lambda_{i-1} + \frac{Q_{i-1}}{\Delta T}). \tag{8}$$

The original objective behind this scheme is twofold: increasing the buffer usage, and curbing the queue occupancy fluctuation to reduce delay jitter.

As it turns out, by holding more traffic in the buffer and allocating less bandwidth than LAVQ, LAVQL does promote the buffer usage efficiency. Besides, its reduced bandwidth allocation brings significantly improved bandwidth utilization. As shown in Fig. 11 and Fig. 12, LAVQL reaches about 14 percent average bandwidth utilization which corresponds to a 40 percent improvement over the 5 to 10 percent industry statistics. Moreover, as shown in Fig. 10(b), when the latency ratio $\frac{\Delta d}{\Delta T}$ varies, the queue occupancy of LAVQL does not fluctuate as much as that of LAVQ. In other words, LAVQL grossly absorbs extra delay jitter induced by the system latency Δd.

6. Conclusions

This article has proposed a new virtual queue occupancy which indicates a more accurate queue status for allocating bandwidth on demand. Two periodic bandwidth on demand schemes, LAVQ and LAVQL, have subsequently been proposed based on this new measure. Providing lossless services, LAVQ achieves better bandwidth utilization than its counterpart LAQ; by taking advantage of redundant buffer spaces, LAVQL once again promotes bandwidth utilization and conceals delay jitter induced by the system latency. Further work is planned on more delicately defining the overflow duration by using different thresholds, where IP datagram dropping strategies may play important roles. In addition, scheduling strategies which can fine-tune delay performances could be further integrated. Last, using the traffic prediction to estimate average arrival rates more precisely will certainly improve the bandwidth allocation accuracy.

Acknowledgments

This work has been supported in part by OpenCon Systems, Inc., the New Jersey Commission on Higher Education via the NJI-TOWER project, and the New Jersey Commission on Science and Technology via the New Jersey Center for Wireless Telecommunications.

References

[1] A.L. B.P. B. Diniz, C.C. Goulart, and J.M. S. Nogueira, "A dynamic bandwidth allocation service for ATM-based networks supporting multimedia applications," Proc. 1st IEEE Int. Conf. on ATM, Colmar, France, pp. 292–299, June 1998.

[2] Y. Afek, M. Cohen, E. Haalman, and Y. Mansour, "Dynamic bandwidth allocation policies," Proc. IEEE INFOCOM, San Francisco, USA, pp.880–887, March 1996.

[3] J. Brassil, "Overcoming reservation latency for IP traffic in an ATM PON," Proc. IEEE GLOBECOM, Rio de Janeiro, Brazil, pp.1155–1161, Dec. 1999.

[4] Q.K. Mehra, B. Ravichandran, J.B. D. Cabrera, D.N. Greve, and R.S. Sutton, "Towards self-learning adaptive scheduling for ATM networks," Proc. 36th Conf. on Decision & Control, San Diego, USA, pp.2392–2398, Dec. 1997.

[5] F. Zhao, M. Hanawa, and M. Takahara, "Traffic measurement-based admission control using rate envelope in ATM networks," Proc. IEEE Int. Conf. on Networks, Singapore, pp.223–229, Sept. 2000.

[6] A. Copley, "Optical domain service interconnect (ODSI): defining mechanisms for enabling on-demand high-speed capacity from the optical domain," IEEE Commun. Magazine, vol.38, no.10, pp.168–174, October 2000.

[7] J. Manchester, J. Anderson, B. Doshi, and S. Dravida, "IP over SONET," IEEE Commun. Magazine, vol.36, no.5, pp.136–142, May 1998.

[8] R. Jain, "IP over SONET," March 1998. http://www.cis.ohio-state.edu/˜jain/talks/ipsonet.htm.

[9] G. Lee, "FlexBand technology optimizes SONET," White Paper, Geyser Networks, Sunnyvale, USA, 2001.

[10] J. Färber, S. Boamer, J. Charzinski, "Measurement and modeling of Internet traffic at access networks," Proc. EUNICE Open European Summer School on Network Management and Operation, München, Germany, pp. 196–203, Aug. 1998.

Jingdi Zeng received the B.E. and M.E. degrees in communications engineering and computer applications, all from Hunan University, P.R.China, in 1995 and 1998, respectively. She is currently pursuing her doctoral studies at the New Jersey Institute of Technology. Her current research interests include resource management and Internet QoS. She is a member of IEEE and Alpha Epsilon Lambda (AEL).

Nirwan Ansari received the B.S.E.E. (summa cum laude), M.S.E.E., and Ph.D. from NJIT, University of Michigan, and Purdue University in 1982, 1983, and 1988, respectively. He joined the Department of Electrical and Computer Engineering, NJIT, in 1988, and has been Professor since 1997. He is a technical editor of the IEEE Communications Magazine, was instrumental, while serving as its Chapter Chair, in rejuvenating the North Jersey Chapter of the IEEE Communications Society which received the 1996 Chapter of the Year Award, currently serves as the Chair of the IEEE North Jersey Section, and also serves in the IEEE Region 1 Board of Directors and various IEEE committees. He was the 1998 recipient of the NJIT Excellence Teaching Award in Graduate Instruction, and a 1999 IEEE Region 1 Award. His current research focuses on various aspects of high speed networks. He authored with E.S.H. Hou Computational Intelligence for Optimization (1997, and translated into Chinese in 2000), and edited with B. Yuhas Neural Networks in Telecommunications (1994), both published by Kluwer Academic Publishers.