Congestion Control for Real-Time Traffic in High-Speed Networks

Henning Schulzrinne
Dept. of Electrical and Computer Engineering
University of Massachusetts
Amherst, MA 01003

James F. Kurose and Don Towsley
Dept. of Computer Science
University of Massachusetts
Amherst, MA 01003

Abstract

Traditional feedback congestion control is not suitable for high-speed networks due to their high bandwidth-to-delay ratio. In this paper, we investigate the possibility of locally controlling short-term congestion for loss-tolerant, but delay-sensitive traffic such as packet voice through selective discarding of packets based on the virtual work found by a packet on arrival to a queue (local deadlines). By analysis and simulation of a multi-stage virtual circuit it is shown that this approach can cut voice-tolerable loss rates in half for high loads. Further, we show that the simple case of using the same local deadline throughout the network performs nearly as well as taking reduced interior traffic into consideration and optimizing loss performance over a set of heterogeneous local deadlines. By example, the issue of establishing control parameters at call setup time is also considered.

1 Introduction

The higher bandwidths promised by broadband integrated services digital networks (BISDN) have made applications with real-time constraints, such as control, command, and interactive voice and video communications, feasible. Many types of real-time traffic are characterized by "perishable", but redundant messages. In other words, excessive delay renders them useless, but a certain degree of loss can be tolerated without objectionable degradation in the grade of service. Real-time packets are lost for several reasons. The packet may arrive at the receiver after the end-to-end deadline has expired after having suffered excessive waiting times in the intermediate nodes (*late loss*). Also, intermediate nodes may shed load by dropping packet as an overload control measure (*drop loss*).

The tolerance for packet loss varies with the type of traffic carried and the measures the receiver is prepared to take to reconstruct lost packets. For speech coded with PCM and adaptive DPCM, losses from 2 to 5% are tolerable without interpolation, while odd-even interpolation raises the threshold of objectionable loss to between 5% and 10% [1]. Compressed video is far less tolerant of lost packets. A variable-rate DPCM coder operating at around 30 Mbits/second [2] is designed for a loss rate of below 10^{-11}.

The goal of this paper is to investigate controlling congestion for real-time traffic by judiciously discarding packets at the intermediate nodes. Packets that stand little chance of making their end-to-end deadline should be discarded as early in the virtual circuit as possible. Discarding applied with care has two beneficial effects. First, *instantaneous* congestion at the discarding node is relieved. This type of congestion is caused by the statistical fluctuations in the arrival process. Secondly, downstream nodes are not burdened with traffic that in all likelihood will not meet its deadline at the receiving end, speeding up service for the remaining packets. Reduced downstream load counteracts *long-term* congestion, expressed as high average load.

Barring clairvoyance, we have three choices in discarding packets, depending on our degree of caution and amount of available information (see Fig. 1):

1. discard a packet if the time spent at the node under consideration exceeds the end-to-end deadline d
2. discard a packet whose time spent in the virtual circuit so far, including time in current node, exceeds the end-to-end deadline d
3. discard a packet if the time to be spent in current node i exceeds a fixed *local* deadline τ_i

The first two schemes above are conservative in the sense that no packet will ever be dropped voluntarily at the intermediate nodes that would not have been late at the destination. The third approach risks the possibility of discarding packets that would have made the end-to-end deadline. The parameter space investigated will cover the first and third possibility, as the second requires "travel history" to be carried with each packet.

We note that in a high-speed wide-area environment, control measures such as backpressure or choke packets [3] lose effectiveness since the high delay-to-bandwidth ratio may render the feedback information obsolete by the time it reaches the controlling agent. Because of this, local, distributed control mechanisms are preferable to centralized algorithms.

After outlining the system model underlying this paper in section 2 and the general equations governing packet loss in section 3, we will in turn investigate two queueing models for the proposed control mechanism, first-in-first-out with bounded *system* time, referred to for brevity as FIFO-BS (section 4), and first-in-first-out with bounded *waiting* time (FIFO-BW) in section 5.[1] For each, the expressions for the conditional distribution of the system or waiting time available in the literature will be reviewed and extended where necessary. Expressions and numerical values for the end-to-end delay distributions are presented and compared to simulation results. Node rejection probability and the

This work is supported in part by the Office of Naval Research under contract N00014-87-K-304, the Defense Advanced Research Projects Agency under contract NAG2-578 and an NSF equipment grant, CERDCR 8500332.

[1] In this paper, *system time* is defined to include service, i.e., transmission, time. *Waiting time* refers to the delay between arrival and start of service.

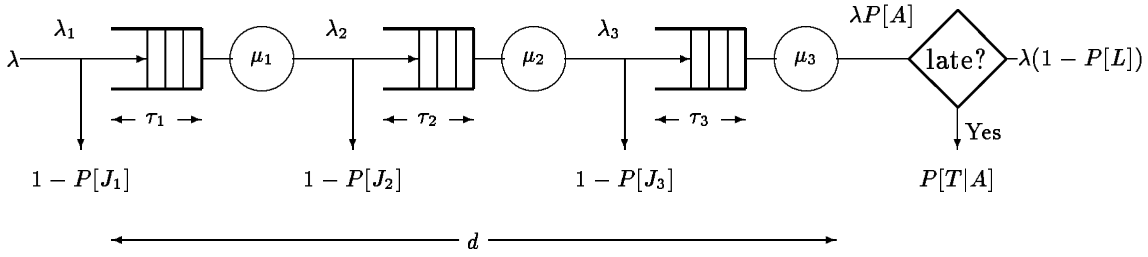

Figure 1: Sample virtual circuit

tail of the delay distribution establish the end-to-end loss probability under the various congestion control schemes. A simple scheme of using the same value for all local deadlines in the VC performs will be shown to perform almost as well as the more difficult scheme of selecting optimal local deadlines for each node. We discuss implementing these policies in high-speed networks and show that they compare favorably to random discarding as a congestion countermeasure. We conclude with some alternative approaches and future work in section 7.

2 System Model and Notation

This work is primarily concerned with packet voice traffic. The approaches are also applicable to other real-time traffic, such as sensor data and control commands, assuming that it shares a similar tolerance for loss and similar traffic characteristics.

We study virtual circuits (VCs) with nodes labeled $i = 1, 2, \ldots M$, typically switching nodes in a wide-area network interconnected by high-bandwidth fiber optic links. Packet losses caused by bit errors and other hardware faults are ignored as the probability of their occurrence falls several orders of magnitude below that of buffer-related losses. Packet arrivals to all nodes are approximated by Poisson processes with arrival rate λ_i. Modeling a superposition of SAD (speech activity detection) voice sources as a Poisson source leads to overly optimistic estimates in the case of a voice multiplexer [4]. However, for the case of voice trunks considered here, the number of voice sources is possibly two orders of magnitude higher than that found at the T1 rate (1.544 Mbps) multiplexer studied by [5] and others. A 150 Mbit channel can support 4380 64 KBit/sec PCM sources of speech activity factor 0.42 with a utilization of 0.8. Thus, according to the arguments of [6], we can conclude that for short time spans, the superposition process is adequately modeled by a Poisson process. The discarding mechanism further limits the interaction period of packets in the queue, improving the approximation [4].

Even assuming that the input traffic to the first node is indeed modeled accurately by a Poisson source, packet dropping at the intermediate nodes will make the input stream to subsequent nodes non-Poisson. The analysis, however, ignores this, based on the premise that the relatively small losses investigated should have limited effects on the delay distribution. Simulation experiments will be used to assess the validity of this assumption, which is also supported by [7, 8].

The service time is assumed to be exponentially distributed with mean $1/\mu$, redrawn at each node (Kleinrock's independence assumption). Without loss of generality, all times are scaled relative to the average service time, i.e., $\mu = 1$. Since we are primarily interested in communication channels, we assume that the packet transmission and waiting time is proportional to the number of bits in the packet or queue, respectively. Also, the service time of each packet is assumed to be known at the instant of arrival to a queue. (The results for FIFO-BW do not change if the packet leaves the queue if it has not started service within the local deadline; the service time of an individual packet does not enter into consideration. For FIFO-BS, we have to assume that the service time is known on joining the queue.)

We investigate a single virtual circuit in isolation, taking into account interfering traffic by reducing the service rate, as in [9]. The length of the virtual circuit obviously depends on the system architecture, but examples in the literature [10] suggest a range of four to eleven, with the values at the lower end appearing more often. Following [11], most of the examples will have five nodes.

In summary, the above assumptions ensure independence of service and arrival sample paths for each node. Nodes are linked only be the reduction in traffic afforded by dropping certain packets. These assumptions are necessary for analytic tractability, but as discussed above, will be validated through simulation.

A number of researchers have investigated the effect of packet dropping on a single packet voice multiplexer (e.g., [12, 13, 14], and that of bit dropping on a virtual circuit [15], but the performance of packet dropping in a VC seems to have been considered only in the context of an ARQ-scheme with variable window size [16]. The latter scheme requires acknowledgements, which are not typically used for real-time traffic, and uses feedback, with the concomitant delay problems. Most importantly, packet dropping is performed without regard to the packet's time constraint.

In the following, probability density functions will be denoted by lower case letters, with w and s standing for waiting and system time, respectively. The respective uppercase letters represent the corresponding cumulative distribution functions. An asterisk indicates the one-sided Laplace transform.

3 Packet Loss in Virtual Circuit

This section outlines some of the general relations governing packet loss in a virtual circuit, independent of the particular queueing model used for the individual nodes. A packet is considered lost if it is either discarded by any one of the M nodes it traverses (because it missed the local deadline τ_i) or if the total system (FIFO-BS case) or the waiting time (FIFO-BW) exceeds a set, fixed end-to-end deadline d (see Fig. 1).

Components of Loss Probability

The arrival event A occurs if a packet reaches its destination, i.e., is not dropped by any of the queue controllers along the VC. In a tandem-$M/M/1$ system, A occurs with certainty. The complementary event will be referred to as D and its probability as the ***drop loss***. J_i represents the event that a packet joins queue i (given that it has traversed nodes 1 through $i-1$), while the event of a lost packet, regardless of cause, will be labeled as L. The probability of an event is shown as $P[\cdot]$.

The probability that a packet is not dropped in any of the M (independent) nodes is the destination arrival probability $P[A] = \prod_{i=1}^{M} P[J_i]$ In other words, $P[A]$ does not include the probability that a packet reaching the destination misses its deadline. In computing $P[J_i]$, the reduced traffic due to packets dropped in nodes $0, 1, \ldots, i-1$ needs to be taken into account, where $\lambda_i = \lambda_{i-1} P[J_{i-1}]$.

Thus, given that the end-to-end conditional cumulative distribution of system time for non-discarded packets is $S(d|J) = P[S \le d|J]$[2], the total loss probability $P[L]$, encompassing both drop loss and late loss, can be computed by the law of total probabilities as

$$P[L] = (1 - S(d|A))P[A] + (1 - P[A]) = 1 - S(d)P[A]$$

The first part of the equation shows the contribution of dropped and late packets to the total loss. In the case of voice, $1 - P[L]$ is the fraction of packets that are actually played out at the receiver.

Since each link is assumed to be independent from all others, the conditional density of the end-to-end system time, $s(t|A)$, is derived from the convolution of the single-node conditional densities, $s_i(t|A)$. By the convolution property of the Laplace transform we obtain,

$$s(t|J) = s_1(t|J) * \ldots * s_M(t|J) = \mathcal{L}^{-1}\left\{\prod_{i=1}^{M} s_i^*(s|J)\right\} \quad (1)$$

where $\mathcal{L}^{-1}$ represents the inverse Laplace transform.

4 FIFO-BS: Bounded System Time

In the first of the two congestion control policies considered, a packet joins queue i ($i = 1, \ldots, M$) only if the virtual work, that is, the amount of time it would take to empty the queue with no new arrivals, plus the service time of the arrival is less than the local deadline, τ_i. In other words, the time spent in the system, including service, is bounded from above by τ_i. In this section, the results provided in [17, 18] will be simplified to a form more amenable to computation. Some of the tedious algebra has been relegated to a separate report [19].

[2]The conditioning on J indicates that we only take packets that were not dropped into consideration.

Laplace Transform of the Pdf of the Virtual Waiting Time

The pdf of the waiting time, $w(t)$, can be expressed as $w(t) = Qh(t)$, where the function $h(t)$ is of no further interest for our purposes. The normalization factor Q is defined as

$$Q = (1-\rho)\left[1 - \rho a + \sum_{n=1}^{\infty} \frac{\rho^n (b^n - a)}{(\rho)_n}\right]^{-1}, \rho \ne 1. \quad (2)$$

where $\rho = \lambda/\mu$, $b \equiv e^{-\mu\tau}$, $a \equiv b^{1-\rho} e^{\rho(b-1)}$. $(x)_n = x(x+1)(x+2)\ldots(x+n-1)$ denotes Pochhammer's symbol.

[17] provides the Laplace transform of the vwt. It can be immediately simplified with the identity $\Gamma(i+1) = i!$, yielding in our notation:

$$\begin{aligned} w^*(s) &= w^*(-1) b^{1+s}\left\{1 + (1+s)\sum_{i=0}^{\infty} \rho^i \frac{\Gamma(\rho - s - 1)}{\Gamma(\rho - s + i)}\right\} \\ &\quad - Q(1+s)\sum_{i=0}^{\infty} (\rho b)^i \frac{\Gamma(\rho - s - 1)}{\Gamma(\rho - s + i)} \quad (3) \end{aligned}$$

(Throughout, we are only concerned with the case $\rho \ne 1$.)

By making use of the identities $\Gamma(\rho + i) = \Gamma(\rho)(\rho)_i$ and $\Gamma(\rho - 1) = \frac{\Gamma(\rho)}{\rho - 1}$, we can avoid the evaluation of the gamma function with complex arguments:

$$\begin{aligned} w^*(s) &= Qab^s\left\{1 + \frac{1+s}{\rho - s - 1}\sum_{i=0}^{\infty} \frac{\rho^i}{(\rho - s)_i}\right\} \\ &\quad + Q\frac{1+s}{\rho - s - 1}\sum_{i=0}^{\infty} \frac{(\rho b)^i}{(\rho - s)_i}. \end{aligned}$$

Thus, the transform equation for the density of the system time of a single node is

$$s^*(s) = \frac{1}{P[J]}\frac{\mu}{s+\mu} w^*(s).$$

The technical report [19] also contains closed-form expressions for the cumulative distribution functions of the waiting and system time of a single queue.

5 FIFO-BW: Bounded Waiting Time

Arriving customers join queue i if and only if the virtual work at queue i at the arrival instant is below τ_i. In contrast to the FIFO-BS case discussed above, the service time of the arriving customer does not enter into the admission decision, thus eliminating the bias against long packets present in FIFO-BS. Here, the time a packet spends waiting for service to begin is bounded by τ_i. Results for this system are summarized in [20] (called FIFO-TO there). Among the policies for a single queue studied in that paper, FIFO-BW performed best in the sense of maximizing the throughput of packets meeting their deadlines, suggesting that it may be a good candidate for a tandem system as well. As in the previous section, we will extend the results for a single queue and then proceed to obtain closed-form expressions for the densities and distributions of the waiting time of a tandem system. For brevity, only results for $\rho \neq 1$ will be shown here.

A customer's conditional waiting time is distributed with density

$$w(t|J) = \frac{\nu\rho e^{-\nu t}}{\alpha} u(\tau - t)u(t) + \frac{1-\rho}{\alpha}\delta(t),$$

where $u(t)$ is the unit step function and $\alpha = \rho e^{-\nu\tau}$.

Recognizing $u(\tau - t) = 1 - u(t - \tau)$, the waiting time in the Laplace domain is derived:

$$w^*(s|J) = \frac{1}{\alpha}\left[\frac{\nu\rho}{s+\nu}\left\{1 - e^{-(s+\nu)\tau}\right\} + (1-\rho)\right] \quad (4)$$

Extending the results in [20], the nth moment can be written as

$$E[w^n] = \frac{\rho}{\alpha\nu^n}\left[n! - \Gamma(n+1, \nu\tau)\right] = \frac{n}{\nu}E[w^{n-1}] - \frac{\rho\tau e^{-\nu\tau}}{\alpha}$$

where $\Gamma(n, x)$ is the complementary incomplete Gamma function.

The IMSL routine DINLAP is able to invert Eq. (1) (with waiting time replacing system time), using Eq. (4), with desired accuracy only for small M and losses above 1%. Therefore, it was attempted to obtain tight bounds on the tail of the waiting time distribution by using the moment generating function in the Chernoff bound or higher moments in Bienaymé's inequality [21]. Both produced bounds too loose to judge the benefits of applying queue control via local deadlines. An algebraic inversion of the Laplace-transform expression proved more fruitful. This was possible because unlike the transform for the FIFO-BS case (Eq. (3)), Eq. (4) is rational and thus amenable to closed-form expansion. As a general strategy, the product of sums is expanded into a sum of products and each term in the sum transformed separately.

The multiplicity of poles in the Laplace-domain convolution expression and the potential for analytic and numerical simplification make it expedient to distinguish four cases when inverting the transform. The cases are classified by the dependence of the local deadline τ_i and the traffic parameter $\nu_i = \lambda_i - \mu_i$ on the node index:

1. homogeneous traffic and local deadlines ($\nu_i = \nu_j, \tau_i = \tau_j \;\forall\; i, j$)
2. homogeneous traffic, but arbitrary local deadlines
3. strictly heterogeneous traffic ($\nu_i \neq \nu_j \;\forall\; i \neq j$) and arbitrary local deadlines
4. arbitrary traffic and deadlines, i.e., $\nu_i = \nu_j$ is allowed

The closed-form expressions for calculating losses for these four cases appear in the appendix.

Performance Evaluation: Numerical Examples and Validation

Having derived expressions for the cdf of the end-to-end system and waiting times, we can now compare the performance of the suggested queue control mechanisms. Limitations of space will restrict our discussion to the FIFO-BW case. In this section, the end-to-end deadline d is fixed so that the packet loss equals a given value when no queue control is applied ("uncontrolled loss").

We first consider the case of homogeneous local deadlines, i.e., $\tau_i = \tau_j = \tau$. It is clear that the optimal τ has to fall between d/M and d since for values of τ less than or equal to d/M, all packets which are not dropped will make their end-to-end deadline; For $\tau \geq d$, any packet dropped at an intermediate node would be late on arrival at the destination.

Since relatively high load and loss should expose the deficiency of ignoring the non-Poisson nature of input streams, the first example uses the parameter set $M = 5$, $\mu = 1$, $\lambda_1 = 0.8$ and $d = 40.47$, resulting in an uncontrolled loss of 5%. The simulation maintains Kleinrock's independence assumption, but not the Poisson nature of the interior arrival streams. The simulation was terminated once the 95%-confidence interval halfwidths computed by spectral estimation [22] and the regenerative method were both less than 10% of the point estimate. Depending on τ, a simulation run consisted of between one and five million packets. An initial transient period estimated at 3000 packets was discarded.[3]

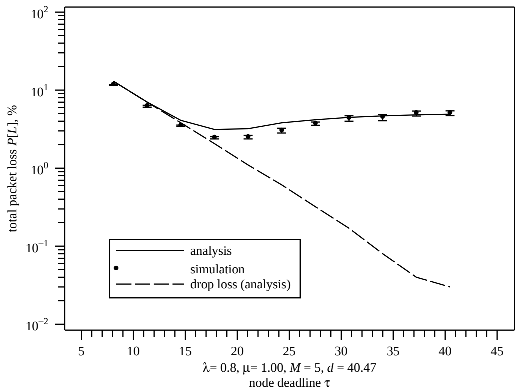

Figure 2: Total and drop loss; analysis and simulation

Fig. 2 compares analytical and simulation results for the first example, plotting the total loss, composed of packets dropped in intermediate nodes and packets missing their end-to-end deadline, as a function of the local deadline τ. The horizontal line at 5% loss indicates the asymptotic uncontrolled case with $\tau = \infty$. The graph shows that the analytical results overestimate end-to-end losses, with simulation and analysis agreeing more closely as τ increases towards infinity and the system approaches a tandem-M/M/1 queueing system. Fortunately, the optimal nodal deadlines seem to fall in the same neighborhood for both analysis and simulation. For this set of parameters, the FIFO-BW dropping mechanism reduces total losses from 5% to 3.1% according to the analysis, to 2.4% according to the simulation. From Fig. 2 we can also conclude that the drop losses incurred by tight deadlines are not compensated for by the reduction in downstream traffic. Since the performance varies little in the region from $\tau \approx 17$ to $\tau \approx 21$, a looser deadline in that range is preferable, providing a margin of safety to compensate for errors in the estimation of the system parameters.

The dashed line traces late loss, showing that for very tight deadlines, almost all packets that have not been dropped at intermediate nodes make their deadline.

[3] The estimate was based on the diffusion approximation of the G/G/1 queue, see [23][24].

Optimal Local Deadlines

While the experiments above used the same τ for all nodes, it seems reasonable to assume that the lower traffic impacting downstream nodes could make heterogeneous deadlines advantageous. To test this hypothesis, an M-dimensional simplex method was used to find a set of τ_i's minimizing the overall loss, taking downstream traffic reductions into account. The optimal homogeneous τ was used as a starting point to minimize the probability of trapping the optimization in a local optimum. From the examples reported in [19], it can be concluded that even for high loads ($\rho = 0.9$) and short VCs such as $M = 2$, heterogeneous deadline yield less than 1% improvement over using a homogeneous τ.

Since the expressions for homogeneous traffic and deadlines are algebraically much simpler and numerically more stable, we examined the performance penalty incurred by ignoring downstream traffic reductions. Clearly, for homogeneous traffic only homogeneous deadlines can be optimal. As the examples in Fig. 3 and 4 demonstrate, the penalty increases with higher loads, higher uncontrolled losses and longer virtual circuits. (In the figures, the horizontal lines denote the asymptotic losses for the uncontrolled system.)

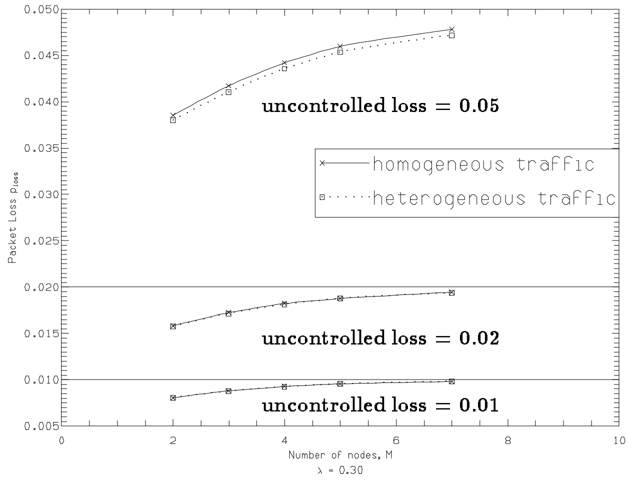

Figure 3: Total packet loss vs. number of nodes, for optimal homogeneous deadlines based on homogeneous or decreasing traffic, $\lambda = 0.30$

Range of Effectiveness

In the course of investigation it became clear that the proposed queue control is advantageous only for a range of parameter values. The contour plot of Fig. 5, containing lines of equal ratio of controlled to uncontrolled loss vs. length of VC, M, and traffic intensity ρ for uncontrolled losses of 10^{-5}, 10^{-3}, 10^{-2} and $5{\cdot}10^{-2}$, shows that the ratio decreases (performance gains realized by queue control increase) with increasing loads, shorter virtual circuits, higher uncontrolled losses and higher loads. For example, at an uncontrolled loss of 5%, overall losses can drop to as little as 40% of the uncontrolled losses for a three-node circuit operating at a load of $\rho = 0.9$. On the other hand, at uncontrolled losses of less than 10^{-3}, even loads of $\rho = 0.9$ and the minimal circuit length of 2 does not push the controlled loss much below 75% of the uncontrolled loss. The parameter region of worthwhile gains agrees well with the operating region of overload control for packet voice systems. However, for the uncontrolled losses found in video applications, little gain can be expected. It should be noted that the results shown in the graphs are conservative estimates since they do not take into account the global traffic reduction afforded by having all virtual circuits apply the control policy, especially significant under loads of 0.8 and above. Also, the simulation results discussed earlier show that actual performance may be slightly better than the analysis predicts.

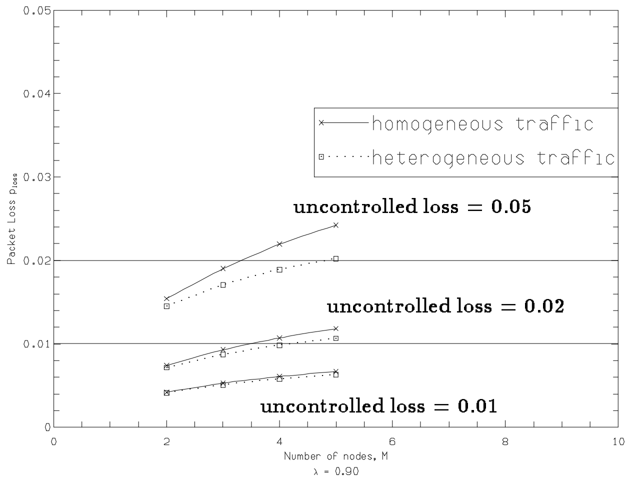

Figure 4: Total packet loss vs. number of nodes, for optimal homogeneous deadlines based on homogeneous or decreasing traffic, $\lambda = 0.90$

Investigations showed that the optimal local deadlines are relatively insensitive to the number of nodes in the network and, to a lesser extent, the VC load. Also, no obvious relationship seems to exists between drop loss and late loss at the optimal operating point.

6 Congestion Control and Robust Local Deadlines

In practical networks, it may not be feasible to measure traffic accurately and adjust the local deadlines accordingly, even with a table-lookup mechanism instead of on-line optimization. However, a static control scenario such as the following would be possible. At call setup, the end-to-end deadline (and, therefore, the playout buffer delay) is set so that the desired loss under non-overload conditions can be met. The local deadlines are then established by table lookup based on the length of the virtual circuit and a tradeoff between overload protection and additional loss incurred under lighter loads. For each packet, the queue controller determines the applicable τ, measured in transmission units, from a look-up table and compares it to the current number of transmission units queued, including those of the packet in transmission. (Here we assume a fixed-rate channel.) These operations are readily implementable in hardware at packet rate.

An example illustrates the approach. Suppose that a 5-node connection can tolerate a loss of 1% without noticeable

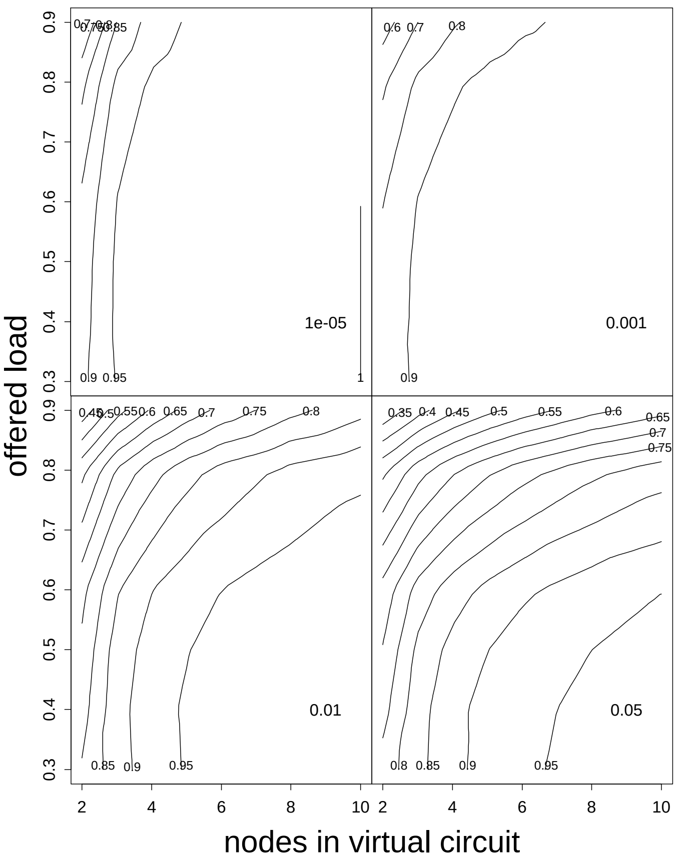

Figure 5: Best achievable ratio of controlled (FIFO-BW) to uncontrolled (M/M/1) loss for homogeneous traffic and deadlines; uncontrolled losses of 10^{-5}, 0.001, 0.01 and 0.05

grade-of-service degradation. A call is allowed into the network only if the average load does not exceed $\rho = 0.8$ at the time of call setup. Correspondingly, the playout buffer is set up for an end-to-end deadline $d = 52.7$. The effect of several choices for τ is shown in Fig. 6. A first attempt for τ uses the value optimal for a load of 0.9, yielding $\tau = 18.3$. Assume that during the call the network load rises to 0.9. If no queue control were applied, the losses would reach an intolerable 32%. With the queue control, $\tau = 18.3$, losses are reduced to 7%, which an interpolation method might be able to cope with. However, this deadline is unduly restrictive for lower loads. For example, at $\rho = 0.8$, losses are twice that of applying no control. Tightening the deadline to $\tau = 15$ does not improve overload performance, but further hampers GOS at $\rho = 0.8$, pushing losses to 3.5%. A more conservative choice, $\tau = 23.3$, ensures that losses at design load and below never exceed 1%, while still limiting overload losses to 9.2%. If load should rise momentarily to $\rho = 0.95$, the overload mechanism will cut losses from an uncontrolled 80% to 20%. As mentioned at the end of section 5, actual network performance will most likely be better as all VCs can be assumed to apply similar control mechanisms.

Figure 7 compares the goodput, defined as $\lambda(1 - P[L])$, of the uncontrolled and controlled system. Beyond a certain load, $\rho = 0.83$ in the example, the goodput for the uncontrolled system actually decreases with increasing load.

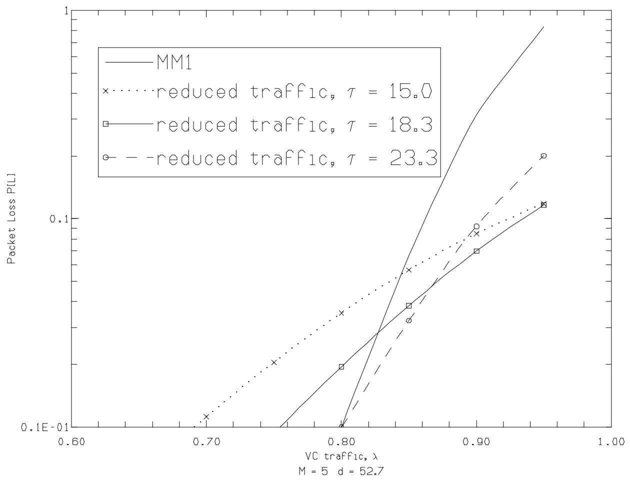

Figure 6: Comparison of overload performance of FIFO-BW to that of uncontrolled system; overload region

Optimal random discarding as proposed in [25] can hold the goodput at the peak value even under overload, as indicated by the horizontal line in the figure. It throttles the input traffic to the point of optimal goodput by randomly discarding packets at the source. Optimal random discarding requires, however, on-line traffic or gradient estimation ($\partial P[D]/\partial\lambda$) to determine the fraction of packets to be discarded. Note also that if we allow the deadline to be adjusted on-line according to the current traffic, FIFO-BW offers a higher goodput than optimal random discarding.

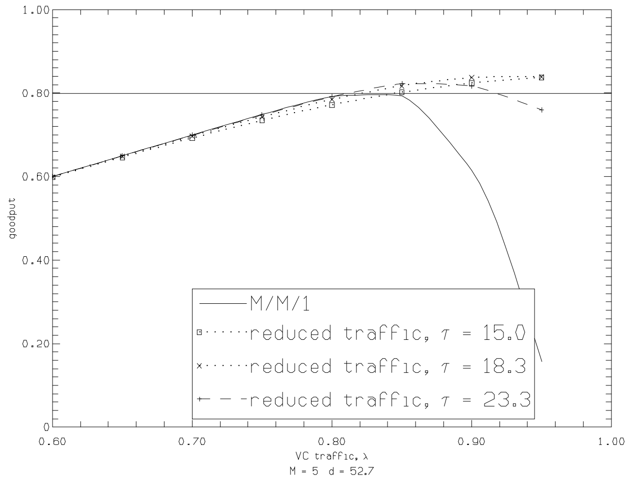

Figure 7: Comparison of goodput: tandem-M/M/1 vs. FIFO-BW with various local deadlines, $M = 5$, $\mu = 1$

7 Conclusion and Future Work

In this paper, we have presented an analytic model for evaluating the loss performance of two queue control schemes, based on rejecting packets at intermediate nodes in a virtual circuit. It was shown that the fraction of lost packets could be reduced substantially for suitable traffic types and virtual circuit size. Even if the exact load is not known, a static overload control mechanism can ameliorate the effects of temporary overload situations. In related work [19], it is shown that analysis technique and performance gains carry over to the case of fixed packet sizes.

Other mechanisms, object of future work, include studying control mechanisms based on minimum-laxity scheduling, where laxity can be defined either as the time to expiration of the end-to-end deadline or as the time to extinction divided by the number of hops still to be traversed. History-based discarding, which admits a packet ohly if the time spent waiting upstream plus the virtual wait falls below a given threshold, offers another possibility, albeit with higher processing overhead. Queue disciplines that favor packets with short waiting times, for example LIFO, also merit investigation and comparison. Finally, the effect of fixed-size packets deserves further inquiry.

Appendix

In this appendix, closed-form expresssions for the waiting-time density and distribution function for the FIFO-BW system will be presented. Homogeneous and heterogeneous traffic leads to Mth order and first-order poles in the Laplace transform expression, respectively, and thus requires separate treatment. A special case for homogeneous traffic simply reduces the number of terms in the expansion. Results for the first three cases enumerated in section 5 will be presented; the fourth case of arbitrary τ_i and ν_i is conceptually similar, but incurs another order-of-magnitude increase in complexity not compensated for by increased insight.

For the first case of homogeneous traffic and deadlines, the closed-form expression for the density can be derived from the Laplace transform as

$$\mathcal{L}^{-1}\left[\frac{1}{\alpha}\frac{\nu\rho}{s+\nu}(1-e^{-(s+\nu)\tau})+\frac{1-\rho}{\alpha}\right]^M$$

$$=\mathcal{L}^{-1}\left\{\frac{1}{\alpha^M}\sum_{k+l+m=M}\binom{M}{k,l,m}e^{-l\nu\tau}(\nu\rho)^{k+l}\right.$$

$$\left.\cdot(1-\rho)^m(-1)^l\frac{e^{-l\tau s}}{(s+\nu)^{k+l}}\right\}$$

For $k+l>0$, the fraction containing terms in s can be readily inverted as

$$\frac{(t-l\tau)^{k+l-1}}{(k+l-1)!}e^{-\nu(t-l\tau)}u(t-l\tau)$$

For $k+l=0$, that is, $k=l=0$, the summand reduces to $(1-\rho)^m$. The cdf can be computed by integrating terms over the interval from $l\tau$ to $d-l\tau$.

The second case allows for local deadlines that depend on the node index. The Laplace transform can be rewritten as

$$\prod_{i=1}^{M}\frac{1}{\alpha^i}\left[\frac{\nu\rho}{s+\nu}(1-e^{-(s+\nu)\tau})+(1-\rho)\right]$$

$$=\left[\prod_{i=1}^{M}\frac{1}{\alpha_i}\right]\sum_{k=0}^{M}\sum_{\vec{l}:|\vec{l}|=M-k}\left[\frac{\nu\rho}{s+\nu}+(1-\rho)\right]^k(-1)^{M-k}$$

$$\cdot\left[\frac{\nu\rho}{s+\nu}\right]^{M-k}\exp\left(-(s+\nu)\sum_{i=1}^{M}l_i\tau_i\right)$$

where the binary vector $\vec{l}$ has components l_i, restricted to the values zero or one. $|\vec{l}|$ denotes the number of ones in $\vec{l}$.

Each term in the inner sum can be readily inverted for a given k as

$$(-1)^{M-k}e^{-\nu\tau'}\sum_{j=0}^{k}\binom{k}{j}(1-\rho)^{k-j}(\nu\rho)^{j+M-k}A_{k,j}$$

where

$$A_{k,j}=\frac{(t-\tau')^{M+j-k-1}}{(M+j-k-1)!}e^{-\nu(t-\tau')}u(t-\tau')$$

for $j+M-k>0$ and $A_{k,j}=\delta(t-\tau')$ for $j=k-M$. Here, $\tau'=\sum_i l_i\tau_i$. The number of terms depends on the value of τ and t and is less than $2^M(1+M/2)$.

The third case allows for traffic that depends on the node index; the condition of strict heterogeneity in λ_i insures single poles. The PDF is epressed in the time-domain as

$$w_M(t)=\sum_{k=1}^{2^M-1}\left[\prod_{i=1}^{M}\frac{1}{\alpha_i}\right]\left[\prod_{j=1}^{M}r_j^{1-b(k)_j}(\rho_j\nu_j)^{b(k)_j}\right]$$

$$\sum_{j\in b(k)}\frac{\sum_{m=0}^{2^{|b(k)|}-1}e^{-\beta_m-\nu_j(t-\tau'_m)}u(t-\tau'_m)}{\prod_{\substack{i\in b(k)\\ i\neq j}}\nu_i-\nu_j}+\prod_{n=1}^{M}r_n\delta(t).$$

$b(k)_i\in\{0,1\}$ stands for the ith digit of k written as a base-2 (binary) number. For brevity, $i\in b(k)$ denotes all indices i such that $b(k)_i$ is 1. The computation of the quantities β_m and τ'_m is based on the expansion of the product of the $|l_k|$ terms of the form $1-\exp(-(s+\nu_i)\tau_i)$. Define the mapping function $h(k)$, whose ith component $h(k)_i$ is the index of the ith one bit in $b(k)$. Then,

$$\tau'_m=\sum_{i=1,i\in b(i)}^{|b(k)|}\tau_{h(k)_i}$$

and

$$\beta_m=\sum_{i=1,i\in b(i)}^{|b(k)|}\nu_{h(k)_i}\tau_{h(k)_i}.$$

The distribution is readily computed by replacing the terms in the summation over m by

$$\frac{1}{\nu_j}\left[1-e^{-\nu_j(t-\tau'_m)}\right]e^{-\beta_m}u(t-\tau'_m)$$

This case requires the evaluation of at most $2M\cdot3^{M-1}$ terms, with the exact value depending on t. The large number of terms for high M (above five), differing in absolute value by several orders of magnitude, forces use of quadruple precision arithmetic and magnitude-sorted summation to control round-off errors.

References

[1] N.S. Jayant. Effects of packet losses on waveform-coded speech. In *Proceedings of the Fifth International Conference on Computer Communications*, pages 275–280, Atlanta, Georgia, October 1980. IEEE.

[2] Willem Verbiest, Luc Pinnoo, and Bart Voeten. The impact of the atm concept on video coding. *IEEE Journal on Selected Areas in Communications*, JSAC-6:1623–1632, December 1988.

[3] K.K. Ramakrishnan and Raj Jain. A binary feedback scheme for congestion avoidance in computer networks with a connectionless network layer. In *Proceedings of the 1988 SIGCOMM Symposium on Communications Architectures and Protocols*, pages 303–313, Stanford, California, August 1988. ACM.

[4] Kotikalpudi Sriram and Ward Whitt. Characterizing superposition arrival processes in packet multiplexers for voice and data. *IEEE Journal on Selected Areas of Communications*, SAC-4(6):833–846, September 1986.

[5] Harry Heffes and David M. Lucantoni. A markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. *IEEE Journal on Selected Areas of Communications*, SAC-4(6):856–867, September 1986.

[6] K. Sriram and D. M. Lucantoni. Traffic smoothing effects of bit dropping in a packet voice multiplexer. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, pages 759–769, New Orleans, March 1988. IEEE.

[7] S. Dravida and K. Sriram. End-to-end performance models for variable bit rate voice over tandem links in packet networks. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, pages 1089–1097, Ottawa, Canada, April 1989. IEEE.

[8] San-qi Li and Jon W. Mark. Simulation study of a network of voice/data integrated TDMs. *IEEE Transactions on Communications*, COM-36(1):126–132, January 1988.

[9] Amit Bhargava, James F. Kurose, Don Towsley, and Guy van Lamport. Performance comparison of error control schemes in high speed computer communication networks. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, New Orleans, April 1988. IEEE.

[10] Kenzo Takahashi, Tadahiro Yokoi, and Yutaka Yamamoto. Communications quality analysis for ATM networks. In *Conference Record of the International Conference on Communications (ICC)*, pages 0423–0427 (13.6), Boston, June 1989. IEEE.

[11] M. De Prycker. Impact of data communication on ATM. In *Conference Record of the International Conference on Communications (ICC)*, pages 0705–0712 (22.4), Boston, June 1989. IEEE.

[12] Nanying Yin, San-qi Li, and Thomas E. Stern. Congestion control for packet voice by selective packet discarding. In *Proceedings of Globecom 1987*, pages 1782–1786, Tokyo, November 1987. IEEE.

[13] Chin Yuan and John A. Silvester. Queueing analysis of delay constrained voice traffic in a packet switching system. *IEEE Journal on Selected Areas of Communications*, SAC-7(5):729–738, June 1989.

[14] Byung K. Kim and Don Towsley. Dynamic flow control protocols for packet-switching multiplexers serving real-time multipacket messages. *IEEE Transactions on Communications*, COM-34(4):348–356, April 1986.

[15] S. Dravida and K. Sriram. End-to-end performance models for variable bit rate voice over tandem links in packet networks. *IEEE Journal on Selected Areas of Communications*, SAC-7(5):718–728, June 1989.

[16] Daniel T. D. Luan and David M. Lucantoni. Throughput analysis of a window-based flow control subject to bandwidth management. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, pages 0411–0417, New Orleans, March 1988. IEEE.

[17] J. W. Cohen. Single server queues with restricted accessibility. *Journal of Engineering Mathematics*, 3(4):265–284, October 1969.

[18] Bezalel Gavish and Paul J. Schweitzer. The markovian queue with bounded waiting time. *Management Science*, 23(12):1349–1357, August 1977.

[19] Henning Schulzrinne, James F. Kurose, and Don Towsley. Congestion control for real-time traffic in high-speed networks. Technical Report TR89-92, Department of Computer Science, University of Massachusetts, Amherst, MA, 1989.

[20] Bharat T. Doshi and Harry Heffes. Overload performance of several processor queueing disciplines for the M/M/1 queue. *IEEE Transactions on Communications*, COM-34(6):538–546, June 1986.

[21] Athanasios Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill Book Company, New York, 2nd edition, 1984.

[22] Philip Heidelberger and Peter D. Welch. A spectral method for confidence interval generation and run length control in simulations. *Communications of the ACM*, 24(4):233–245, April 1981.

[23] Hisashi Kobayashi. Application of the diffusion approximation to queueing networks II: Nonequilibrium distributions and applications to computer modeling. *Journal of the ACM*, 21(3):459–469, July 1974.

[24] Hisashi Kobayashi. *Modeling and Analysis: An Introduction to System Performance Evaluation Methodology.* Addison-Wesley, Reading, Massachusetts, 1981.

[25] Christos G. Cassandras, Michelle Hruby Kallmes, and Don Towsley. Optimal routing and flow control in networks with real-time traffic. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, pages 784–791, Ottawa, Canada, April 1989. IEEE.