On Router Control for Congestion Avoidance

Marine Auvert*, Håkan Hjalmarsson*, Karl H. Johansson*, and Gunnar Karlsson†

∗ Dept. of Signals, Sensors & Systems, Royal Institute of Technology, kallej@s3.kth.se
† Dept. of Microelectronics and Information Technology, Royal Institute of Technology, gk@imit.kth.se.

Abstract

This short paper deals with active queue management for computer networks. The goal is to develop control mechanisms for routers in heterogeneous networks that reduce traffic fluctuations. The proposed control strategy operates with local information (such as estimated arrival rates) and actively use the buffers to smooth traffic, and thus it avoids the buildup and propagation of traffic bursts.

1 Introduction

A continually increasing use of the Internet and its widening set of offered communication services increase the demands on the control of the network. The control objective is often to improve traffic throughput and to accommodate different service demands better. Traffic congestion in statistically multiplexed computer networks is a major concern. Today's congestion control for the Internet is implemented as an end-to-end protocol, namely the transport control protocol (TCP) [1, 2, 3, 4]. The service policy in routers is predominantly work conserving: packets are sent out at the full rate of the output link. This allows congestion to propagate to downstream routers and might create a wide area of congestion. There are (at least) two reasons for why it is difficult to attenuate this behavior with existing congestion control mechanisms such as TCP: (1) congested network traffic has often both long-range and short-range dependencies, and is therefore hard to control at the single time scale of the end-to-end protocol; and (2) the actual control actuation of the end-to-end protocol on the congested area is implicit and complex, since there are several protocol instances that interact. Park [5] points out the need for control strategies working over longer time scales compared to what is the case today. Recent models capturing the self-similar nature of the traffic [6, 7] open up new possibilities for the use of model-based control. Our idea is to use feedback control to smooth the pattern of departing packets of individual routers. The importance of smooth network flows has been emphasized by Massoulié and Roberts [8].

Initial work on router control strategies for reducing traffic fluctuations is presented in this paper. The main contribution is to introduce a particular idea based on a so called virtual service rate, which can be implemented in today's routers. Some preliminary simulation results are presented. Smoothing has been used in network control before, but then on an end-to-end control basis. For example, smoothing has been proposed for reducing bit rate variability for video transmission [9], which seems not to have been studied from a feedback control point of view.

2 Congestion Control

Traffic congestion in communication networks can be reduced for example by introduce new control mechanisms in the routers or by conventional end-to-end control protocols such as TCP.

2.1 Router Control

By actively using the buffers in the routers to smoothen traffic, it is possible to improve the overall efficiency of a packet-switched network. If each router sends out packets at the full rate, the utilization of the individual link is maximized at each time instant. Such a local optimization might not be globally optimal, however. This is due to limited buffer capacity in each router of the network. If one of the routers overflows, and thus packets are dropped, retransmission leads to decreased overall throughput of the network. A better strategy is then to spread out the temporary storage of packets to several buffers. A globally optimal solution needs a centralized control structure where information about the traffic throughout the network is transferred and where the control decisions are taken. This is not realistic for networks such as the Internet. However, it is possible to improve the performance also with local information and control.

Although routers forward packets at the full rate of the output links, it would be possible to use the output buffers to smoothen the pattern of departing packets to avoid aggregation of bursts. Each packet has, of course, to be sent at the bit rate of the link. But after a packet has been sent, the server could wait before serving the next packet. The length of the introduced idle period could depend on such dynamic information as the fullness of the buffer, the past departure pattern, and the predicted arrival pattern. The hope is that the regulated traffic will make congestion more rare, which in turn means that losses are reduced and delays are less varying. This would in turn lead to higher throughput and less capacity wasted on retransmissions. The control law should determine the lengths of the idle periods between packets, so that the variance of the departure process is minimized.

2.2 End-to-End Control

A fundamental question is how much can be gained by using feedback information to regulate flows in a network. The end-to-end control in TCP uses acknowledgements from the receivers as the only explicit feedback information whereas the available bit rate (ABR) service in an ATM network relies on information from the switches. Similar feedback is suggested also for the Internet protocol by means of explicit congestion notification [10]. Explicit feedback increases both the network complexity and the requirements on the nodes. We therefore believe that it is of interest to quantitatively assess the performance benefits of explicit feedback information from the network. This would for instance tell how conservative it is to separate the traffic flow control problem into local controls in the nodes and end-to-end control. The ideal objective is to produce curves of performance gain vs. feedback bit rate. Work in this area is yet sparse [11]. The issue of what type of feedback information to use also needs to be addressed. An interesting suggestion by Kelly and co-workers [12] is that the shadow prices for resources should be used. It is critical to consider the delay dependence in this analysis since the benefits of feedback diminish as the network delays increase.

3 Network Flaps

It is possible to influence the distribution of the outgoing traffic from a router by adding an idle time period to the service time.

3.1 The Idea

Packets are sent out at the maximum transmission rate in routers today. This leads to that traffic bursts propagate downstream in the network. To avoid the accumulation of bursty traffic, which may create a congested area of the network, the proposed idea is to smooth out the distribution of the departure times of the routers. This idea can be compared to the action of rubber flaps on a conveyer belt, where the flaps even out the distribution of the goods. Therefore, we call the mechanism added to the routers *network flaps*.

3.2 Prototype System

Consider a simple subnetwork consisting of two routers in series modeled as the queueing system in Figure 1. This is

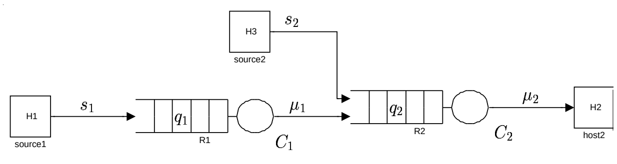

Figure 1: Prototype network with two buffers.

the simplest configuration to illustrate the idea of network flaps and is based on a bottleneck topology. The queues contain $q_i(t)$ packets, $i = 1,2$, at time instant t, where $0 \leq q_i(t) \leq q_{i\max}$. The incoming traffic comes from two virtual sources having time varying source rates $s_i(t)$, $i = 1,2$. The output link capacities are fixed and equal to C_1 and C_2, respectively. By assuming that $C_1 > C_2$, the second link represents a bottleneck, and possibly the two flows aggregating at the second buffer create an overflow. The output rates $\mu_i(t)$, $i = 1,2$, are limited by the corresponding capacity. The arrival rates of the two buffers are $\lambda_1(t) = s_1(t)$ and $\lambda_2(t) = s_2(t) + \mu_1(t)$.

It is common in most applications to let the departure rate attain its maximum value $\mu_i = \mu_{i\max}$ as soon as there are packets to send, i.e., packets in buffer i are sent out at maximum rate. This leads to maximal utilization of the link at each time instant. When the traffic is bursty and the load of the network is high, this approach however may lead to performance losses. The evolution of the queue lengths are shown in Figure 2 for the system in Figure 1, when the inter-arrival times for the source rates have exponential distribution and $\mu_i = \mu_{i\max}$. The first queue never builds up due to that the capacity C_1 is basically high enough to let all packets be served as soon as they arrive. The second link clearly creates a bottleneck since the capacity C_2 is not high enough. Therefore q_2 grows. In this particular example $q_{1\max} = q_{2\max} = 50$, so the bottleneck leads to loss of packets, as shown in the lower plot of Figure 2.

3.3 Virtual Service Time

We suggest that bursty traffic should be smoothed by delaying packets. This can be interpreted as considering the output rate μ_i as a control variable, which belongs to the interval $[0, \mu_{i\max}]$. An equivalent interpretation is to introduce the notion of a *virtual service time* T_v, which is equal to the sum of the actual service time T_s and a controlled delay T_c, i.e., $T_v = T_s + T_c$. Of course, the control signal $T_c \geq 0$ only takes positive values, and for a conventional router we have $T_c \equiv 0$.

A simple illustration of the effect of a virtual service time is to let T_c be a positive constant. The result of this is shown in Figure 3. Note how the first buffer is utilized to store packets, so that the total number of dropped packets in this case is less than in Figure 2. It is reasonable to let T_c be chosen such that T_v is approximately constant over a

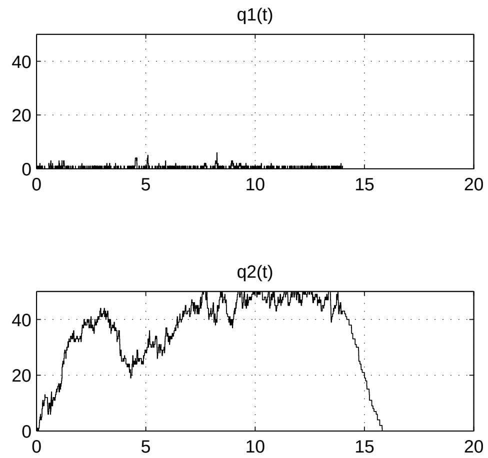

Figure 2: Simulation of the system in Figure 1. The departure rates are limited by the fixed capacities of the links.

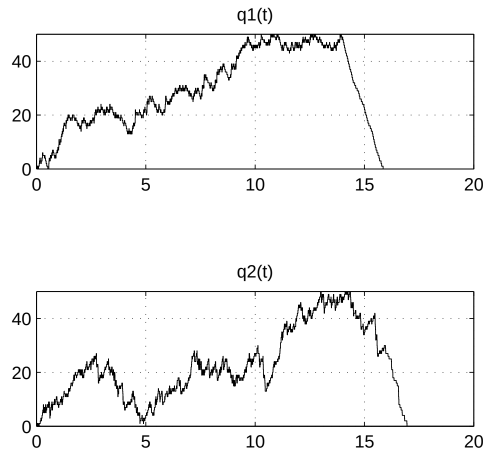

Figure 3: Introducing a virtual service time distributes the packet load over both buffers and reduces the total loss of packets.

certain time interval. A natural choice is to let this constant be equal to the (estimated) mean of the inter-arrival times.

Note that the control strategy above should be applied when the load of the (sub)network is high. When the load is low, it is always optimal to use the maximum sending rate. This is captured by the switched control strategy with three states illustrated in Figure 4. We assume that only local information is available to the router, i.e., the router may estimate the arrival rate λ and measure the queue length q. If λ is small, it can be interpreted as if there is no risk for congestion and thus $T_c = 0$. If λ is large and $q \ll q_{\max}$, one should use the control actuation available in T_c to smooth the distribution. If both λ and q are large, one should use less control actuation to avoid the own queue to become full.

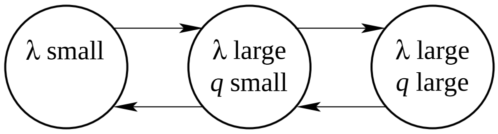

Figure 4: Finite state machine that illustrates the three states of a router controller with network flaps. The controller state depends on the arrival rate λ and the queue length q.

3.4 Preliminary Analysis

The following heuristic analysis gives hint on when network flaps are useful.

One bottleneck: Consider a configuration consisting of two routers in series. Assume Router 1 has an infinite size buffer ($q_{1\max} = \infty$), while the buffer of Router 2 is finite ($q_{2\max} < \infty$). Assume the link connecting the two routers has infinite capacity, but that the outgoing link from the second router has finite capacity C packets/s. Router 1 can be seen as the aggregate of a complete network. The downstream router and its outgoing link will act as a bottleneck.

Note first that the average transmission rate is limited by C and, hence, that if the sender transmits an even flow of F packets/s, where $F < C$, then there is no gain in using network flaps. Under this condition there is no extra delay introduced by the network.

Let now the transmitted flow be bursty. We will consider the extreme case that the sender transmitts FT packets in one second every T seconds, with T being large. The upstream router will buffer these packets and then subsequently deliver them to the downstream router.

Using network flaps, ideally the transmission can be made controlled such that no packets are lost in Router 2. Hence, the receiver will receive the packets with the rate C packets/s and the average packet delay will be $FT/2C$. Clearly, this is optimal.

Without network flaps, Router 1 will (with its infinite buffer size) deliver all packets at once to Router 2. If $FT > q_{2\max}$, this will imply that $FT - q_{2\max}$ packets will be lost. Define LAT, the loss alert time, to be the time it takes for the sender to become aware of the packet loss and to initiate a retransmission of lost packets. Once the sender is notified of the packet loss, the lost packets are immediately sent out. This means that, as long as the bottleneck buffer has not become empty so that the router has become idle, no loss in performance is obtained compared to using network flaps. If, however, $LAT > q_{2\max}/C$, Router 2 will become empty before the re-sent packets arrive and this will cause a loss of performance. It is easy to see that under the assumption that FT is much larger than $q_{2\max}$ (i.e., the peaks of the bursts are sufficiently high), the average packet delay in the network becomes

$$\frac{FT}{2C} \cdot \frac{LAT}{B/C}$$

which obviously is larger than the optimal packet delay $FT/(2C)$ obtained by network flaps. Note that the loss alert time relative to the bottleneck delay B/C plays a role for when network flaps are useful. The loss alert time can be approximated by the round-trip time.

Two bottlenecks: We will now show that network flaps may increase the capacity of a network. Consider a network with three routers. Router 1 and Router 2 are connected to Router 3 by two separate links with capacity C_1 and C_2, respectively. The outgoing link of Router 3 has capacity C_3. Suppose that $C_1 + C_2 > C_3$, so that the downstream router with corresponding outgoing link is the bottleneck when both Router 1 and Router 2 are active. Suppose also that $C_1 < C_3$, so that the first link is the bottleneck when only Router 1 is active. Consider again a bursty scenario, when suddenly senders transmit packages above the network capacity.

Using network flaps, ideally Routers 1 and 2 would adapt their rates to their fair shares (relative their rates) $C_3C_1/(C_1+C_2)$ and $C_3C_2/(C_1+C_2)$, respectively, in order to not overload the bottleneck. The overall rate would then be C_3.

Now, if the upstream routers transmit at full rates but Router 3 favours Router 2, so that this router is allocated more than its fair share of the capacity, then the transmission from Router 2 will finish before the transmission from Router 1. In the instant Router 1 will become the bottleneck, the transmission rate will be reduced from C_3 to C_1. The worst-case scenario is when Router 3 first transmits all packets from Router 2 before the packets from Router 1 are considered. In this case the overall capacity of the network becomes $C_3(C_1 + C_2)/(C_3 + C_2)$, which clearly is less than C_3.

3.5 Simulations

Consider the system in Figure 1 again. Let the inter-arrival times for the source rates s_i, $i = 1, 2$, be Pareto distributed with distribution function

$$F(x) = P(X \leq x) = 1 - (\alpha/x)^{\beta}, \qquad x \geq \alpha$$

where α, β are positive constants. Figure 5 shows the response of the system when no control is applied. Note how the burstiness of the arrival rates affects the queue lengths. Also, note that the second queue fills up quite early, although the first is hardly used.

We may now ask if by applying network flaps, it is possible to smooth the traffic. A naive approach is to simply use linear control, based on the queue lengths. Figure 6 shows a control structure where the output rate μ_1 of the first router is controlled using linear feedback from the difference between the queue length q_1 and a set-point value $q_{1\mathrm{sp}}$. The gain of the controller is chosen according to a linear quadratic control design (e.g., [13]). Figure 7 shows

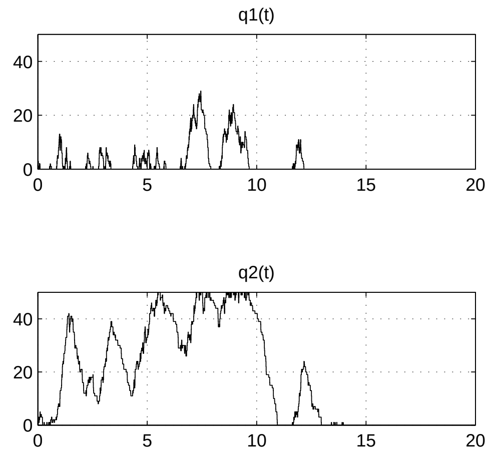

Figure 5: Simulation of protype network in Figure 1 with (bursty) Pareto distributed sources. No control is applied, but the departure rates are maximal.

the result of this control. Note how the first queue builds up. This reduces the load on the second.

Figure 8 shows a control structure where the output rate μ_1 of the first router is controlled using linear feedback from $q_2(t) - q_{2\mathrm{sp}}$. Again the gain of the controller is chosen according to a linear quadratic control design. Figure 9 shows the result of this control.

Both the decentralized and the centralized control structure improve the performance (lower the total number of dropped packets). The second strategy gives the better behavior, but note that it requires the sending of information from the second router to the first (which might not be desired from an implementation point of view).

4 Conclusions

We have discussed router control for congestion avoidance. Our aim is to address the problem by a combination of router controls and end-to-end controls. In this paper we focus on the former. Some promising initial results was shown for the network flaps, our proposal for router control. However, more investigation is needed. We believe that this research area presents several challenging problems from a control theory as well as from a computer network point of view and we hope to present more conclusive results in the near future.

References

[1] V. Jacobson. Congestion avoidance and control. In *Proc. of SIGCOMM*, pages 314–329, 1988.

[2] J. Walrand and P. Varaiya. *High-performance communication networks*. Morgan Kaufmann, 2nd edition, 2000.

[3] L. L. Peterson and B. S. Davie. *Computer networks: a systems approach*. Morgan Kaufmann, 2nd edition, 2000.

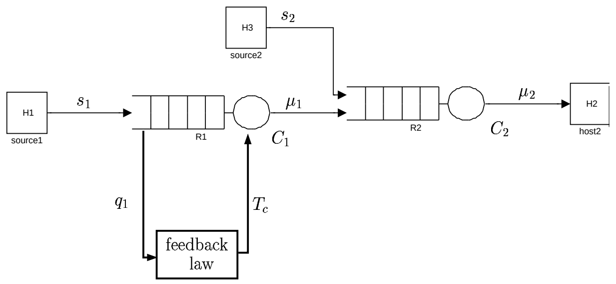

Figure 6: Control structure where the departure rate μ_1 of the first router is controlled using information about q_1. This corresponds to decentralized control, where only local information is used in the controller implementation.

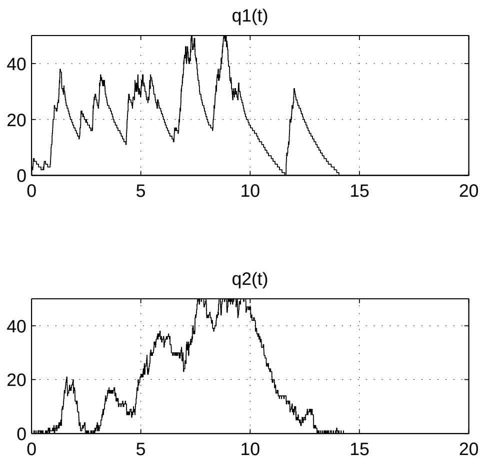

Figure 7: Simulation with control of the departure rate μ_1 of the first router using information about the queue length q_1. The control structure is shown in Figure 6.

[4] S. Mascolo. Congestion control in high-speed communication networks using the Smith principle. *Automatica*, 35(12):1921–1935, 1999.

[5] K. Park. Future directions and open problems in performance evaluation and control of self-similar network traffic. In *Self-Similar Network Traffic and Performance Evaluation*. Wiley-Interscience, 2000.

[6] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, 1994.

[7] V. Paxson and S. Floyd. Wide-area traffic: the failure of Poisson modeling. *IEEE/ACM Trans. on Networking*, 3(3):226–244, 1995.

[8] L. Massoulié and J. Roberts. Bandwidth sharing: objectives and algorithms. In *Proc. IEEE INFOCOM*, pages 1395–1403, New York, NY, March 1999.

[9] Z. Jiang and L. Kleinrock. A general optimal video smoothing algorithm. In *Proc. of IEEE INFOCOM*, San Francisco, CA, 1998.

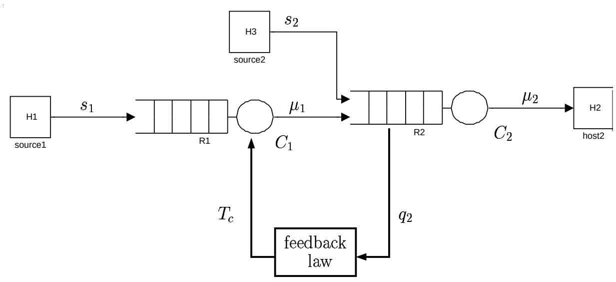

Figure 8: The departure rate of the first router is controlled using information about q_2. This requires information from the second router being send to the first, which corresponds to a centralized control structure.

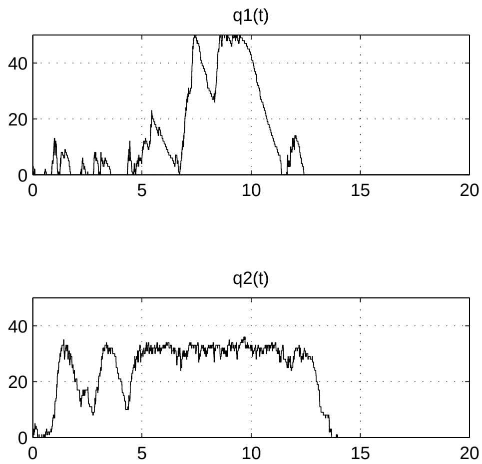

Figure 9: Simulation with control of the first router using information about the queue length q_2 as shown in Figure 8.

[10] S. Floyd. TCP and explicit congestion notification. *ACM Computer Communication Review*, 24(5):10–23, 1994.

[11] F. Paganini, J. Doyle, and S. Low. Scalable laws for stable network congestion control. In *IEEE CDC*, Orlando, FL, 2001.

[12] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of Operational Research Society*, 49:237–252, 1998.

[13] K. J. Åström and B. Wittenmark. *Computer-Controlled Systems*. Prentice Hall, Englewood Cliffs, N.J., 3rd edition, 1997.