Measurement-based admission control for elastic traffic

S. Ben Fredj, S. Oueslati-Boulahia, J. W. Roberts

France Telecom R&D
38, rue du Général Leclerc, 92794 Issy les Moulineaux, France
{slim.benfredj, sara.boulahia, james.roberts}@francetelecom.com

We propose and evaluate an implicit measurement-based scheme for elastic flow admission control in the Internet. We first discuss the nature of IP traffic and present a simple fluid flow model of statistical bandwidth sharing on an isolated bottleneck link. This model is used to guide our choice of admission control algorithms. The main contribution of the paper is to demonstrate by means of detailed packet level simulations that the proposed scheme is efficient and perfectly feasible.

1. INTRODUCTION

It is useful to distinguish two main categories of IP traffic: stream traffic corresponding to audio and video communications and elastic traffic corresponding to the transfer of digital documents. While admission control is generally agreed to be necessary for the former, it seems to be a common understanding that, because the rate of elastic traffic is controlled, there is no need to limit the number of users concurrently sharing some piece of network bandwidth. We argue to the contrary, that admission control is essential for elastic traffic.

Elastic traffic derives from the transfer of a certain mass of documents. We assume that the arrivals of individual flows composing this mass constitute a stationary process. Thus, although transfers can be spread elastically over time by adjusting their rate, it is not possible to reduce the amount of data to be transferred, equal to the flow arrival rate times the average flow size. If this demand is less than capacity, all transfers can be completed and quality of service is generally good. If demand exceeds capacity, on the other hand, congestion necessarily ensues. We suggest admission control is necessary in this case as a form of overload control.

It is inconceivable to perform admission control for individual elastic flows using signalling and resource reservation, since most flows are very small. A measurement-based implementation is essential to ensure the necessary reactivity and to avoid the well known problems of scalability. The rejection of a flow must be realized *implicitly*, simply by discarding the packets which manifest its presence. The loss of the first packets of a flow is generally sufficient signal to the application that the transfer cannot proceed immediately.

Proposals for implicit admission control for elastic traffic have already appeared in the literature. Kumar et al. have even implemented such a scheme on the link from their campus to the Internet which successfully alleviated the effects of congestion [1]. Mortier el al. propose an alternative method which they have evaluated using simulation and implemented on a test bed [2]. The present work is part of our ongoing research on flow aware networking [3–6] by the present authors and their colleagues.

We first discuss the nature of IP traffic and present a simple fluid flow model of statistical bandwidth sharing on an isolated bottleneck link. This model is used to guide our choice of admission control algorithms and to illustrate the potential for service differentiation in overload by means of selective admission control. We propose that admissibility be based on an estimation of the bandwidth which a new flow would acquire and investigate two alternative criteria for deriving this: the measured rate of a permanent "phantom" connection, and the current packet loss rate experienced by traffic on the considered link or path. The main contribution of the paper is to evaluate the proposed schemes by means of detailed packet level simulations. It is demonstrated that, although the erratic traffic fluctuations at packet level make the admission control decisions considerably less precise than predicted by the fluid model, the implementation of implicit measurement-based admission control is perfectly feasible.

2. MODELLING ELASTIC TRAFFIC

We discuss the characteristics of elastic traffic and recall the basic fluid flow model of statistical bandwidth sharing.

2.1. Traffic characteristics

It is well known that the transfer of digital documents under the control of TCP constitutes the majority of Internet traffic. Such traffic is elastic in the sense that TCP connections adapt their transmission rate to the network congestion state. It is more natural and easier to characterize this traffic at the level of the flow or session rather than that of the packet. A flow for present purposes is defined to be the stream of packets corresponding to the transfer of some document (Web page, file, MP3 track,...). A session is a grouping of flows having some common attribute: a Web session, an FTP connection, an e-commerce transaction,... The defining feature of the session is that different sessions are independent.

A simple model of traffic at flow level is to assume flows with a size drawn independently from a certain distribution start at the epochs of a particular arrival process. When the number of sources is large, a natural choice is the Poisson process. These assumptions are not incompatible with observed self-similar packet level characteristics and appear sufficiently realistic to meet our present objectives of illustrating the advantages of admission control and demonstrating the feasibility of a measurement-based implementation. It has moreover been shown that performance results derived for the Poisson model are in fact valid under more general and realistic traffic assumptions [7]. As previously mentioned the distribution of flow size has a heavy tail. This distribution has the characteristic that most flows are very small (so-called "mice") while the majority of traffic is contained in very long flows (so-called "elephants").

2.2. Statistical bandwidth sharing

Elastic flows share bandwidth dynamically under the control of TCP. The degree of fairness achieved by TCP is variable depending on many factors including the connection round trip time, the maximum window size and congestion on other links. For present purposes, however, we make a number of simplifying assumptions about the way bandwidth is shared and limit attention to an isolated bottleneck link. Specifically we assume the bottleneck bandwidth is shared perfectly fairly with instant readjustment whenever new flows begin or existing ones end. Note, however, that detailed packet level simulations of TCP connections are used in Section 5 to validate the concept of measurement-based admission control.

With the assumed traffic model, the shared link behaves like an M/G/1 processor sharing queue [4,3]. Let the flow arrival intensity be λ flows/sec, the mean flow size σ bits, the link capacity C bits/sec and denote by ρ the link load $\lambda\sigma/C$. Assuming $\rho < 1$, the distribution of the number of flows in progress $\pi(n)$ is geometric:

$$Pr[\text{flows} = n] = \rho^n(1-\rho), \tag{1}$$

and the expected duration $R(s)$ of a flow of size s is:

$$R(s) = \frac{s}{C(1-\rho)}. \tag{2}$$

This simple model usefully illustrates the important distinction between throughput performance in underload and in overload. If ρ is not too close to 1, flow throughput, measured by the ratio $s/R(s)$ is satisfactory. In practice, for most shared links of reasonably high capacity, $C(1-\rho)$ is much higher than rate limitations due to causes not considered here (the user's modem, the server, the TCP maximum receive window,...). Such links are virtually transparent with respect to their impact on perceived throughput performance. If, however, $\rho > 1$ our simple processor sharing model would be unstable, the number of flows in progress increasing indefinitely.

Congestion for elastic traffic thus appears as an essentially binary phenomenon: either demand is within capacity and quality of service is excellent or demand exceeds capacity and quality of service is very poor. The objective of introducing admission control is to attenuate the negative impacts of congestion experienced in overload conditions.

2.3. User behavior

In practice when demand exceeds capacity the number of flows in progress does not increase indefinitely. As their bandwidth share diminishes, some flows will be interrupted due to user impatience or aborts triggered by TCP or higher layer protocols. A Markovian model of impatience was proposed in [4]. It showed notably how impatience leads to ineffective link utilization due to bandwidth wasted on flows which do not complete. A more general model introduced in [8] shows that impatience mainly affects the transfer of elephants, the response time of mice generally being sufficiently short even when mean throughput is very low. Of course, aborted flows are likely to be reattempted, further exacerbating the state of congestion. Both bandwidth wastage and discrimination against large flows are absent if admission control is used to reject new flows whenever the mean throughput would otherwise tend to become too low.

3. ADMISSION CONTROL AND THE FLUID MODEL

We generalize the fluid flow model of bandwidth sharing to account for admission control and show that effective service differentiation can be realized by using distinct admissibility thresholds for different traffic classes.

3.1. Choice of admission threshold

With the model assumed in Section 2.2 we additionally apply an upper limit N on the number of admitted flows. In other words, new flows are rejected if the per-flow rate would decrease below the threshold C/N. The blocking probability is given by:

$$B[\rho, N] = \frac{(1-\rho)\rho^N}{1-\rho^{N+1}} \tag{3}$$

This probability is very small when $\rho < 1$ for any moderately large value of N. For instance, a threshold equal to 1% of link capacity (i.e., $N = 100$) gives a blocking probability better than 0.001 for $\rho < 0.96$. On the other hand, when $\rho > 1$, the blocking probability attains the constant fluid limit $(\rho - 1)/\rho$ as soon as N is greater than 50.

The expected per-flow throughput (measured by $s/R(s)$) is virtually independent of N and equal to $C(1-\rho)$ when $\rho < 1$. In overload, on the other hand, throughput drops with increasing N being approximately equal to C/N for moderately large N.

It is important to note that admitting more flows does not reduce the blocking probability in overload and therefore only deteriorates perceived performance. An optimal choice of admissibility threshold should produce negligible blocking in normal load while maintaining sufficiently high throughput for admitted flows in overload. A reasonable compromise for the present system is a throughput threshold of between 0.5% and 2%.

3.2. Selective admission control

An interesting possibility afforded by the implementation of flow admission control is the use of different thresholds to discriminate between traffic classes. Consider again a bottleneck link of capacity C and suppose we have Poisson flow arrivals from two classes with the same size distribution contributing loads ρ_1 and ρ_2, respectively. Flows of class i are blocked when the number of flows in progress of either class is greater than or equal to N_i. We assume $N_1 > N_2$ so that flows of class 1 receive priority service.

We have the following expressions for the blocking probabilities:

$$B_1 = \rho^{N_2}\rho_1^{N_1-N_2}P_0 \tag{4}$$

$$B_2 = \frac{\rho^{N_2}(1-\rho_1^{N_1-N_2+1})}{1-\rho_1}P_0 \tag{5}$$

where

$$P_0 = \left(1 + \frac{\rho \times (1-\rho^{N_2})}{1-\rho} + \frac{\rho_1 \times \rho^{N_2} \times (1-\rho_1^{N_1-N_2})}{1-\rho_1}\right)^{-1}$$

Based on the discussion in the previous section we set $N_1 = 100$ and consider what would be a suitable choice for N_2. Figures 1 and 2 plot the blocking probability of each class as a function of N_2. We observe that class 1 is effectively protected for a wide range of N_2. Expected response time is the same for both classes and approximately equal to $C(1-\rho)$ when $\rho < 1$ and C/N_1 when $\rho > 1$. On the basis of these results a reasonable value for N_2 would be 50.

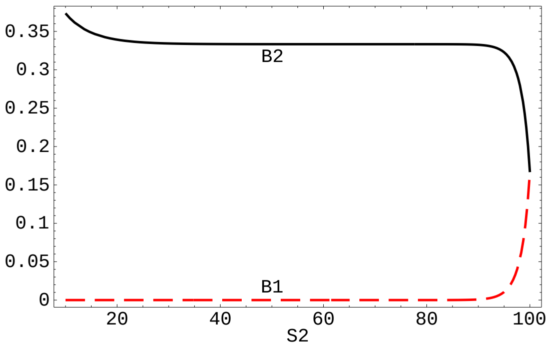

Figure 1. Blocking probability against low priority threshold ($\rho_1 = \rho_1 = 0.6$)

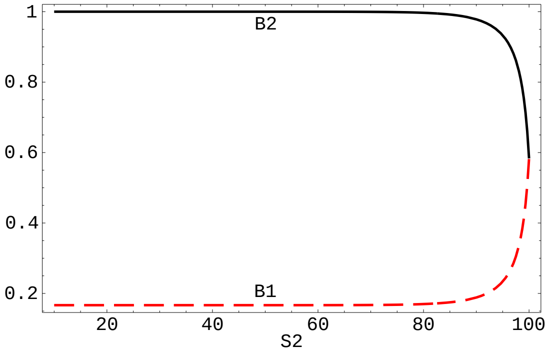

Figure 2. Blocking probability against low priority threshold ($\rho_1 = \rho_1 = 1.2$)

4. MEASUREMENT-BASED ADMISSION CONTROL

In this section we describe the method we propose based on estimating available bandwidth and discuss implementation issues.

4.1. Implicit admission control

Given the very high flow arrival rate on any network link and the small size and duration of most of them, it is inconceivable to implement an explicit admission control procedure based on an exchange of signals between source and network. The admission control procedure must be implicit, with new flows identified on the fly and rejected using existing protocol semantics at transport layer and above.

One possibility used by Kumar et al. [1] and Mortier et al. [2] is based on detecting the SYN and SYN/ACK packets that initialize a TCP connection. If congestion is such that flow rejection is considered necessary, connection set up is aborted by discarding these packets or by sending an RST (reset) packet to the sender. In both cases, the application will recognize that the required transfer cannot take place.

This solution is relatively easy to implement but has a number of disadvantages. It is not possible to detect flows which occur as bursts in a persistent TCP connection, for example. More significantly, admission control can be applied at the TCP connection level only whereas it may be preferable to admit or refuse entire sessions rather than individual flows. Finally, to apply selective admission control it would be necessary to examine additional fields of the SYN or SYN/ACK packet header.

A more general approach is to maintain a list of flows in progress and to systematically compare the flow identity of all arriving packets with this list. The flow identity would be determined from certain fields in the packet (IP and TCP) header (e.g., source and destination addresses and port numbers, the flow identity field). It may also be envisaged to define a specific flow ID field as in IPv6 providing added flexibility in the designation of what constitutes a ‘flow’. It might, for instance, correspond to the succession of Web pages consulted in an e-commerce session, an admission control entity more natural than a single TCP connection for this application.

If a packet belongs to an existing flow it is forwarded; if not, either the new flow is added to the list if it is accepted, or the packet is simply discarded. The discard of the first packets of a flow is generally sufficient signal to the source that resources are unavailable. Flows would be overwritten or erased from the table whenever the time since the last packet exceeds a certain threshold. An important advantage of the latter approach is that the same table of flows could be used in an adaptive flow aware routing scheme [5].

4.2. Admissibility criteria

Several criteria may be used to determine whether the link can accept a new flow. In [1] admission control depends on an estimation of current load. The link enters a blocking state whenever this exceeds a certain threshold (90%, say). It will again accept new flows when the level next decreases below another, lower threshold (80%, say). In [2] the authors use an admission control procedure initially designed for inelastic traffic. This consists in estimating the likely probability of packet loss due to buffer overflow and rejecting a new flow whenever this would exceed an assumed limit value.

The parameters used to calibrate the above algorithms are very loosely related to flow throughput performance. It is possible, for example, that although a link is momentarily saturated, the number of connections is small and any new flow would in fact acquire a satisfactory throughput. We here propose an alternative approach based on estimating "available bandwidth" defined as the bandwidth a new flow would acquire on sharing capacity fairly with the flows already in progress. A significant advantage of this approach is that it applies equally to a single link and to a network path. This is a desirable feature when admission control and routing decisions are performed in the edge routers of an MPLS domain, say.

4.3. Estimating available bandwidth

Estimation of available bandwidth in the Internet is a subject of current research with a variety of objectives. Tools such as Pathchar [9,10] and Netchar [11] have been developed either to estimate the underlying link capacity or the available bandwidth given current traffic levels. However, the objective of these tools is generally to obtain estimates of the long term average available capacity. They are not immediately applicable to the present purpose of estimating in real time the ability of a link or path to accept a new flow. They are also generally more precise than necessary at the cost of considerable implementation complexity.

It is important to understand that the solution we seek does not need to be excessively accurate. A rough estimate of available bandwidth is sufficient to determine whether a new flow can be accepted or not. Recall that admission control is not proposed to ensure flows have a strictly guaranteed minimum throughput. The objective is simply to avoid the negative impact of congestion occurring in a situation of demand overload. We note finally that measurement-based admission control is intrinsically self-correcting in that repeated errors become increasingly less likely.

To perform the bandwidth estimation, our first idea is to implement a "TCP phantom" connection, as proposed by Afek et al [12]. The phantom connection sends a continuous stream of dummy packets and reacts to packet loss precisely as would a regular TCP connection. The bandwidth is measured simply by averaging the short term rate of acknowledged packets.

TCP phantom is easy to implement using existing protocol stacks. However, it might in some cases be considered to generate an unreasonable overhead. A possible enhancement

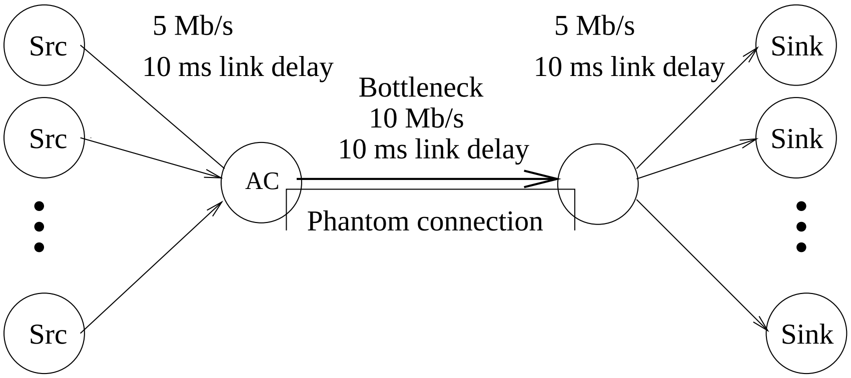

Figure 3. Simulated network topology

requiring a slight modification to TCP would be to use packets without payload but to adjust the congestion window as if the packets had a normal length. The rate of the phantom could also be limited to a maximum value somewhat greater than the admissibility threshold, reducing the amount of artificial traffic in the absence of congestion.

A second approach is to measure the current packet loss rate and use the relation between the latter and the throughput of TCP connections. In practice, it is not straightforward to estimate the loss rate on the link or path in question. One possibility would be to generate a stream of probe packets and measure their loss rate. Alternatively, one might use regular data packets as probes with an added sequence number allowing the detection of loss. The relation between available bandwidth and packet loss can be studied empirically in order to calibrate the admission threshold. We have found this approach preferable to relying on TCP models such as that of Padhye et al. [13].

5. SIMULATION RESULTS

To evaluate admission control algorithms under realistic traffic conditions at packet level we have performed a number of simulation experiments using NS2 [1]. We evaluate both TCP phantom and loss rate based bandwidth estimators.

5.1. Simulated configuration

We considered the simple dumbbell topology shown in Figure 3. All links have the same fixed delay of 10 ms. Admission control is performed on the 10 Mbit/s bottleneck link. The link buffer has a capacity of 50 packets. Four source nodes transfer data to four sinks via 5 Mbit/s feeder links.

TCP connections are generated by each source node according to a Poisson process. Each connection is used to transfer a stream of 1 Kbyte packets representing a document of a certain size and then terminated. The document size is drawn from the following distribution: 90% of documents are "mice" with size uniformly distributed between 1 and 9 packets; the remainder, deemed "elephants" have size uniformly distributed between 10 and 400 packets. This choice is made for the sake of simplicity, performance being largely

[1] http:/www.isi.edu/nsnam/ns/

independent of the size distribution.

Each TCP connection in the simulation is thus identical to a flow. Clearly, in there is no difficulty here in identifying the start of a new flow or in recognizing the flow to which different packets belong.

5.2. TCP phantom estimator

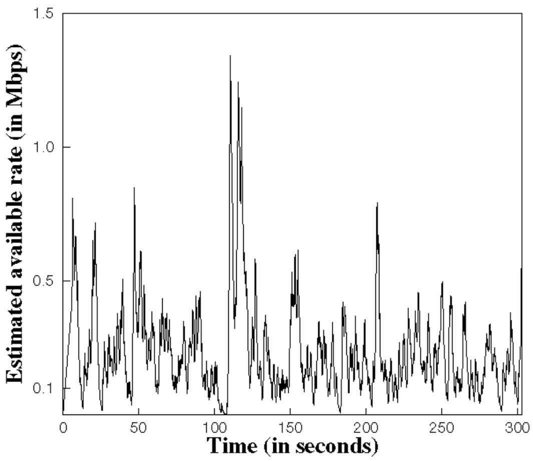

Figure 4. Available bandwidth estimation, load 1.2

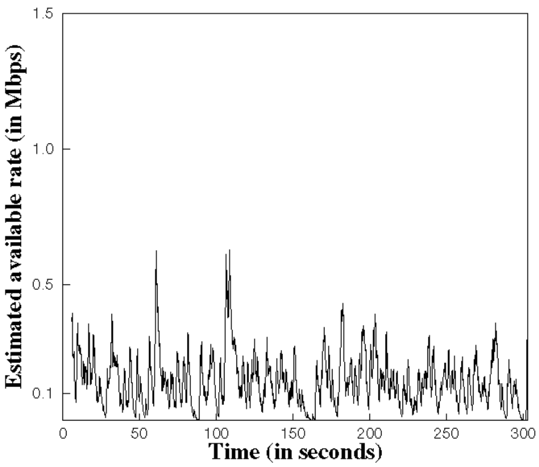

Figure 5. Available bandwidth estimation, load 1.6

The TCP phantom connects the extremities of the bottleneck link. Its goodput, equal to the rate of acknowledged packets, is measured in fixed time intervals of length δ seconds. Let τ_n denote the number of bits acknowledged in the interval $((n-1)\delta, n\delta]$ and let β_n be the available bandwidth estimate derived at time $n\delta$. We apply exponential smoothing with parameter α, $0 < \alpha < 1$:

$$\beta_n = \alpha \times \beta_{n-1} + (1-\alpha) \times \tau_n/\delta. \tag{6}$$

The values of δ and α are not highly critical to the accuracy of the method. Following a series of initial experiments we settled on $\delta = 0.1$ (corresponding to 5 times the RTT of the phantom connection) and $\alpha = 0.9$.

Figures 4 and 5 plot the available bandwidth estimation β_n as a function of time $n\delta$. In this experiment, the admission threshold is set to 100 Kbit/s (1% of the bottleneck link capacity). We observe that the estimated availability varies rapidly about the threshold value. The amplitude of the variations decreases as the offered load increases.

Figure 6 shows how the blocking probability depends on the admission threshold in underload and overload. The figure contrasts the simulation results (dots) with the predictions of the theoretical fluid flow model of Section 3.1 (lines). Figure 7 shows corresponding results for the throughput realized by the phantom connection.

Discrepancies between the analytical and simulation results are due mainly to the fact that TCP does not fully exploit the link capacity (the buffer sometimes empties) and the

Table 1
Impact of the rate threshold on the realized throughput (in %) of connections

Threshold	$\rho = 0.9$			$\rho = 1.4$		
	All	> 100	Phantom	All	> 100	Phantom
0.5%	1.8	4.9	9.5	1.2	2.1	1.7
2%	1.9	5.1	12.6	1.4	3.0	4.2
4%	1.9	5.4	13.5	1.5	3.3	5.4

extra load induced by the phantom connection which is not accounted for in the fluid model. Nevertheless, the simulation results confirm the observations made in Section 3.1 on the choice of admission control threshold.

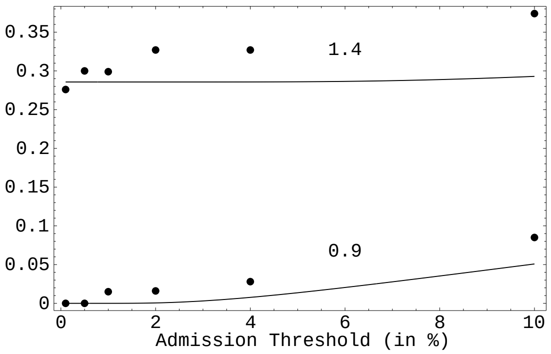

Figure 6. Blocking probability, ρ =0.9, 1.4

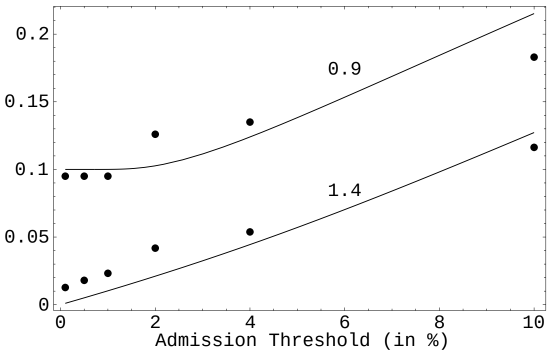

Figure 7. Normalized throughput of the phantom connection (dots), ρ =0.9, 1.4

In underload ($\rho = 0.9$) we observe that blocking is negligible for any threshold smaller than 0.5% and while the phantom throughput effectively attains the residual bandwidth $(1 - \rho)$. In overload ($\rho = 1.4$) blocking is given approximately by the fluid limit $(\rho - 1)/\rho$ but increases with the threshold due to the inefficiency of TCP which cannot saturate the link when the number of connections is small.

Table 1 compares the throughput realized by the admitted connections depending on their size. Results for "all connections" are dominated by the throughput of mice which is severely limited by TCP slow start even when the number of simultaneously admitted flows is small. The throughput of large transfers of more than 100 packets depends more on TCP congestion avoidance and tends to that of the TCP phantom.

In conclusion, we recommend a rate threshold around 0.5% to ensure transparency in normal load. Simulations confirm that any value between 0.5% and 2% is acceptable in overload. A threshold higher than 4% tends to be inefficient since the admitted connections are not able to completely saturate the link (high blocking and no compensating increase in throughput). Thresholds lower than 0.5% increase the response times and do not reduce the blocking probability.

Table 2
Impact of the loss threshold on the realized throughput (in %) of connections and on the blocking (in %)

Loss threshold	$\rho = 0.9$			$\rho = 1.4$		
	All	>100	Blocking	All	>100	Blocking
1%	3.1	13.3	11.2	2.7	10.7	37.1
5%	2.8	11.1	0	1.6	3.9	27.3
10%	2.8	11.2	0	1.1	1.8	26.7

5.3. Loss rate estimator

In this section we evaluate the effectiveness of the second admission control approach based on the measured loss rate. In the simulations, we simply measure the loss rate averaged over all packets using the bottleneck link. The loss rate is measured on 0.1 second intervals and averaged using an exponential smoothing parameter $\alpha = 0.9$.

It proves difficult to precisely calibrate the observed loss rate with a target available bandwidth. Table 2 gives the average throughput realized by all flows and by large flows (> 100 packets), respectively, together with the blocking rates for different loss thresholds under two load conditions.

We observe that a threshold smaller than 1% is overly conservative and leads to significant blocking in normal load. For thresholds greater than 5% blocking is relatively stable and roughly equal to the fluid limit. Realized throughput decreases as the threshold increases, notably for large transfers. An admission threshold of 5% appears as a reasonable choice for the present configuration.

5.4. Selective admission control

In this section we examine the precision of service discrimination realized by applying different admission thresholds to distinct traffic classes.

TCP phantom estimator

The analytical results in Section 3.2 predict that performance is largely insensitive to the value of the threshold of class 2 flows as long as it is not too close to that of class 1. Here we show results for $\rho_1 = \rho_2 = 0.45$ and $\rho_1 = \rho_2 = 0.7$. The admission threshold of class 1 is set to 0.5% and we consider two different values for the admission threshold of class 2: 2% and 4%.

Table 3
Blocking (in %) using TCP phantom, ρ=0.9 and 1.4, threshold of class 1 =0.5%

Class 2 rate threshold	$\rho = 0.9$			$\rho = 1.4$		
	Class 1	Class 2	All	Class 1	Class 2	All
2%	0	2.7	1.3	12.0	49.7	30.7
4%	0	6.7	3.3	4.5	61.4	32.7

Table 3 shows the results obtained. Discrimination is effective but is not as clearcut as predicted by the fluid model. On the basis of these results, the choice of a class 2 threshold of 4% appears as the best choice.

Loss rate estimator

We set the threshold on the admissible loss rate for the privileged class to 5% and consider the thresholds of 1% and 2% for class 2. Table 4 gives the blocking rates observed for both classes.

Table 4
Blocking (in %) using measured loss rate, ρ=0.9 and 1.4, loss threshold of class 1=5%

Class 2 loss threshold	$\rho = 0.9$			$\rho = 1.4$		
	Class 1	Class 2	All	Class 1	Class 2	All
1%	0	14.4	7.2	0	69.8	34.6
2%	0	7.8	3.9	0.3	63.1	31.5

We observe that discrimination is effective: class 1 hardly experiences any blocking. However, the stricter threshold of 1% leads to high blocking for class 2. A looser loss threshold of 2% is effective in terms of discrimination and results in better accessibility for the class 2 flows.

6. CONCLUSIONS

Admission control for elastic flows in the Internet appears as a necessary safeguard against the potentially harmful effects of overload. In order to ensure sufficient reactivity and to avoid problems of scalability, such admission control must be implicit and measurement-based.

A simple fluid model of statistical bandwidth sharing suggests that the choice of admissibility criterion is not highly critical and that performance is essentially robust to imprecision in the implemented procedure. We choose to base admissibility on the amount of bandwidth which a new flow would acquire when sharing link capacity fairly with the flows currently in progress. Admission control is then inoperative in normal traffic conditions and only leads to flow rejection in overload. Use of distinct class-based admission thresholds constitutes an effective service differentiation device preserving higher priority classes from the effects of overload while handling as much lower priority traffic as possible.

To test the feasibility of implementation we have performed extensive and detailed packet level simulations. To estimate available bandwidth we have tested two approaches: 1) simulate an artificial "phantom" TCP connection and measure the bandwidth it currently receives, and 2) use the fact that TCP adjusts bandwidth in reaction to packet loss and calibrate admission thresholds with respect to the observed loss rate. Both approaches work satisfactorily, though the particularities of TCP not taken into account in the fluid model (notably the influence of slow start) lead to some additional noise in the decision criteria. The loss-based estimation avoids the overhead constituted by the phantom connection but might be more difficult to implement in practice.

The results presented here are preliminary. Additional analyses and simulations for a bottleneck link integrating stream and elastic traffic are presented in [14]. We are continuing our investigations by simulation of more extensive network topologies. We are also setting up an experimental test bed and plan to perform trials on real traffic. A particular concern is to demonstrate the scalability of the proposed implicit measurement-based admission control framework.

REFERENCES

1. A. Kumar, M. Hegde, and S.V.R. Anand. NETMASTER : Experiences in Using Nonintrusive TCP Connection Admission Control for Bandwidth Management of an Internet Access Link. *IEEE Communications Magazine*, May 2000.
2. R. Mortier, I. Pratt, C. Clark, and S. Crosby. Implicit Admission Control. *IEEE Journal on Selected Areas in Communications*, December 2000.
3. L. Massoulié and J.W. Roberts . Bandwidth Sharing and Admission Control for Elastic Traffic. *Telecommunications Systems*, 15 (2000) 185-201 (also in *ITC Specialist Seminar, Yokohama*, 1998).
4. L. Massoulié and J.W. Roberts. Arguments in Favour of Admission Control for TCP Flows. In P. Key and D. Smith (editors), *Teletraffic Engineering in a Competitive World, Proceedings of ITC 16*, pages 33–44. Elsevier, June 1999.
5. J.W. Roberts and S. Oueslati-Boulahia. Quality of Service by Flow Aware Networking. *Phil. Trans. Royal Society London*, 2000. Available at : http://www.enst.fr/˜oueslati.
6. S. Oueslati-Boulahia and E. Oubagha. A Comparative Study of Routing Algorithms for Elastic Flows in a Multiservice Network. In *ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management*, Monterey, 2000.
7. T. Bonald, A. Proutière, G. Régnié, and J.W. Roberts. Insensitivity Results in Statistical Bandwidth Sharing. In *Proceedings of ITC 17*, Salvador, Brazil, 2001. (in these proceedings).
8. S. Ben Fredj, T. Bonald, A. Proutière, G. Régnié and J.W. Roberts. Statistical Bandwidth Sharing: A Study of Congestion at Flow Level. in *Proceedings of ACM SIGCOMM'01*, August 2001.
9. Pathchar: A Tool to Infer Characteristics of Internet Paths. Presented at the Mathematical Sciences Research Institute (MSRI), April 1997. Slides available from ftp://ftp.ee.lbl.gov/pathchar.
10. K. Lai and M. Baker. Measuring Bandwidth. In *Proceedings of IEEE INFOCOM'99*, 1999.
11. A.B. Downey. Using Pathchar to Estimate Internet Link Characteistics. In *Proceedings of ACM SIGCOMM'99*, 1999.
12. Y. Afek, Y. Mansour, and Z. Ostfeld. Phantom: A Simple and Effective Flow Control Scheme. In *Proceedings of ACM SIGCOMM'96*, 1996.
13. J. Padhye, V. Firoiu, D. Towsley, and J. Kurose. Modeling TCP Throughput: A Simple Model and its Empirical Validation. In *Proceedings of SIGCOMM'98*, 1998.
14. N. Benameur, S. Ben Fredj, S. Oueslati-Boulahia and J.W. Roberts. Integrated admission control for streaming and elastic traffic. In *Proceedings of QofIS 2001*, Cuimbra, Portugal, September 2001.