Measurement-Based Admission Control for Flow-Aware Implicit Service Differentiation

Jordan Augé†
UPMC Sorbonne Universités and CNRS,
Paris, France.
<jordan.auge@lip6.fr>

Sara Oueslati
Orange Labs
Issy-les-Moulineaux, France
<sara.oueslati@orange-ftgroup.com>

James Roberts
INRIA,
Paris-Rocquencourt, France.
<james.roberts@inria.fr>

***Abstract*—It has previously been shown that the combined use of fair queuing and admission control would allow the Internet to provide satisfactory quality of service for both streaming and elastic flows without explicitly identifying traffic classes. In this paper we discuss the design of the required measurement based admission control (MBAC) scheme. The context is different to that of previous work on MBAC in that there is no prior knowledge of flow characteristics and there is a twofold objective: to maintain adequate throughput for elastic flows and to ensure low packet latency for any flow whose peak rate is less than a given threshold. In the paper we consider the second objective assuming realistically that most elastic and streaming flows are rate limited. We propose a MBAC algorithm and evaluate its performance by simulation under different stationary traffic mixes and in a flash crowd scenario. The algorithm is shown to offer a satisfactory compromise between flow performance and link utilization.**

I. Introduction

In this paper we return to the much studied question of admission control and, in particular, to the design of an effective measurement-based admission control (MBAC) algorithm. Despite extensive research over many years it is fair to say no entirely satisfactory algorithm has been proposed for the admission of variable rate flows, even in the favourable case of "bufferless multiplexing". The subject is clearly out of fashion among networking researchers but admission control is still an essential ingredient of effective QoS control and many significant issues remain open.

We consider MBAC in a particular context. This is the proposal to realize service differentiation using two flow-based router mechanisms, known as Cross-protect, that avoid the need for flow rate estimates or explicit class of service marking [10]. These mechanisms are a per-flow fair queuing scheduler and an MBAC that uses load characterizations provided by the scheduler. Fair queuing ensures max-min fair sharing of link bandwidth. Flows of rate less than the fair rate are not backlogged and their packets experience low latency. In Cross-protect, delay for such flows is minimized by sending their packets to a priority queue. Implicit differentiation is realized on recognizing that streaming flows that require low latency indeed have relatively low peak rates, and generate no queueing. Admission control is necessary to maintain the fair rate sufficiently high and to ensure the load on the priority queue remains within bounds.

Admission control in Cross-protect can only be based on the limited traffic measurements made possible by the scheduler. These are basically byte counts of packets emitted in the priority queue in successive time intervals and a running estimate of the current fair rate. We have no *a priori* knowledge of flow characteristics and can only detect flow completions by the absence of new packets in a timeout interval.

When a significant proportion of traffic is elastic and flows can attain the fair rate, it is relatively easy to protect the performance of flows in progress simply by rejecting new flows when the fair rate goes below a certain threshold (set typically to 1% of the link rate and typically only exceeded in situations of severe overload [1]). In this paper we consider a more problematic but typical scenario where the vast majority of flows are limited in rate (e.g., by user DSL access rates) and consequently, in normal load conditions, are handled with priority by the Cross-protect scheduler.

While the available information on traffic characteristics is quite limited, the considered context is highly favourable to efficient statistical multiplexing. By assumption, the link rate is much higher than the maximum peak rate of flows intended for priority service. The latter is necessarily less than the 1% (or so) fair rate threshold and typically much smaller (e.g., 4 Mbit/s video flows sharing a 2.5 Gbit/s OC48 backbone link). High efficiency is thus possible while maintaining excellent per-flow performance.

In the following sections we first discuss some existing MBAC algorithms that are relevant to our problem. Our approach is then presented in Section III and evaluated in Section IV under a number of demand scenarios. Conclusions are presented in Section V.

II. Related work on MBAC

Measurement-based admission control has, of course, been widely researched over many years. However, only a small number of proposed algorithms make the kind of

†The author is affiliated with LIP6 Computer Science Laboratory and LINCS (Laboratory of Information, Network and Communication Sciences)

978-0-9836283-0-9

This paper was peer reviewed by subject matter experts for publication in the Proceedings of ITC 2011

minimal assumptions about traffic characteristics that are appropriate for the present context. We only review these contributions in the present section.

Jamin et al [7] propose a simple MBAC algorithm called *measured sum*. A new flow is admitted if the sum of its nominal rate and the estimated rate of aggregate flows is less than a utilization target times the link bandwidth. The estimated rate of aggregate flows is derived using a time window estimator. This algorithm is clearly very simple and takes no explicit account of a performance objective or the nature of traffic. It therefore has limited predictability.

Gibbens et al. [4] propose a decision theoretic approach, where a new flow is admitted only if the current aggregate load is less than a certain threshold. In the simplest variant proposed, the flow peak rate is the only flow information taken into account. This method thus corresponds with present requirements and presented results provide useful confirmation that efficient multiplexing is possible even in this restricted framework. Unfortunately, to compute the admission threshold requires prior knowledge about the level of offered load and flow burstiness that is not available in our context.

Grossglauser and Tse [5], [6] propose an algorithm based on measured overall traffic which could be adapted to fit Cross-Protect requirements. This is based on a Gaussian approximation of aggregate demand considered in their model as a fluid arrival rate. The admission condition is designed to satisfy a target probability for this overall demand exceeding link capacity. New flows are admitted while:

$$C - A_t - r > \alpha_q \widehat{\sigma}_t \tag{1}$$

where C is the link capacity, A_t is the measured aggregate load (in bits/s) at time t, r is the flow peak rate, $\alpha_q = Q^{-1}(\epsilon)$ with ϵ the target loss probability and $Q(.)$ the complementary distribution function of a $N(0,1)$ Gaussian random variable, and $\widehat{\sigma}_t$ is the estimated standard deviation of the aggregate load. When a flow is admitted its peak rate is added to A_t. This prevents momentary overloads under heavy traffic due to too many flow admissions following a low estimate A_t.

To estimate A_t and $\widehat{\sigma}_t$, Grossglauser and Tse introduce the notion of critical time scale ($\tilde{T}_h$), broadly equal to the time scale over which the impact of an admission decision persists. $A(t)$ should be estimated by exponential averaging over the critical time scale while the variance should be measured over a much longer time scale. They establish that $\tilde{T}_h = \frac{T_h}{\sqrt{n}}$ where T_h is the average flow holding time and n is the number of flows the link can carry.

Despite significant differences between proposed MBACs, it turns out that they all result in the same trade-off between utilization and perceived per-flow performance [3]. The algorithms differ in terms of their predictive capability: the choice of parameter values for the admission condition resulting in a given performance target. Breslau et al. [3] show that all the algorithms they tested are quite poor at predicting performance as measured in terms of the packet loss ratio. Reasons why it is intrinsically difficult to control performance parameters like the packet loss rate by means of MBAC are discussed by Bean [11]. Simply stated, the problem is that inappropriate admission decisions leading to excess traffic produce a loss rate that may easily be ten times greater than the target for a certain time. The average loss rate can only be restored by ensuring a much lower loss rate for a period more than ten times longer. It appears that we can, in fact, only hope to derive algorithms that are reasonably efficient and this is the limit of our ambition.

III. An MBAC for Cross-protect

We discuss the design of an MBAC that can ensure streaming flows of peak rate less than a given threshold p experience negligible packet loss.

A. Admission criteria

Flow admissibility in Cross-protect is determined from the values of two measures of congestion, the fair rate FR and the priority load PL. These values are determined as running averages in successive time slots of length τ and admissibility in slot t is determined from values calculated in slot $t-1$.

FR is the rate a flow would acquire on output if it always had packets to send and can be estimated from data available to the scheduling mechanism [8], [9]. It is estimated by a long term average representing the rate acquired over an interval commensurate with the time scale of flow arrivals and departures. PL is an exponential moving average of the traffic in bits arriving to the priority queue in one slot. Packets sent to the priority queue are those arriving to flows that are not currently backlogged.

When a significant proportion of traffic is from elastic flows that are not limited in rate elsewhere on their path, the most significant admission criterion is FR. Priority load remains low and is such that packet loss and delay for streaming flows of peak rate somewhat less than FR are negligibly small. Admission control is rather easy in this case since elastic flows are naturally tolerant of any imprecision in the fair rate estimate [8], [9].

In practice, however, the large majority of elastic flows have a limited peak rate and many will be handled together with streaming flows in the Cross-protect priority queue. FR is then not critical while the measured priority load PL includes traffic from flows of peak rate greater than p that we do not seek to protect. The difficulty in designing the Cross-protect MBAC resides in differentiating the impact of flows whose peak rate is greater than or less than the threshold p without explicitly distinguishing them.

We assume flows of peak rate less than or equal to p will suffer negligible packet loss and delay if the traffic in bits

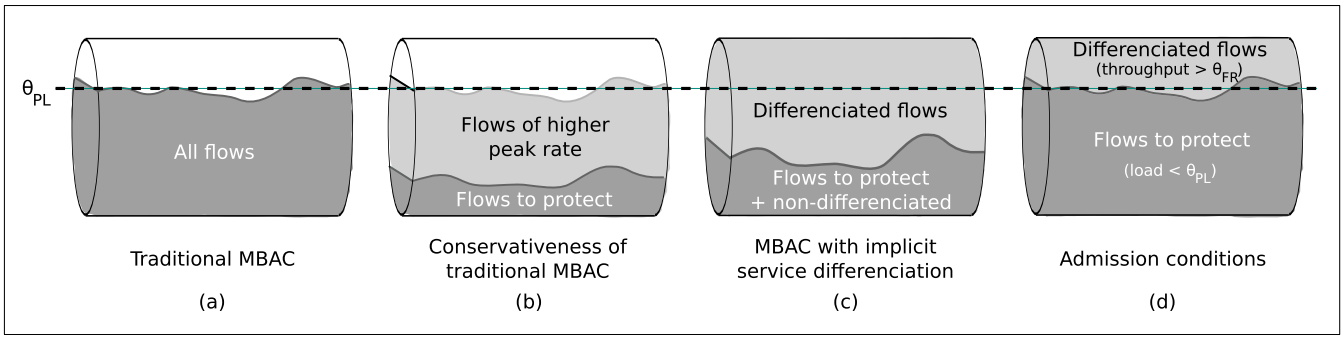


Fig. 1. Sketch of possible utilization regimes compared to target PL threshold θ_{PL}

forwarded to the Cross-protect priority queue in a slot of length τ remains less than capacity $C\tau$ with probability greater than $1-\epsilon$. We suppose that ϵ can be calibrated to ensure sufficiently low packet loss and delay.

B. Mean and variance of priority load

Let b_t be the measured load in bits/s sent to the priority queue in slot t. The mean priority load estimate A_t is the moving average:

$$A_t = (1-\alpha) \times A_{t-1} + \alpha \times b_t. \tag{2}$$

where parameter $\alpha = 1 - \tau/\tilde{T}_h$ and $\tilde{T}_h$ is the critical time scale defined in [6].

To apply the approach from [6], we estimate the variance as follows:

$$\begin{aligned} \widehat{\sigma}_t^2 &= D_t - E_t^2 \text{ where} \\ D_t &= (1-\beta)D_{t-1} + \beta(b_t - A_t)^2, \\ E_t &= (1-\beta)E_{t-1} + \beta(b_t - A_t). \end{aligned} \tag{3}$$

and set the smoothing parameter β to $(1-\alpha)/10$. Estimates of the flow holding time would need to be established periodically by measurement. It is also worth noting that the distribution of flow durations is not exponential as assumed in [6] but heavy-tailed. Fortunately, it appears as in previous work that the choice of smoothing parameters α and β is not highly critical.

To apply condition (1) using these estimates proves too conservative for our purpose when some flows have a peak rate greater than p. We assume such flows are able to adjust their rate to the current fair rate in case of aggregate rate overload and do not require protection. For example, in Figure 1b), measured fair rate and priority load are the same as in Figure 1a), but we could accept more flows without violating the latency requirements for flows of rate less than p. Applying measured mean and variance in (1) would seek to maintain load below the threshold θ_{PL} and prevent the system from entering the favourable state represented in Figure 1c). In this state, flows of rate greater than p become backlogged and no longer contribute to priority load. The sketch in Figure 1d) shows an ideal situation where the admission control perfectly discriminates flows of peak rate greater than and less than p, and the traffic mix allows the link to be saturated.

C. A Poisson approximation

We assume the number of inelastic flows of peak rate less than or equal to p in progress in a given measurement slot has a Poisson distribution. This would be the case in the absence of blocking for a quite general traffic model where flows occur in sessions of alternating flows and think times and the session arrival process is Poisson. If packets are of constant maximal length L, the number of packets arriving in any slot not greater than L/p also has a Poisson distribution[1]. This suggests we can estimate the variance of the rate in a slot from the measured mean yielding the following simple MBAC.

Choose a slot length $\tau = L/p$. Given the measured priority load A_t bits/s, let $m_t = A_t\tau/L$ be an estimate of the number of packets and deduce the variance estimate $\widehat{\sigma}_t^2 = m_t L^2/\tau^2 = A_t p$. Apply this estimate in admission condition (1).

By ignoring flow blocking and the fact that some flows emit packets of size less than L, the Poisson assumption leads to conservative admission decisions. However, given the assumption that p is a small fraction of link capacity C, the approximation nevertheless allows reasonably high utilization. Table I gives the link load threshold corresponding to particular values of C/p and ϵ. Admission control would be applied in slot t when A_{t-1} is greater than this threshold.

C/p	100	100	1000	1000
ϵ	0.001	0.01	0.001	0.01
α_q	3.09	2.33	3.09	2.33
Threshold	0.73	0.79	0.91	0.93

TABLE I
ADMISSION THRESHOLDS

Note that if flows of peak rate greater than p are included in the priority load estimate A_t, the variance estimated by (3) would be greater than the Poisson estimate $A_t p$. The MBAC derived from [6] would be too

[1] A flow of rate r independently sends a packet in this interval with probability r/p.

conservative, preserving states like Fig. 1b). We expect the Poisson estimate to more readily allow transition to states like Fig. 1c).

On the other hand, if the aggregate traffic has many flows of rate much smaller than p, the Poisson approximation may be too conservative. We therefore propose a more refined MBAC where the variance estimate is the minimum of $A_t p$ and that calculated by (3).

The Cross-protect MBAC algorithms are called Poisson and MinVar (for minimum variance) algorithms. Their performance is evaluated for some test scenarios in Section IV below.

D. Limiting the number of arrivals per slot

Admission control is especially useful in exceptional events when demand considerably exceeds capacity. This occurs in particular when a failure somewhere in the network leads to traffic being rerouted over a considered link. Traffic will increase to attain in time a new stationary load. The MBAC should be able to reject excess traffic in this regime to maintain performance objectives. However, the main impact of the failure is for flows in progress on the interrupted path to appear suddenly as a burst of apparently new flows appearing on links of the fail-over path. A link on this path will generally be uncongested before the failure and therefore open to accept new flows. If, however, all "new" flows that arrive in the time slots following the failure are accepted, the link will immediately become heavily congested.

To alleviate this effect we limit the number of new flows accepted in any slot, as in [6]. Given current estimates A_t and $\widehat{\sigma}_t^2$, we accept a maximum of n new flows where $(n+1)p + A_t + \alpha_q \widehat{\sigma}_t > C$. This may still not be sufficient given the time lag before the traffic of the new flows appears in the load estimates A_t. Note further that, as flow durations generally have a heavy-tail distribution, the flows that fail-over to the new link will have an expected *residual* duration that is greater than the mean. The impact of wrong positive admission decisions is therefore more severe than for genuinely new flows.

The back-off strategy proposed in [4] where, once one flow is blocked, no other flow is accepted until an ongoing flow terminates is not applicable in our system since flow terminations are not signalled. The envisaged solution is to interrupt the protection of a sufficient number of on-going flows to alleviate congestion noting that interrupting flows is in all cases necessary when the combined traffic on failed and back-up paths exceeds remaining capacity.

E. Instability of measured priority load

The measured priority load can vary suddenly as flows with a common peak rate P become backlogged and are therefore no longer handled with priority. If P is close to p, this can momentarily lead to an anomalous situation where A_t is below the admission threshold and the (long term) average fair rate stays above its threshold even though flows of rate p are backlogged.

To limit the impact of this phenomenon, we set the priority load to C in any slot where a flow of rate p would be backlogged. The latter condition can be easily deduced from the parameters of the priority fair queueing algorithms defined in [8], [9]. It is an instantaneous measure of the current fair rate while FR is a long term average.

This modification also proves useful when flows have a nominal peak rate p but in practice are subject to jitter, becoming momentarily backlogged when the inter-packet interval is too small (see Section IV-F).

F. Impact of the length of the sampling interval

The choice of the discretization interval will generally be guided either by technical constraints, or by a target peak rate for the flow we want to estimate. We envisage intervals of length kL/p for $k \geq 1$.

Lower values of k fail to take proper account of the low variance of flows of rate less than p. On the other hand, larger values of k tend to make the MBAC unresponsive to sudden changes, as in envisaged flash crowd scenarios.

IV. Evaluation of the algorithms

We evaluate our Poisson and MinVar propositions using extensive ns-2 simulations, and compare the results with an algorithm inspired from [6] and equation 3, that we denote GT[2].

A. Simulation set-up

The topology for the simulation is the traditional dumbbell, illustrated by the bold part in figure 2, with a central link capacity $C = 10$Mb/s. The deliberate choice of such a low value for C is intended to keep simulation time short, since performance results only depend on the C/p ratio (confirmed by simulations).

Traffic is composed of UDP flows of exponentially distributed duration with mean $T_h = 60$s. They arrive as a Poisson process. The flows are generally assumed to generate an on/off packet arrival pattern with given peak rate in the on-periods (from 50 to 300kb/s). Packet size is constant and set to 1000 bytes. Simulations are run for 2000s multiple times (25), and we focus on the stationary regime by discarding the first 200s. The target overflow ϵ is set to .01. We simulate a stationary load equal to 100, 120 and 140% of link capacity. Unless specified, the arrival rate of the flows is such that the limit on the amount of traffic admitted in a slot is inoperative and thus has no impact. The sampling interval is $\tau = kL/p$ with $k = 1$ or a small integer, as specified.

[2]Source of ns2 modules and scripts used for simulation are available online at http://jordan.auge.free.fr/research/mbac

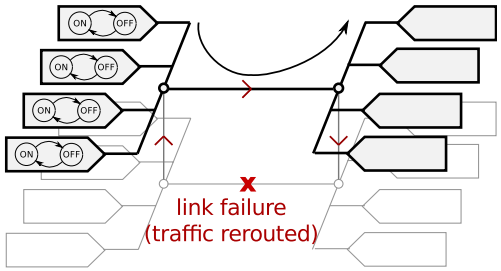


Fig. 2. Simulation set-up

B. Performance criteria

Monitoring the performance of both streaming and elastic flows requires selecting metrics that reflect differentiation realized by our scheduler. The overflow probability (ov) represents the proportion of slots where the instantaneous priority load is greater than link capacity. It is representative of streaming flow performance, provided their packets are correctly sent with priority, which is measured by the *backlog probability* (bk considered for each class).

We also measure the *blocking probability* (bl), which represents the proportion of flows that are not admitted, and the *packet loss rate* (lo), the proportion of packets that are lost due to buffer overflow. The combination of both metrics shows how the excess traffic is discarded during moments of congestion. An effective MBAC should make bl as low as possible while guaranteeing the quality of service of both classes of flows. We also track the proportion of times each admission criterion is responsible for flow blocking (bl_{PL} and bl_{FR}). Finally, we measure *link utilization* (ut) which, while it depends on the traffic mix, gives insight into how well the resources of the link are exploited.

C. Utilization versus overflow probability

The relation between utilization and overflow probability represents the optimal performance that can be attained by the algorithm in stationary conditions. It is derived for the considered MBAC algorithms by varying the admission criterion over a range. For GT, we vary ϵ and note the realized utilization and overflow probability. For either Cross-protect algorithm (Poisson and MinVar), denoted XP in the figure, we simply vary an admission threshold on the observed load A_t, without relying on either mean or variance. Figure 3 presents results where flows all have the same peak rate of 100Kb/s. The figure also plots the same relation for a load that has an exact Poisson distribution.

Unsurprisingly given results reported in [3], the relation is broadly the same for the two MBACs. A subtle difference occurs at high loads. As the link becomes saturated, flows in Cross-protect are momentarily backlogged and no longer contribute to the priority load allowing more admissions and higher utilization. Such high threshold values correspond to situations where flows of rate p become backlogged, of course, and performance in this case is not adequately measured by ov.

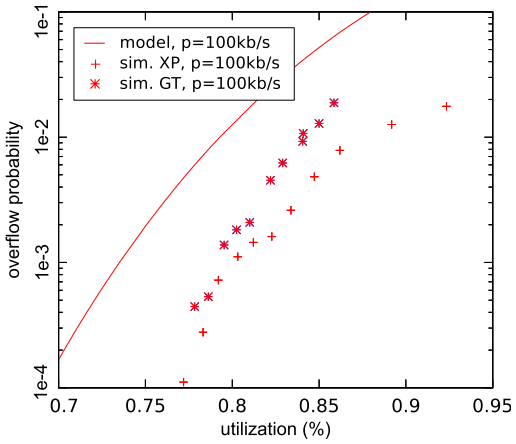


Fig. 3. Utilization against overflow probability for XP and GT MBACs compared to a Poisson distributed slot load

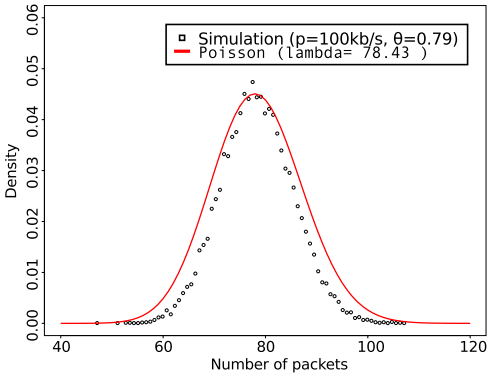


Fig. 4. Load density for $\rho = 1.20$, $\theta = 0.78$ and $p = 100$kb/s, with fitted Poisson distribution

From the plot, we can deduce that the load per slot is less variable than Poisson since achievable utilization is higher for the same overflow objective. Figure 4 shows the empirical histogram of the number of packets arriving in one slot for flows of peak rate 100Kb/s under the Poisson MBAC. A fitted Poisson distribution is shown to correspond quite closely suggesting our approximation is reasonable (the real traffic being less variant). Results not presented here show the same behaviour for different peak rates values.

D. Predictiveness of XP MBAC

While most MBAC algorithms have the same performance frontier, they differ in their predictiveness, i.e., the predictability of their performance for given parameter settings. We evaluate the predictiveness of Poisson and MinVar for homogeneous peak rate flows. The protected rate p may be equal to, greater than or less than the actual flow peak rate r. Results with 95% confidence interval are presented in Table II for a sampling slot size $\tau = L/p$.

	p	r	ov	ut	bl	bk	lo
Poisson	50	50	3.98e-4 ±0.92e-4	84.08 ±0.02	31.37 ±0.38	2.82e-3 ±0.52e-3	0.00 ±0.00
	100	100	3.11e-4 ±0.64e-4	78.44 ±0.06	35.20 ±0.61	8.14e-3 ±1.80e-3	0.00 ±0.00
	100	50	3.05e-3 ±0.20e-3	78.79 ±0.02	35.80 ±0.29	1.85e-5 ±2.19e-5	0.00 ±0.00
	100	300	2.71e-4 ±0.86e-4	97.63 ±0.96	3.04 ±1.49	76.27 ±3.32	15.60 ±1.14
MinVar	50	50	5.42e-3 ±0.31e-3	88.28 ±0.05	27.99 ±0.33	6.30e-2 ±0.44e-2	0.00 ±0.00
	100	100	3.24e-3 ±0.15e-3	83.59 ±0.08	30.98 ±0.59	1.32e-1 ±0.10e-1	0.00 ±0.00
	100	50	7.69e-3 ±0.30e-3	81.25 ±0.06	33.68 ±0.40	6.54e-4 ±5.62e-4	0.00 ±0.00
	100	300	2.13e-4 ±0.56e-4	98.33 ±0.69	1.84 ±0.98	78.91 ±2.56	16.21 ±1.08

TABLE II
PERFORMANCE OF CROSS-PROTECT MBAC FOR TRAFFIC WITH FLOWS OF SAME PEAK RATE

1) Flows of know peak rate: The results when $r = p$ are presented in Table II, for $p = 50$Kb/s and $p = 100$kb/s. The overflow probability remains one or two orders of magnitude less than our objective, leading to a link utilization lower than what could have been possible (deduced from Fig. 3). However, the loss in utilization is slight and quality of service constraints are respected. Over-conservatism is a common pitfall for all MBAC algorithms. Results for MinVar MBAC show that using measured variance improves performance, as expected since the traffic aggregate is in fact less variant than Poisson due to blocking. Performance is even better in this respect with $k = 2$.

2) Flows with lower peak rates: Typically, the value of p will be set sufficiently high to protect all streaming flows including many with a much lower peak rate. Lines in Table II for $p = 100$ and $r = 50$ show the impact of assuming a peak rate higher than that actually present. With the Poisson MBAC, the threshold only depends on p so that achieved performance with $r = 50$kb/s is roughly similar to $r = 100$kb/s. This illustrates the cost of using the simple Poisson MBAC which nevertheless remains slight with this link size (10Mb/s). MinVar is more effective in this case when the sampling interval is equal to kL/p and $k \geq 2$ (results not shown here) and, in fact, has the same performance as the GT MBAC in this case (utilization rises from 81.25 ± 0.06 % to 88.02 ± 0.06 %).

3) Flows with higher peak rates: The last case where $r > p$ illustrates the major advantage of the (XP) MBAC. When admission criteria are set for $p = 100$, flows with peak rate $r = 300$ are backlogged and hardly contribute to priority load. The link is fully utilized and flows lose a high proportion of packets. Blocking ensues when their long term rate goes below the fair rate threshold $C/100$ (which corresponds to high load).

E. Performance with a mixture of flow peak rates

To investigate the differentiation realized in a heterogeneous context, we consider two classes of flows, with peak rate under and above p, and overall load $1 < \rho_1+\rho_2 < 1.40$.

Figure 5 represents realized differentiation in a set of experiments where $r_1 = 50$kb/s and $r_2 = 300$kb/s, respectively with Poisson (a) and MinVar with parameter $k = 1$ (b) and $k = 2$ (c). The x and y axes represent ρ_1 and ρ_2, with constant overall load in the same shade of gray. The symbol indicates the level of differentiation: ✓ (full), ∼ (good, differentiation occurs after some time), ≈ (episodic differentiation), × (no differentiation).

With Poisson, differentiation occurs for $\rho = 1.20$ and $\rho = 1.40$ as soon as the load of the low peak rate class is not too high. MinVar improves differentiation while maintaining QoS, especially when k is properly sized (here $k = 2$). MinVar with $k = 2$ manages to differentiate the traffic in all the configurations with $\rho = 1.20$ and $\rho = 1.40$.

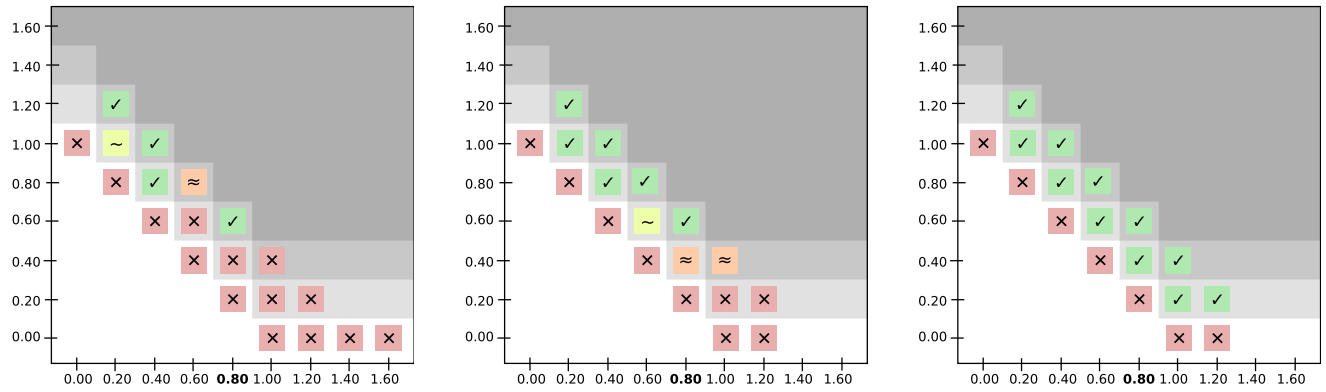
When looking more closely at the performance metrics, we see that differentiation is variable depending on the traffic parameters. Discrimination is reflected in the different values of blocking rate bl and packet loss lo. It is necessary in all cases to eliminate at least the extra load ($bl + lo > (\rho - 1)/\rho$). In cases without discrimination, this is realized uniquely by blocking and utilization remains relatively low. In others, flows with the higher rate lose packets reducing the blocking proportion (which is the same for both flow classes) and utilization is close to 100%. Discriminated flows are served depending on the available bandwidth left by priority flows, and their quality of service is ensured by the lower bound on the long term fair rate.

Episodic discrimination occurs as follows. The start of the simulation corresponds to a transient regime where the priority load includes all incoming traffic. As the link begins to saturate, the highest peak rate flows become backlogged and cease to contribute to priority load. Note that the MBAC allows the link to attain this regime since the variance estimate is given by the Poisson approximation. This leads to overflow implying higher rate flows become backlogged. Once differentiation has occurred, PL decreases back to a value allowing for new flow arrivals. This process continues until the MBAC no longer accepts any more flows.

Sometimes, the traffic mix is such that the MBAC does not allow the link to become saturated. This arises generally when the rate variance estimate is low because both peak rates are relatively small or demand from the high rate flows is slight. Cases when p is small compared to the threshold on FR will be favourable to differentiation. In all cases, we see that flows with peak rate lower than p are protected, since they have negligible overflow and backlog probability, and they encounter no losses.

F. Impact of jitter

In practice, flows of some nominal peak rate r are subject to variable packet delays and acquire jitter as they are multiplexed in successive router queues. It is important to understand how this impacts MBAC performance. A worst case assumption, according to the so-called negligible jitter conjecture in [2], is that flows emit packets as a Poisson process of rate r during on periods. In



(a) (b) (c)

Fig. 5. Differentiation realized by Poisson and $k = 1$ (a), and MinVar with $k = 1$ (b) and $k = 2$ (c). In each figure, the target peak rate is $p = 100$kb/s, and the load of flows of class 1 ($r_1 = 50$kb/s) and 2 ($r_2 = 300$kb/s) is represented respectively on the x- and y-axis.

practice this is likely to be much worse than realized jitter, especially in a network with Cross-protect routers where fair queuing tends to restore the original packet spacing of jittered flows. However, it is of some independent interest to evaluate the MBAC under Poisson traffic at packet level.

We have re-evaluated the performance of both algorithms in the previous scenarios when flows suffer jitter. For lack of space, we can only highlight here our main observations.

When the flow rate $r < p$, jitter is not high enough to have an impact on performance. The difference occurs when r is close to p. Packets of jittered flows are temporarily delayed by the scheduler when their instantaneous rate is higher than p. This enables the MBAC to saturate the link when traffic characteristics allow. New flows are then blocked by the FR condition, which is a sign of high load. While GT only considers the traffic aggregate, CrossProtect temporarily delays packets which have a high instantaneous rate to leave some room for admitting new flows. In this way, it actually reduces the degree of jitter, which is a supplementary advantage of the differentiation.

G. Performance in flash crowds

We now consider the non-stationary cases introduced in Section III-D, limiting discussion to simple scenarios where all traffic is UDP[3]. The flows have the same properties as in the previous homogeneous case (on/off arrival pattern, exponential size distribution). The topology simulated is presented in Figure 2 and consists of two parallel dumbbells that share the same central link in case of failure. The initial load on both is 60%. The simulation duration is 1000s; the central link of the second dumbbell fails after 500s which causes the rerouting of all traffic on the other link.

We set $p = 100$kb/s and consider flows of peak rate $r_a = 100$kb/s (a worst case for performance). Figure 6 plots the evolution of the FR (left) and PL (right) estimators alongside their instantaneous values. After the failure, rate r_a flows are backlogged for a few seconds until the link regains the stationary regime presented in Section IV-D. The instantaneous fair rate becomes less than p, which indicates that too many flows were admitted. During the flash-crowd event, the blocking of flows is solely due to the limiting condition introduced in Section III-D. Adding the value of the target peak rate p to the measured load A_t thus corrects the estimator which is now updated quickly enough due to the smoothing factor. It would otherwise be ineffective since the critical timescale is no longer representative of the rate of flow arrivals and departures. Simulations confirm that performance is much more severely degraded if we remove this additional safeguard.

In order to understand the impact of the peak rate on the performance of streaming flows, we now also consider flows of rate $r_b = 50$kb/s. Figure 7 plots the streaming backlog probability with flows of rate r_a (left) and r_b (right), with both algorithms and a number of different parameter settings. During a few seconds after the rerouting, a large fraction of streaming flows are backlogged. With flows of rate r_b, the backlog probability is much smaller, which is encouraging since the peak rate p would in practice be set considerably higher than the actual rate of the flows to be protected.

While encouraging, these preliminary results clearly illustrate the difficulty of controlling traffic through admission control in the considered fail-over scenario. The condition that limits the number of incoming flows per slot might prove to be over conservative when flows actually have a peak rate much less than p or, as observed in real traces, if there are many flows consisting of a single packet. Blocking new arrivals after detecting congestion as we do here might not be sufficient for non-stationary regimes. This remark is especially true if we consider more realistic scenarios with TCP flows and heavy-tailed distributions. Future work will consider adjusting smoothing factors

[3]Rerouted TCP flows might fall back to *slow start* mode, and increase their rate progressively after being accepted which, combined with the rate adaptation, makes it less tractable

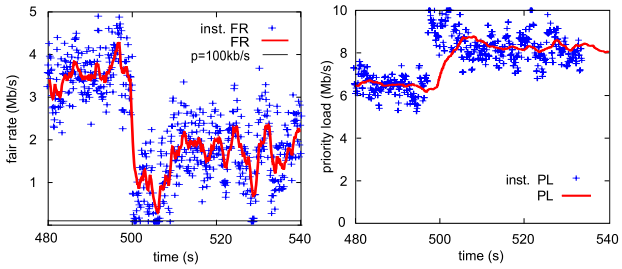


Fig. 6. Evolution of fair rate (left) and priority load (right) during a flashcrowd with MinVar, $k = 2$, $p = 100$kb/s et $r = 100$kb/s

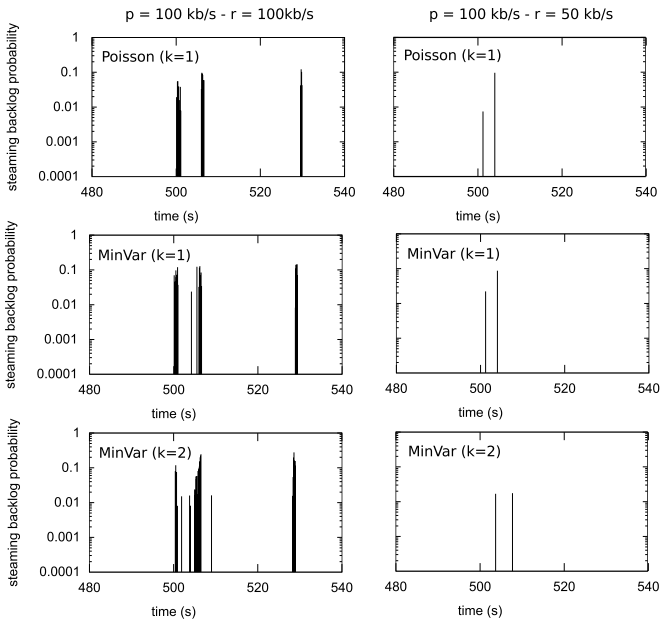


Fig. 7. Streaming backlog probability in two scenarios: $p = 100$kb/s, $r = 100$kb/s (left) and $p = 100$kb/s, $r = 50$kb/s (right), with Poisson, $k = 1$ (top), MinVar, $k = 1$ (middle), and $k = 2$ (bottom).

depending on the change in variance characteristic of such situations, or using change detection algorithms (as in signal theory), provided that those solutions remain simple. However, it is likely that more radical measures are needed to prevent a general performance degradation, including pre-emptive interruption of some flows in progress.

V. Conclusions

Flow-aware networking based on the Cross-protect mechanisms allows performance guarantees for streaming and elastic traffic without the significant complication of having to mark packets for class of service discrimination. It remains, however, to design and calibrate the necessary measurement-based admission control algorithms. Previous work suggests a simple algorithm based on the estimated fair rate is sufficient when a significant proportion of traffic is composed of elastic flows that are bottlenecked at the link in question. However, in a backbone network the majority of flows are bottlenecked elsewhere, notably by the user access links whose rate is usually much smaller than link capacity. Traffic is then typically composed of a mixture of flows with limited peak rate. For some the rate is less than the target rate p (for which flows are intended to be handled with priority) while for others it is greater than p. The challenge is to design an MBAC that allows the latter to become backlogged (they are supposed to be able to adjust their rate) while always preserving priority handling for the former. The challenge is compounded by the fact that the MBAC cannot rely on any knowledge of flow characteristics.

In the paper we have proposed a simple MBAC based on measured mean and variance of load offered to the Cross-protect priority queue. This differs from the algorithm proposed by Grossglauser and Tse [6] in that the variance is used only if it is less than the variance estimated on assuming all flows have the target rate p. This algorithm is shown in our simulation results to enable the required discrimination when flow rates are not too close.

Evaluations in a flash-crowd scenario are less encouraging. Several factors combine to make it difficult to find an adequate compromise between accepting too many new flows, leading a significant period of performance degradation, and being over conservative and rejecting many more flows than is strictly required. Further analysis of this scenario is necessary given its practical importance. Our preliminary results suggest simple admission control may not be sufficient. It is likely that interruption of a subset of flows in progress is additionally necessary in order to restore an acceptable load level.

References

[1] N. Benameur, S. Ben Fredj, S. Oueslati and J. Roberts. Quality of service and flow-aware admission control in the Internet. In Computer Networks, Vol 40, pages 57-71, 2002.

[2] T. Bonald, A. Proutière, J. Roberts. Statistical performance guarantees for streaming flows using expedited forwarding, In Proc. of IEEE INFOCOM 2001

[3] L. Breslau, S. Jamin and S. Shenker. Comments on the Performance of Measurement-Based Admission Control Algorithms. In IEEE INFOCOM 2000, Tel Aviv, Israel, March 2000.

[4] R.J. Gibbens, F.P. Kelly and P.B. Key. A Decision-Theoretic Approach to Call Admission Control in ATM Networks. In IEEE Journal on Selected Areas in Communications, Vol 13, No 6, pages 1101-1114, August 1995. oban1

[5] M. Grossglauser and D. Tse. A Framework for Robust Measurement-Based Admission Control. IEEE/ACM transactions on Networking, 5 (1), February 1997.

[6] M. Grossglauser and D. Tse. A Time-Scale Decomposition Approach to Measurement-Based Admission Control. IEEE/ACM Transactions on Networking, 2003.

[7] S. Jamin, S.J. Shenker and P.B. Danzig. Comparison of Measurement-Based Admission Control Algorithms for Controlled-Load Service. IEEE Infocom 1997.

[8] A. Kortebi, S. Oueslati and J. Roberts. Cross-protect: implicit service differentiation and admission control. Proceedings of HPSR 04, Phoenix, 2004.

[9] A. Kortebi, S. Oueslati and J. Roberts. Implicit Service Differentiation using Deficit Round Robin, ITC19, Beijing, August 29-September 2, 2005.

[10] S. Oueslati and J. Roberts, A new direction for quality of service: Flow aware networking, NGI 2005, Rome, April 18-20, 2005.

[11] N.G. Bean, Estimation and Control in ATM Networks, TRC Report, 1994