Evaluating Fairness in Aggregated Traffic Marking

Igor B. H. de A. Alves, José F. de Rezende, Luis Felipe M. de Moraes
Laboratório de Redes de Alta Velocidade (RAVEL) & Grupo de Teleinformática de Automação (GTA)
COPPE - Universidade Federal do Rio de Janeiro (UFRJ)
Caixa Postal 68511 - 21945-970 - Rio de Janeiro - RJ - Brasil
igor@ravel.ufrj.br, rezende@gta.ufrj.br, moraes@cos.ufrj.br

Abstract- **This article analyses the behavior and evaluates the performance of a fair traffic marker implementation, called Fair Marker (FM), as proposed in [1]. The implementation exploits the duality between buffer allocation and token consumption in a token bucket marker in order to enforce fairness among different flows originated from the same subscriber network in a differentiated services (diffserv) domain. The results obtained show that fairness can be achieved when parameters are correctly set. In this paper, we establish well-defined guidelines to help configure the FM parameters. The results also show that the FM cannot provide fairness in the allocation of excess bandwidth. In order to overcome this problem, we present an extension to the original proposal called Three-Color Fair Marker (TCFM). As an additional contribution, marking strategies proposals are discussed and classified.**

I. INTRODUCTION

The need to offer different service levels on the Internet has encouraged the research in differentiated services (diffserv or DS) [2, 3]. The DS proposal is based on a set of simple mechanisms that treat packets differently according to the marking of the DS field in the IP header. Before entering in a DS domain, the field is marked with a certain value (or *codepoint*) that determines the treatment that should be supplied to the packet inside the domain.

In the standardization groups, different treatments (Per-Hop Behaviors or PHBs) are being specified together with the associated codepoints. Two PHBs, now in wide discussion, are the Express Forwarding (EF-PHB) [4] and the Assured Forwarding (AF-PHB) [5].

In order to provide the desired level of service, traffic conditioning is performed on DS boundary nodes. Traffic conditioners may contain markers, meters, droppers and shapers to bring traffic into compliance with an established profile. Several markers were proposed in the literature [1, 6, 7, 8, 9]. The function of these mechanisms is to mark traffic according to the service profile contracted by the user. The behavior of the markers has a great impact on the service level, in terms of bandwidth, obtained by TCP flows that cross a DS domain. This behavior has been studied in several scenarios: TCP flows with different round-trip times (RTTs), with different service expectations (target rates), and in the presence of congestion-insensitive flows (non-adaptive). In the last scenario, problems of fairness were observed in the allocation of excess bandwidth (non-contracted bandwidth) between TCP and non-adaptive flows, such as UDP flows. However, a scenario still few explored is the fairness in bandwidth allocation among flows of an aggregation when marking is performed on the aggregated traffic instead of per flow.

In this work, we present traffic marker classifications based on two criteria; discuss the need of aggregated traffic marking, while pointing out fairness problems this sort of marking brings; and present a solution to the fairness marking problem, called fair-marker (FM) [1]. We describe a possible implementation of the FM using the algorithm specified in [10]. With the objective of evaluating the behavior of the FM in different scenarios and comparing it with other traffic markers, it was implemented in the ns-2 simulator. The results show that a correct tuning of the FM can guarantee a high degree of fairness in the allocation of the assured bandwidth among the traffic flows that compose an aggregation. We present configuration guidelines for achieving these efficiencies. We also propose and evaluate an extension to the FM to address concerns regarding inefficiencies in sharing excess bandwidth.

This article is organized in the following way. Section II gives the fundamental concepts involved with the AF-PHB service and active queue management. In Section III, several proposals for traffic markers are discussed and classified. In Section IV, the fair-marker is described in terms of objectives, operation and implementation used. In Section V, the simulation results are presented and discussed. Finally, in Section VI, the conclusions and the perspectives of this work are presented.

II. FUNDAMENTAL CONCEPTS

The AF-PHB provides the delivery of IP packets in four independent classes, called AFx classes (where, x = 1, 2, 3 or 4). For each class, there is a certain amount of resources, such as buffer and bandwidth, allocated by each DS node. In each AF class, an IP packet can be marked, either by the user or by the DS domain, within three levels of loss precedence (codepoint = AFx1, AFx2 or AFx3). In case of congestion within an

AF class, a congested DS node preferentially discards packets with a higher loss precedence value. Normally, the DS nodes perform active queue management by using RED [11], one for each loss precedence level. Each RED aims to reduce the effects of the congestion before it becomes necessary to discard packets with lower loss precedence values.

In [12], the authors present four general categories of RED policies when multiple loss precedence levels are used in packet marking. These categories originate from the way one calculates the average queue size and sets drop thresholds for each RED algorithm. In this work, we use the Multiple Average/Multiple Thresholds (MAMT) category. For this policy, one average queue size is calculated for each precedence level, where the number of packets of a certain level is equal to the sum of packets of this level and, if it exists, of the inferior levels. In addition, each precedence level has different drop thresholds. For instance, the RIO queue (RED with *IN* and *OUT*) [6], belongs to this category. The average queue size for *IN* packets (in profile) is calculated using solely the number of *IN* packets, while the average queue size for *OUT* packets (out profile) is calculated using the number of *IN+OUT* packets. Different drop thresholds are defined for each level.

III. Traffic Markers

In this section, we focus on several of the strategies used to mark packets in order to classify and understand their differences. Marking strategies can be classified into three categories based on the criteria used for the marking. Devices can mark packets: (i) based on the state of all individual flows of an aggregation, called *per-flow marking*, (ii) based only on the aggregation state, without any knowledge about individual flows, called *per-aggregation marking* or, (iii) based on a partial knowledge of individual flows, called *flow aware per-aggregation marking*.

When per-flow marking is performed on aggregated traffic, the device responsible for marking packets must deal with individual flows states. In this aspect, per-aggregation marking is easier to manage and is more suitable for customers who generate a huge number of individual flows. An example of a customer with this characteristic is a web-server. The number and the dynamics of short-term flows generated by this kind of customer can prevent devices from performing per-flow marking. The large number of states associated with metering needed to perform per-flow marking makes this strategy non-scalable. Furthermore, giving each flow a fraction of the target rate of the aggregated traffic can lead to an inefficient utilization of the reserved bandwidth. In this case, "idle" flows waste their shares while preventing "active" flows from increasing theirs. Nevertheless, per-aggregation marking can introduce unfairness within aggregated flows. The unfairness can be caused by different RTTs, different target rates, different link bandwidth, or different levels of congestion experienced by individual TCP flows within the network. In the flow aware per-aggregation marking category, devices responsible for marking are not aware of how many flows may be marked, not even parameters associated to a particular flow are kept. However, the marker maintains a partial state of the flows being marked. This bounded number of states can be a major advantage in certain scenarios.

Most of studies on diffserv networks deals with per-flow marking [6, 7, 12, 13, 14, 15, 16]. The marking strategies presented in [16] focus on aggregated sources, while providing additional mechanisms to deal with unfairness within aggregated flows. The authors make a comparison of per-aggregation marking with per-flow marking, and three different strategies are proposed to alleviate the unfairness due to different RTTs and target rates. A similar study is made in [16]. In this work, Nandy *et al.* propose some strategies to mitigate the effect of RTTs, UDP/TCP interactions, and different target rates.

Concerning the mechanism used to check the traffic conformity to the service profile, packet marking can be further classified in two broad categories: *token-bucket based marking* and *average rate estimator based marking*. This classification is completely orthogonal to the one earlier described, i.e. all marking strategies can be classified independently according to both criteria.

Token-bucket based marking comprises all strategies that include one or more token-bucket mechanisms measuring the amount of data that individual (or aggregated) flows generate in any time interval. Recent works on diffserv networks mostly use token-bucket based marking [7, 8, 9, 12, 13, 14, 17]. To improve the fairness in the allocation of excess bandwidth between adaptive and non-adaptive traffic inside an AF class, new token-bucket based marking strategies were developed [8, 9]. The effectiveness of three loss precedence levels was evaluated in [12, 13]. In [7], the authors show some advantages of token-bucket based marking in respect to average rate estimator based marking. In [17], Sahu *et al.* makes a performance analysis of token-bucket based marking for TCP by means of an analytical model.

In the average rate estimator based category, marking is performed according to the measurement of the average sending rate of individual (or aggregated) flows. The works in [6, 16, 18] study this marking category. In the initial proposal [6], when the estimator measures an average rate that surpasses a certain threshold for a given flow, packets are marked as *OUT* with a linearly increasing probability. Clark and Fang propose the use of a time sliding window (*TSW*) rate estimator and an intelligent marker. In [16, 18], authors propose some extensions to the *TSW* in order to improve fairness.

IV. The Fair-Marker

The fair-marker (FM) consists of a token-bucket based marker that performs flow aware per-aggregation marking. It focuses on distributing tokens fairly among individual flows from an aggregation. In order to achieve this purpose, it maintains information regarding the consumption of tokens by the monitored flows. Though, to avoid state explosion, FM only keeps states of flows that consumed tokens during the last time interval corresponding to the time needed to fill the token-bucket, denoted $TBFT$ (Token Bucket Fill Time). FM uses an analogy between a token-bucket and a queue, where maintaining states from flows that consumed tokens during last $TBFT$ is similar to keeping states from flows that have packets in a queue. One can imagine a packet consuming tokens as a situation similar to having it substituting the same tokens in the bucket. In practice, FM keeps a complementary queue (with the same size of the bucket) where these *packet traces* are stored. Whenever a token is generated, the *trace queue* is consulted as to whether the number of tokens in the bucket is enough to remove packet traces from the queue. To obtain a fair marker, packet traces are queued according to a fair buffer allocation algorithm.

For each arriving packet, the FM must determine how many packet traces can be removed from the complementary queue. This is equal to the number of tokens accumulated since the arrival of the last packet. After removing these traces, each individual flow that still has traces in the queue has its state updated. These flows are the ones that consumed tokens during the last $TBFT$. Next, packets are marked according to the current number of tokens in the bucket. If this number is insufficient, the packet is marked as *OUT* and its trace is not placed in the complementary queue. Otherwise, the consumption of tokens in last $TBFT$ determines if a packet can consume tokens or not. The fair algorithm is used to determine if the packet trace is queued or not. In case it can be queued, tokens are consumed and the packet is marked as *IN*. Otherwise, the packet is marked as *OUT*.

The fairness in the token distribution achieved by the FM is a function of the fair buffer allocation algorithm used. In our implementation, we use FRED (Flow Random Early Drop) [10], which is a modified version of RED [11]. Besides the two minimum ($minth$) and maximum ($maxth$) thresholds, FRED introduces two new thresholds corresponding to the minimum ($minq$) and maximum ($maxq$) number of packets that each flow can have in the queue. FRED also controls the instantaneous ($qleni$) and the average ($avgq$) number of packets per flow, favoring the flows that have fewer packets than the average. Further, FRED punishes flows that try to exceed the maximum number of packets allowed per flow ($maxq$). More details about the FRED algorithm can be found in [10].

V. Simulations

The first topology we use in our simulations is illustrated in Fig. 1. Four different scenarios are created in order to verify the interaction between TCP and UDP flows composing the same aggregation and the impact of different RTTs: homogeneous TCP (same RTTs) with and without CBR (1 and 2), and heterogeneous TCP (different RTTs) with and without CBR (3 and 4). In all scenarios, the traffic is generated by ten FTP/TCP Reno traffic sources from nodes n to nodes $n + 10$, where $n = 1, 2, ..., 10$. Scenarios including CBR traffic have one

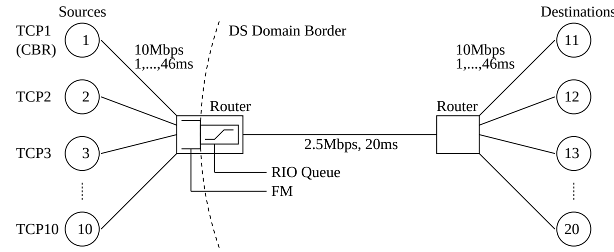

Fig. 1: Topology 1.

(extra) CBR/UDP traffic source from node 1 to node 11, which has a transmission rate of 2.5Mbps (100% of the bottleneck capacity). All packets are 1000 bytes long. The upper bound on the advertised window for the TCP connections corresponds to 90 packets. In the homogeneous TCP scenarios, the propagation delay between sources/sinks and routers is 1ms. In the TCP heterogeneous scenarios, this value varies from 1ms to 46ms in an arithmetic progression of ratio 5ms. Consequently, the minimum RTT for the TCP connections varies from 44ms (TCP1) to 224ms (TCP10). The RIO queue has a capacity of 50 packets ($qlim$), and parameters for *IN* and *IN + OUT* packets are equal to $[0.3 * qlim, 0.6 * qlim, 0.002, 0.1]$ and $[0.14*qlim, 0.3*qlim, 0.002, 0.1]$[1], respectively. The FM has a bucket size (b) of 50 packets and a token rate (r) of 1Mbps (40% of the bottleneck). The parameters of the trace queue $minq$, $maxq = minth$ and $maxth$ assume the values 2, 4, 8, 16, 32 and 50, respecting the inequalities $minq < (maxq = minth) < maxth$, what leads to twenty different configuration sets. The other parameters are maintained constant with values $wq = 0.002$ and $maxp = 0.02$.

In each scenario, we run five simulations for each configuration. The total amount of time for each simulation is 50s. The sources start transmitting at a random time uniformly distributed between 0s and 5s. All the results are computed using the interval from 10s to 50s in order to remove the transient. For each flow (i), the number of packets marked as *IN* and delivered to the corresponding destination (x_i) is calculated. Next, we calculate the fairness index (fi) using (1) [19], where $N = 11, 10$ in the scenarios with and without CBR, respectively.

$$fi = \frac{(\sum_{i=1}^{N} x_i)^2}{N * \sum_{i=1}^{N} (x_i)^2} \qquad (1)$$

With the objective of verifying the influence of FRED parameters on the FM behavior, we compare the twenty configurations using the ranking method described in [19]. In each scenario and for each configuration, the average fi obtained considering the five simulations. From these results, we state configuration guidelines for the FRED trace queue parameters in order to maximize fairness in assured bandwidth allocation.

[1] [minth,maxth,wq,maxp]

(i) lower values of $maxth$ (close to $minth$) degrade the FM performance. In these cases, the maximum average number of *IN* packets in the FRED trace queue decreases, leading to a lack of space for all connections at the same time. In the scenarios with CBR, this non-adaptive connection always occupies its share in the trace queue. On the other hand, TCP connections fight against each other for space due to its bursty nature. In the TCP heterogeneous scenarios, the fairness index also decreases since the TCP connections with longer RTTs are more sensitive to the increased number of packet drops (due to the lower $maxth$). Therefore, we recommend $maxth$ close or equal to b. (ii) when $maxq$ surpasses a certain threshold in comparison with $maxth$, the performance of FM decreases. This is due to the increasing in the number of packets that can be marked as *IN* during a $TBFT$, which reduces the capacity of the FRED algorithm to punish the CBR flow and the TCP flows with smaller RTTs. On the other hand, $maxq$ should not be very small so that TCP flows don't be punished for trying to exceed $maxq$ packets in the trace queue. (iii) $minq$ practically doesn't affect the performance of the FM since the most important issue is that flows don't occupy more than their fair share of space in the trace queue. It is clear therefore that a good performance of the FM in terms of fairness in the distribution of tokens depends on the correct adjustment of its parameters. According to our results, a recommended configuration is $maxth = b$, $minq \leq 10\%b$ and $2 * minq \leq maxq \leq 25\%maxth$.

Next, we compare the FM with the classical token-bucket marker (TB) using the topology of Fig. 2, which depicts a more realistic situation. The monitored traffic consists of ten traffic sources of TCP Reno from nodes 1,...,10 to nodes 51,...,60, and a CBR/UDP traffic source from node 1 to node 51 with a transmission rate of 2.5Mbps (100% of the bottleneck). Ten additional TCP traffic sources, from nodes 11,...,20 to nodes 31,...,40, use the best-effort service and compete with the monitored traffic. The token rates of both markers (r) vary from 200kbps to 2Mbps (8% to 80% of the bottleneck link capacity). The parameters $minq$, $maxq = minth$ and $maxth$ assume the values 2, 8, 8 and 50, respectively. The RIO parameters for *IN* and *IN + OUT* packets are equal to $[0.5 * qlim, 0.8 * qlim, 0.002, 0.02]$ and $[0.2 * qlim, 0.5 * qlim, 0.002, 0.1]$ respectively. Five simulations are run for each value of r.

Fig. 3 shows the fairness in the assured bandwidth sharing. The bars in each point define the confidence interval of 95%. The FM performs better than the TB, obtaining fairness indexes above 0.9 for values of r up to 50% of the bottleneck capacity. The TB, independently of the value of r, gives low fairness indexes (about 0.1) , i.e. the CBR flow practically obtains all the assured bandwidth. Fig. 4 shows the fairness indexes for the excess (*OUT* packets) bandwidth sharing. The FM and the TB have the same performance in terms of excess bandwidth since both mechanisms treat *OUT* packets in the same way, without any action in the sense of guaranteeing fairness. Therefore, the FM needs an additional mechanism to deal with the problem of the excess bandwidth allocation.

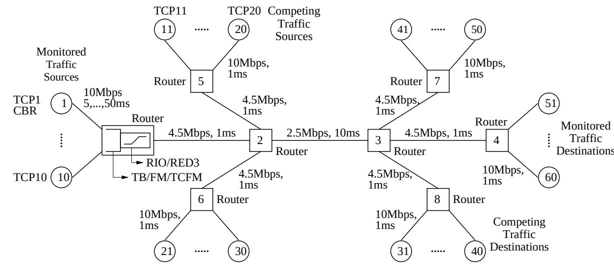

Fig. 2: Topology 2.

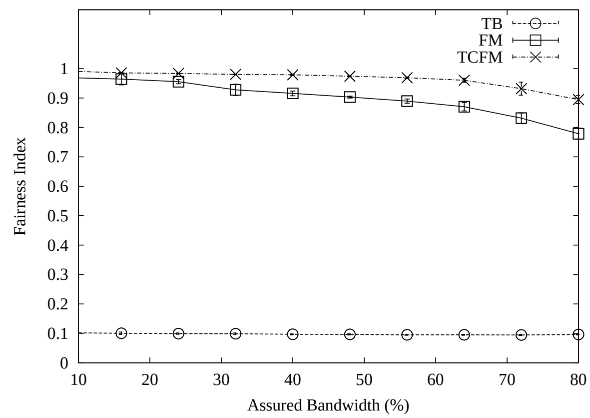

Fig. 3: Fairness in the Assured Bandwidth Sharing

In order to overcome this deficiency, we propose an extension to the FM, called Three Color Fair-Marker (TCFM). The TCFM is obtained from the FM by adding another (yellow) token-bucket with its corresponding FRED trace queue. The green and yellow token-buckets try to fairly distribute their tokens so as to provide fairness in the assured and excess bandwidth, respectively. A packet will be marked as *GREEN* if there are enough tokens in the green bucket and it can be queued in the green trace queue. A packet will be marked as *YELLOW* if at least on condition above is not satisfied and the same things happen for the yellow bucket and its trace queue. Otherwise, a packet will be marked as *RED*. In [20], a very similar marker was proposed. It is a fair version of the Two Rate Three Color Marker described in [9], obtained by embedding a FRED trace queue in each one of its token-buckets.

For the purpose of evaluating the TCFM we test it under the same situation. However, since we are making use of three loss precedence levels, we replace the RIO queue by a RED3 queue with parameters values of $[0.6 * qlim, 0.8 * qlim, 0.002, 0.025]$ for *GREEN* packets, $[0.4 * qlim, 0.6 * qlim, 0.002, 0.05]$ for *GREEN+YELLOW* packets, and $[0.2 * qlim, 0.4 * qlim, 0.002, 0.1]$ for *GREEN+YELLOW+RED* packets. The green profile rate CIR (Committed Information

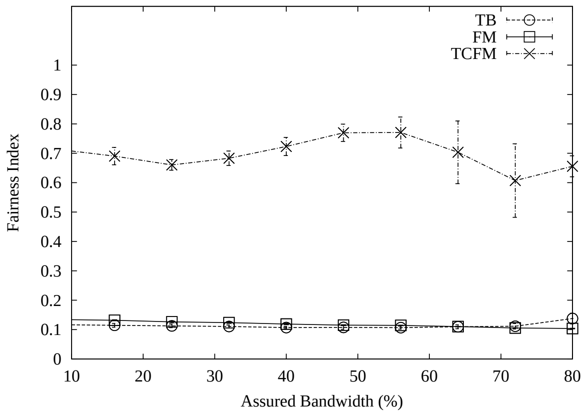

Fig. 4: Fairness in the Excess Bandwidth Sharing

Rate) is varied from 8% to 80% of the bottleneck capacity (as before) while the yellow profile rate EIR (Excess Information Rate) always correspond to 2.5Mbps - CIR. The green bucket size CBS (Committed Burst Size) and the yellow bucket size EBS (Excess Burst Size) are equal to 50 packets. Both FRED trace queues have the same settings of the FM trace queue.

These preliminary results show that the TCFM provides the same or better performance than the FM in the assured bandwidth sharing (Fig. 3). Furthermore, the TCFM provides a considerable improvement in the excess bandwidth sharing among the flows of the same aggregation (Fig. 4). This improvement can be explained by the impact of having a second FRED trace queue to fairly distribute yellow tokens among the flows. Consequently, this improvement in performance reflects in the total bandwidth sharing.

VI. Conclusions

In this work, we classified the different types of existent markers, pointing out the need of marking aggregated traffic in the entry of a DS domain and the problem of fairness among the flows that compose an aggregation.

Next, we presented an implementation of the fair traffic marker defined in [1], which uses the FRED active queue management algorithm so as to obtain fairness among the flows of an aggregation in a DS domain implementing the AF-PHB. As a result from a first study, we presented configuration guidelines for the FM parameters, since it was evidenced that the FM performance can be degraded as function of an inadequate adjustment of its parameters.

Finally, in a second study concerning the assured bandwidth sharing among flows of an aggregation, it was shown that the FM outperforms the classical token-bucket. However, in terms of the excess bandwidth sharing, the FM is unable to assure fairness since no differentiated treatment is supplied to the *OUT* packets. Then, we presented and evaluated an extension to the FM, called Three Color Fair-Marker (TCFM). The TCFM provides significant improvements in the excess bandwidth allocation. Future plans include a more deep analysis of the FM and the TCFM, using formal methods described in [19], regarding the adjustment of their parameters and their behavior in scenarios in which other factors of interest are considered.

References

[1] H. Kim, "A fair marker," Internet draft, work in progress, Apr. 1999.

[2] K. Nichols, S. Blake, F. Baker, and D. L. Black, "Definition of the differentiated services field (DS field) in the IPv4 and IPv6 headers," RFC 2474, Dec. 1998.

[3] S. Blake, D. L. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for differentiated services," RFC 2475, Dec. 1998.

[4] V. Jacobson, K. Nichols, and K. Poduri, "An expedited forwarding PHB," RFC 2598, June 1999.

[5] J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski, "Assured forwarding PHB group," RFC 2597, June 1999.

[6] D. D. Clark and W. Fang, "Explicit allocation of best effort packet delivery service," *IEEE/ACM Transactions on Networking*, vol. 6, pp. 362–373, Aug. 1998.

[7] J. Ibanez and K. Nichols, "Preliminary simulation evaluation of an assured service," Internet draft, work in progress, Aug. 1998.

[8] J. Heinanen, T. Finland, and R. Guerin, "A single rate three color marker," Internet draft, work in progress, May 1999.

[9] J. Heinanen, T. Finland, and R. Guerin, "A two rate three color marker," Internet draft, work in progress, May 1999.

[10] D. Lin and R. Morris, "Dynamics of random early detection," in *Proceedings of ACM SIGCOMM'97*, pp. 127–137, Cannes, France, Sept. 1997.

[11] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, pp. 397–413, Aug. 1993.

[12] M. Goyal, A. Durresi, P. Misra, C. Liu, and R. Jain, "Effect of number of drop precedences in assured forwarding," Internet draft, work in progress, June 1999.

[13] O. Elloumi, S. D. Cnodder, and K. Pauwels, "Usefullness of three drop precedences in assured forwarding service," Internet draft, work in progress, July 1999.

[14] J. F. de Rezende, "Assured service evaluation," in *Proceedings of IEEE Globecomm'99*, Rio de Janeiro, Brazil, Dec. 1999.

[15] S. Sahu, D. Towsley, and J. Kurose, "A quantitative study of differentiated services for the Internet," in *Proceedings of IEEE Global Internet'99*, Rio de Janeiro, Brazil, Dec. 1999.

[16] B. Nandy, N. Seddigh, P. Pieda, and J. Ethridge, "Intelligent traffic conditioners for assured forwarding based differentiated services networks," in *Proceedings of IFIP High Performance Networking (HPN 2000)*, Paris, France, June 2000.

[17] S. Sahu, P. Nain, D. Towsley, C. Diot, and V. Firoiu, "On achievable service differentiation with token bucket marking for TCP," in *Proceedings of ACM SIGMETRICS'00*, Santa Clara, CA, June 2000.

[18] W. Lin, R. Zheng, and J. Hou, "How to make assured services more assured," in *Proceedings of the 7^{th} International Conference on Network Protocols (ICNP'99)*, Toronto, Canada, Oct. 1999.

[19] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. New York, NY: John Wiley and Sons Inc., Apr. 1991, pp. 36 & 387.

[20] L. W. Ilias Andrikopoulos and G. Pavlou, "A fair traffic conditioner for the assured service in a differentiated service internet," in *Proceedings of IEEE International Conference on Communications (ICC 2000)*, New Orleans, LA, June 2000.