

# Source and channel adaptive rate control for video streaming over Diffserv networks.

Reda B. Hosny and Christine Guillemot  
IRISA-INRIA  
Campus Universitaire de Beaulieu, 35042 Rennes cedex, France.  
rhosny@irisa.fr

## Abstract

This paper addresses the issue of quality of service of video streaming applications over networks supporting service differentiation, according to the architecture defined by the IETF. After studying a proper setting of key mechanisms such as scheduling and queueing management, this paper focuses on the problem of source rate control that would capture and would adapt to the behavior of AF (Assured Forwarding) channels. The ultimate goal is to optimize the bandwidth usage within the limits of the negotiated service level agreement (SLA). For this we derive an analytical model of AF channels throughput. This model relies on equations governing traffic conditioning and queueing management at the edge router of the diffServ domain. The overall strategy can be regarded as an open loop rate control based on the TBF fluid model as well as on multi-level RED rate control equations. Experimentation results with H.263+ non-layered and layered (with SNR scalability) compatible video streams evidence the benefits in terms of PSNR, and of graceful degradation in presence of congestion, against a simple TMN8 constant bit rate video source rate control.

## 1 Introduction

Delivering temporally-constrained and high rate continuous streams over a network offering no guarantees in terms of bandwidth, packet loss or delay jitter is a very challenging issue faced today by both the networking and the video coding communities. A first trend in order to optimize the end-to-end quality of service (QoS) consists in adapting the applications to the network characteristics. A number of experiments have already shed the light on the benefits of such adaptation mechanisms. By handling delay jitter, packet loss and adapting the source rate accordingly, they allow to improve the quality of video communications over best-effort networks. A second trend consists in deploying new scheduling and queueing disciplines in order for the network to better meet multimedia streams QoS requirements. In this context, the IETF is studying models and architectures for supporting differentiated services. The key idea advocated by the DiffServ group of the IETF consists, by using a priority mechanism, in offering services with gradually increasing performances. In the DiffServ architecture, traffic is classified, metered and marked at the edge of the network. Traffic with similar QoS requirements is marked, by setting the DiffServ code point (DSCP) in the header of each IP packet, and is treated in an aggregated fashion in the core routers. The DSCP specifies a per-hop behavior (PHB) to be applied to the packets of a given class within a provider's network (in the core network). Therefore, the core routers do not need to maintain per flow states, but use instead scheduling

and buffer management for aggregates of flows. A PHB denotes a combination of forwarding, classification, scheduling, and drop behaviors at each hop. Two PHB's have been standardized [1], [2] so far: the Expedited Forwarding (EF) PHB for support of delay and jitter sensitive traffic and the Assured Forwarding (AF) set of PHBs. The set of AF PHBs is intended to provide different levels of forwarding assurances for IP packets at a node and to be used to support multiple priority service classes.

In the following we concentrate on AF classes, assuming that the EF class will be reserved for critical low bandwidth streams (e.g. audio, voice). We consider an architecture where traffic shaping and conditioning at the edge routers is based on a Token Bucket Filter (TBF) mechanism. The assured forwarding per-hop-behavior (AF PHB) in the core routers is supported by a class based queueing (CBQ) scheduler coupled with a queueing management policy based on multi-level RED operating within a shared virtual buffer. The paper first studies working parameters of multi-level RED, in order to support PHBs in AF classes that would be suitable for video streaming. In particular, we estimate the appropriate values of dropping probabilities that would allow to guarantee isolation and fairness between classes. The corresponding operating region of multi-level RED is said to be "stable".

Before the communication starts, the source negotiates its traffic characteristics and gets from the bandwidth broker the corresponding service parameters (delay, losses, bandwidth). The parameters of the marking and shaping functions applied at the edge router are then set accordingly. Considering here a TBF for marking and shaping, the key parameters are the assured rate and the maximum burst size of the TBF mechanism. To encode a video source with a constant bit rate constraint leads to very bursty streams. The bursts may exceed the maximum burst size of TBF, hence lead to packet losses. On the other hand, an increase of the maximum burst size parameter of TBF will impede any real shaping, and the bursty nature of the resulting output streams will induce higher loss rates for lower priority streams. Therefore, we seek here an optimal usage of the resources available within the limits of the negotiated SLA.

We derive a model of the AF channel throughput. This model relies on equations governing token bucket and multi-level RED functions. The resulting model coupled with a source marking strategy is used to control the rate of the video source. This can be seen as a source adaptive traffic shaping. The procedure leads to smoother traffic while conforming to real-time video constraints. The overall strategy can be regarded as an open loop rate control based on the TBF fluid model as well as on multi-level RED rate control equations. The performance of the approach is first estimated in terms of PSNR against a constant bit rate (CBR) system based on the TMN8 [3] rate control mechanism for H.263+ compliant mono-layered streams. In a second step, the TBF based rate control algorithm is applied on the base layer of H.263+ SNR scalable streams. The rate of the enhancement layer is varied in order to have an overall constant rate in average. The impact of the number of marking levels is also studied. The results obtained evidence graceful degradation over a wide range of loss rates, as well as the ability to handle short-term congestion. The analysis also leads to guidelines for configuring buffer management and for setting layering parameters for video streaming applications over DiffServ environments.

## 2 Background and Diffserv architecture

### 2.1 Elements of service differentiation

Key performance parameters for continuous media applications such as video streaming over the Internet are end-to-end delay, delay jitter, loss rate and bandwidth. The above parameters will

affect differently the various multimedia applications.

Queuing disciplines and scheduling techniques have a strong impact on the achievability of the above performance criteria. The simplest approach for selecting a packet to drop is the well-known tail-dropping mechanism. Because packets are dropped deterministically from the end of the queue, sub-optimal traffic patterns can emerge [4]. In order to increase the overall network throughput while maintaining low delays, active queue management mechanisms have been introduced. The usage of active queue management - random drop [5], early packet discard, early random drop [6], and random early detection (RED) [4] - assumes that the routers have the capability of detecting incipient congestion. A RED gateway detects congestion by monitoring the average queue size and randomly discards packets if the average queue size exceeds some lower bound so as to notify connections of incipient congestion. The dropping probability increases with the average queue size. However, these mechanisms, as such, are not amenable to support service differentiation. A queue management algorithm called RIO - RED with IN and OUT - has been defined, as an extension of RED, in order to discriminate low priority (out-of-profile) packets from high priority (in-profile) packets in times of congestion. By supporting two RED algorithms with different levels of dropping probability, RIO allows to perform preferential dropping of out-of-profile packets over in-profile packets. This mechanism extended to a higher number of levels is known as multi-level RED.

Tuning the parameters of queue management algorithms in order to satisfy appropriate service differentiation criteria, is not an easy task. In the following, we first try to study some key Multi-level RED parameters in order to satisfy isolation and fairness between classes, and will rely in the sequel on these parameters to derive and validate an analytical model of throughput of video UDP streams in AF channels.

## 2.2 Building blocks of the Diffserv architecture

Before analyzing the operation and behavior of queuing disciplines, in order to eventually set the corresponding parameters, we first describe the main elements of the Diffserv architecture that we have considered and that form the basis of our experimentation testbed. At the edge routers, a classifier reads the DiffServ codepoint (DSCP), if already set, selects and routes packets to a traffic conditioner (TC). Fig. 1 depicts the structure of a traffic conditioner. The classifier is a mechanism used to select the PHB for a traffic flow. In the following we will use class based queueing (CBQ). CBQ is a priority queuing mechanism that is used here to classify packets at the level of aggregated flows.

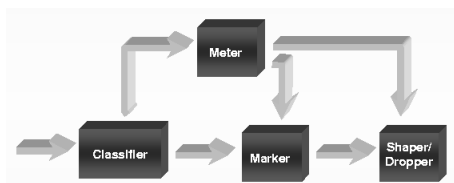


Figure 1: Traffic Conditioner Building Blocks

The role of the TC is to ensure that the flows are in line with the SLA, by monitoring the traffic of each stream (meter) to see if it is within the required profile, by triggering re-marking, dropping, or shaping, if it is out of profile. The metering block measures each traffic stream, informs the marker, shaper, and dropper mechanisms about the quantitative and qualitative status of the stream in order to trigger the corresponding policies.

An Exponential Weighted Moving Average (EWMA) estimator is used to calculate the expected rate at time  $t$ . It estimates the average sending rate at each packet arrival as  $Avg = (1 - w)avg + w * diff$ , where  $diff$  is a function of the time given by ( $diff = t - s/R_a$ ). The term  $s$  denotes the size of the recently-transmitted packet in bytes,  $R_a$  is the TBF assured rate, and  $t$  is the inter-departure time between the current packet and the previous one. The quantity  $w$  is a weighting factor used to adjust the sensitivity of the algorithm to traffic fluctuations.

The marking and shaping functions are supported by a TBF mechanism coupled with a multi-level RED queueing discipline. On the basis of the average rate measured by the EWMA estimator, the marking function proceeds as follows : It marks the packets with the colors *red*, *yellow* or *green*, depending on the measured throughput of the stream against the assured rate  $R_A$  and the maximum burst size or peak rate controlled by the TBF depth. If the estimated average rate is less than or equal to the CIR, packets of the stream are marked *green*. If the estimated average rate is greater than the CIR but less than or equal to the maximum burst size, packets are marked *yellow* with probability  $P_0$  and *green* with probability  $(1 - P_0)$ .  $P_0$  is the fraction of packets contributing to the measured rate beyond the CIR. If the estimated average rate is greater than the maximum burst size, packets are marked *red* with probability  $P_1$ , *yellow* with probability  $P_2$  and *green* with probability  $(1 - (P_1 + P_2))$ .  $P_1$  is the fraction of packets contributing to the measured rate beyond the maximum burst size.  $P_2$  is the fraction of packets contributing to that part of the measured rate between CIR and maximum burst size. An AF class is therefore configured with a committed rate and a threshold above which the classifier begins to assign a higher drop precedence to the packets. This (re)-marking function coupled with multi-level RED leads to a shaping of the streams, ensuring that the traffic stream which is forwarded is in conformance with the SLA.

The absorption of the bursty arrival of packets characterizing video flows may require the maximum burst size to be set to a large value. However, a large value of this parameter may in turn contribute to increase the burstiness of the output streams at the expense of an increased loss rate for lower priority streams. We show in the following that a proper video source rate control may overcome this difficulty. By tailoring the shaping of outgoing video streams to AF channel characteristics (AF channel throughput), smoother traffic can indeed be obtained while maximizing the bandwidth usage and at the same time conforming to the SLA.

### 3 DiffServ testbed and parameter setting

The testbed consists of four machines a server, a client and two LINUX PC-based routers supporting the EPFL DiffServ kernel (version 8) [7]. The links used between the two routers are 10Mbps Stackable Hub. The server and client machines are used as respectively sources and sinks for traffic generators. The TTCP and NETPERF tools have been used for generating TCP and UDP traffic. A modified version of TTCP is actually used for marking streams at the source. MPEG video traffic is emulated by using Netspec (3.0). It gives the ability to create parallel sessions of MPEG traffic with various DSCP values.

Several tests have been carried out in order to validate isolation and fairness between TCP and MPEG UDP traffic. The first tests carried out aim at finding working parameters for Multi-level RED queues that would ensure immunity of low rate traffic flows competing with higher rate traffic flows in the same AF class. We have set the queue limits and weighting EWMA factor for AF11 to values recommended in the literature [8], i.e.  $W_p = 0.002$ ,  $Thresh_{min} = 12Kbytes$  and  $Thresh_{max} = 32Kbytes$ . The average packet size has been set to 1000 bytes. The values for

AF12 and AF13 are shifted in order to avoid overlapping of the thresholding values as shown in Fig. 5. We have then studied the effect of the dropping probability factor ( $D_p$ ) on the properties of isolation and fairness between TCP and MPEG UDP traffic. The tests are carried out for the AF1X class in presence of competing EF (virtual leased lines) CBR traffic sources.

The experiments have shown that for a proper setting of multi-level RED thresholds there is a range of values for the dropping probability factor  $D_p$  ( $0.01 \leq D_p \leq 0.1$ ) guaranteeing on one hand isolation for different AF class traffic flows in an over-provisioned<sup>1</sup> network while at the same time assuring fairness between these streams in an under-provisioned network. Fig. 2 shows the achieved rate with respect to the target rate in different classes for different values of the dropping probability factor  $D_p$ . We can observe (see fig. 2) that below this value RED is too gentle and the congestion state will be very likely reached, hence AF11 as well as AF12 and AF13 will suffer from severe losses (congestion dropping). Also above this value RED is too aggressive and leads to unpredictable losses even without reaching a congestion state (early dropping). The corresponding operating mode of RED will be called the “stable” region of RED in the sequel. In the following, we have set this parameter to  $D_p = 0.02$  in a bounded rate traffic class. The tests have also shown that for the special case of low rate traffic, the immunity against high rate traffic can be guaranteed even for other values of dropping probabilities. This case corresponds to the state when RED can absorb all traffic dynamics due to its low rate nature and the dropping probability function is almost never used.

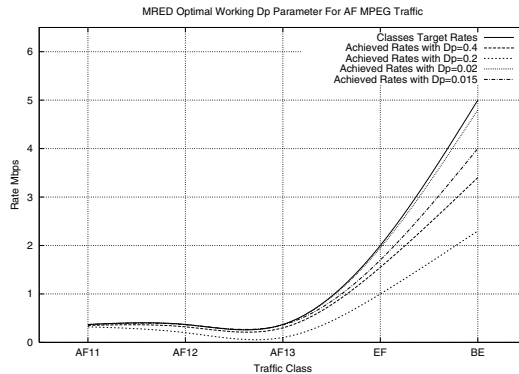


Figure 2: Study of MRED dropping probability with respect to isolation and fairness between TCP and MPEG UDP traffic.

## 4 Analytical model of TBF-MRED based AF channel throughput

The encoding of video sources with CBR constraints leads to streams of a bursty nature. The burst may indeed exceed the TBF depth  $L$ , leading to packet losses. In the following we try to adapt the video source to the AF channel characteristics, within the limit of the negotiated SLA, in order to avoid exceeding the maximum burst size of the TBF, while at the same time maximizing the bandwidth usage. The goal is therefore to design a source rate control that would capture and adapt to the characteristics of the AF channel. In order to do so, we consider the architecture depicted in fig. 3. The architecture comprises a rate control and a shaper/marker

<sup>1</sup>Over-provisioned means that the total available bandwidth is higher than the cumulated requested bandwidth

on the server side that will “mimic” the behavior of the shaper (TBF) /marker coupled with Multi-level RED on the edge router.

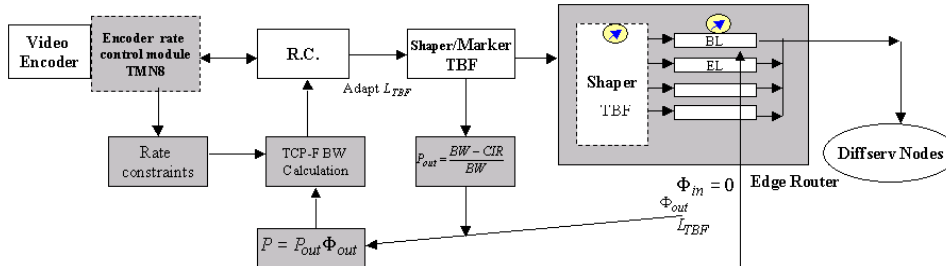


Figure 3: Rate control building blocks.

The design of a network adaptive source rate control requires first to derive an analytical model of the AF channel throughput. In order to do so, we analyze the behavior of TBF+MRED-based traffic marking and shaping performed on the edge router.

#### 4.1 Traffic shaping in RIO queues

In this analysis, for sake of simplicity, we consider a two-level RED, called RIO. Traffic shaping in RIO queues is governed essentially by marking and dropping probability of in and out of profile packets. Let  $\phi_{in}$  and  $\phi_{out}$  be the probabilities of marking packets respectively in-profile and out-of-profile, and let  $P_{in}$  and  $P_{out}$  be the multi-level RED dropping probability values for respectively in-profile and out-of-profile packets. The overall loss rate can then be expressed as

$$P = P_{in}\phi_{in} + P_{out}\phi_{out}. \quad (1)$$

The dropping probability for out-of-profile packets is given by

$$P_{out} = \frac{r_{avg} - CIR}{r_{avg}} \quad (2)$$

where  $r_{avg}$  is the average bandwidth estimated by the EWMA rate estimator, and where  $CIR$  is the committed rate negotiated for a given AF class. We first assume that the probability of dropping in-profile packets is given by  $P_{in} = 0$  (this corresponds to an under-subscribed<sup>2</sup> network). This leads to

$$p = P_{out}\phi_{out} = \frac{r_{avg} - CIR}{r_{avg}} = \phi_{out} - \phi_{out} \frac{CIR}{r_{avg}}. \quad (3)$$

#### 4.2 TBF+RED marking/shaping equations

The goal is to derive an analytical model of the video UDP based traffic in AF channels implemented with TBF+RED. We consider first the equations modelling the TCP throughput on a channel comprising TBF+RED [9]

$$R_{TBF} = \begin{cases} \frac{1}{2}(R_A + \sqrt{R_A^2 + \frac{6}{\phi_{out}RTT^2}}) & R_A \leq \frac{1}{RTT}(\sqrt{2(L + \frac{1}{\phi_{out}})} + 2\sqrt{2L}) \\ \frac{3R_A}{4} + \frac{3\sqrt{L+\phi_{out}}}{2\sqrt{2*RTT}} & R_A > \frac{1}{RTT}(\sqrt{2(L + \frac{1}{\phi_{out}})} + 2\sqrt{2L}) \end{cases} \quad (4)$$

<sup>2</sup>Under-subscribed network means  $P_{in} = 0$  and  $P_{out} > 0$

and show that, for a given range of the TBF depth  $L$ , these equations turn out to provide a valid model for the AF channel. The above set of equations evidences two phases: a flow phase (first equation) and a saturation phase (second equation). Targeting a flow phase operating mode, we consider the first equation, i.e,

$$R_{TBF} = \frac{1}{2}(R_A + \sqrt{R_A^2 + \frac{6}{\phi_{out}RTT^2}}) \quad (5)$$

$$\text{if } R_A \leq \frac{1}{RTT}(\sqrt{2(L + \frac{1}{\phi_{out}})} + 2\sqrt{2L}), \quad (6)$$

where  $R_A$  represents the TBF assured rate parameter. The quantities  $RTT$  and  $\phi_{out}$  represent respectively the round trip time and the RED dropping probability value.

The non linear relationship between the TBF depth value  $L$  (packets) and the increase in total throughput  $R_{TBF}$ , ( $R_{TBF} \propto \sqrt{L}$ ) for a dropping probability value  $Dp = 0.02$ , is depicted in Fig. 4. There is a limit value  $L_{max}$  beyond which the increase of  $L$  has no more effect on the throughput. This boundary value corresponding to the limit between the flow and saturation phases is not fixed, it depends on the dropping probability.

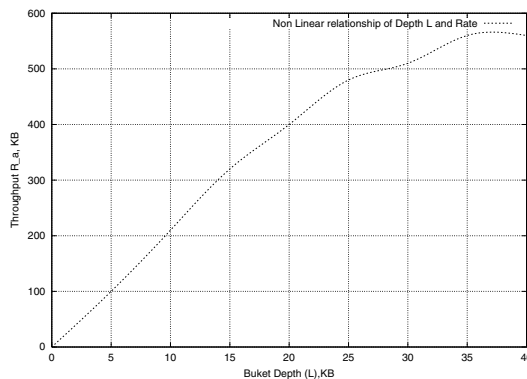


Figure 4: Non linear relation between  $L$  and  $R_a$ .

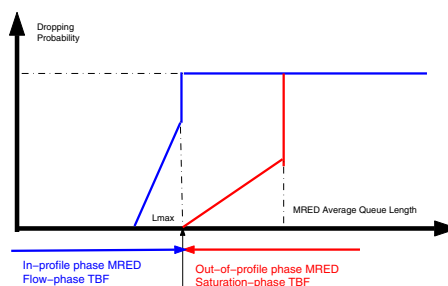


Figure 5: Multi-level RED flow and saturation phases.

In order to show that the above equation is a valid model of an AF channel, we have made several experiments with MPEG UDP streams in AF channels based on a TBF+RED implementation. For different values of the dropping probability  $Dp$ , we have increased the TBF depth  $L$  up to the observation of saturation. At the point of reaching saturation, the maximum depth  $L_{max}$  and the average rate have been measured. Fig. 6(a) shows these maximum values of the

TBF depth  $L_{max}$  for the “stable” region of TBF+RED, i.e., when fairness and isolation between classes are guaranteed and TBF is in a flow phase. Fig. 6(b) shows the corresponding achieved rate for different values of dropping probability  $D_p$ , and for the maximum TBF depth values  $L_{max}$  shown in Fig. 6(a). It can be verified easily that the achieved rate measured turns out to be very close to the rate given by Eq. 5, showing that Eq. 5 is a valid analytical model of the throughput of video UDP-based traffic in the AF channel implementation based on TBF+RED.

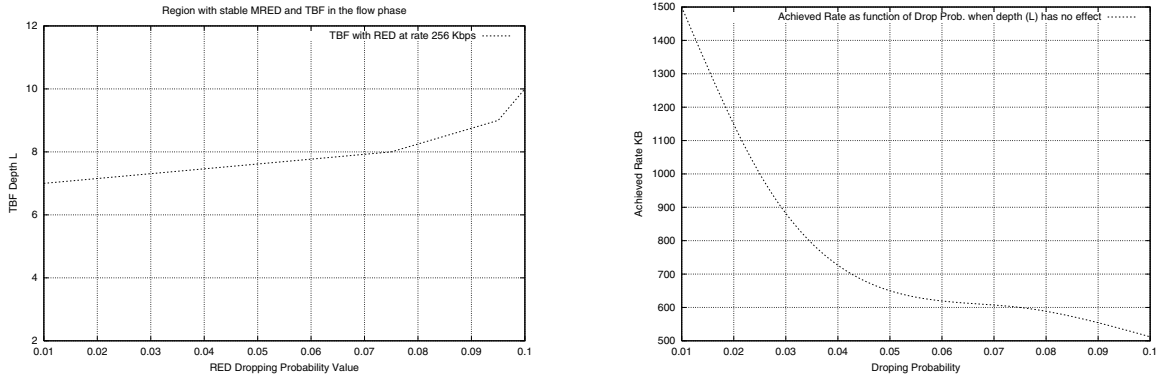


Figure 6: a-Relation between TBF buffer depth and RED dropping probability for the AF channel. b-TBF Achieved rate in terms of the dropping probability for the TBF depth  $L$  corresponding to RIO saturation phase limit.

## 5 Modelling the throughput of TCP-compatible aggregated flows in AF channels

In order to further validate the above analytical model, we show that, in the particular case of traffic isolation and fairness, of an under-subscribed domain and provided that the TBF depth is lower than the maximum value  $L_{max}$  (corresponding to the saturation point), the rate control equation given in Eq. 5 turns out to be equivalent to a throughput model of an aggregation of TCP-compatible flows.

Indeed, let us consider the TCP-compatible rate control equations for an aggregate of flows expressed as <sup>3</sup>

$$R_{TCPC} = \sum_{i=1}^N \frac{MSS}{RTT_i} \frac{C}{\sqrt{p_i}} \quad (7)$$

where  $C$  is a constant equal to 1.22. Assuming that the aggregated flows have the same  $RTT$  and  $D_p$ , the above equation simplifies as

$$R_{TCPC} = N \frac{MSS}{RTT} \frac{C}{\sqrt{p}}. \quad (8)$$

Then, substituting the expression of the loss rate given in Eq. 3 in Eq. 7 above, leads to

$$R_{TCPC} = N * \frac{MSS}{RTT} \frac{C}{\sqrt{\phi_{out} - \phi_{out} \frac{CIR}{R_{TCPC}}}} \quad (9)$$

<sup>3</sup>This could be extended to the more reliable equation  $R_{TCPC} = \frac{MTU}{RTT \sqrt{\frac{2p}{3} + T_0 \sqrt{\frac{2T_p}{3} p(1+32p^2)}}}$ .

which in turn, by solving this quadratic equation, leads to

$$R_{TCPC} = \frac{CIR}{2} + \sqrt{\left(\frac{CIR}{2}\right)^2 + \left(\frac{N * C * MSS}{RTT * \sqrt{\phi_{out}}}\right)^2}. \quad (10)$$

This expression turns out to be similar to the throughput provided by the TBF analytical model above, i.e., to

$$R_{TCPC} = \frac{1}{2}R_A + \sqrt{\left(\frac{R_A}{2}\right)^2 + \left(\frac{N \times C \times MSS}{\sqrt{\phi_{out}} \times RTT}\right)^2}, \quad (11)$$

when  $CIR = R_A$ , i.e., when the rate assured by TBF is the rate negotiated in the contract. From this study we conclude that the above TBF+RED rate control equation can in turn be exploited to control the rate of the transmitted streams in order to maintain the TBF+RED mechanism into the flow phase (at the limit of the saturation phase). This should lead to an optimal bandwidth usage within the limits of the SLA.

## 6 Source rate control

Let us now consider the overall rate control architecture depicted in Fig. 3, and explicit the rate control algorithm coupled with a marking/shaping on the server side. The traffic shaping and marking functions on the sender side are also based on a TBF mechanism governed by the same equations as above (Eq.4). When the condition (6) is verified, the assured rate  $R_A$  increases with the token bucket depth  $L$ , up to the saturation point. Reaching the saturation phase the AF channel may start to drop packets. The goal of the rate control algorithm is to exploit this threshold value  $L_{max}$  to maintain the channel in the flow phase, i.e. to avoid any losses for the highest priority information ( $p_{in} = 0$ ). In order to do so, it will adapt the rate of the information of the highest priority (base layer in the case of layered coding), up to the limit of saturation of TBF+RED. Beyond this point, the rate control algorithm falls back to a classical CBR rate control mechanism such as the TMN8 algorithm [3]. If the encoding rate becomes lower than the rate corresponding to  $L_{max}$  the rate control decreases the sender TBF depth  $L$  to the proper value corresponding to the current frame bit budget. The rate control algorithm for one layer video proceeds as follows:

1. Initialize RED parameters :  $R_A =$  SLA committed rate (CIR), maximum burst size. Derive the TBF depth parameter  $L$  by solving the conditional equation 6.
2. Apply the marking function to the base layer by choosing the class or priority and the corresponding DSCP points.
3. Calculate the loss rate  $P$  from the RED dropping probability and out-of-profile marking probability.
4. From  $P$  estimate the TCP-compatible rate prediction  $R_{TCPC}$ .
5. If the source coding rate  $R_S$  of the high priority information is above the predicted TCP-compatible rate  $R_{TCPC}$  given by Eq. 11 ( $R_S > R_{TCPC}$ ), increase the TBF depth  $L$  as  $L = L + \Delta L$ .
6. Else if  $R_S \leq R_{TCPC}$ , adapt  $L$  by solving Eq. (6).

7. Compare  $L$  with  $L_{max}$ , the limit of saturation (where the TBF depth has no more effect on the maximum achieved rate). If  $L \geq L_{max}$ , no more increase for current depth, and the rate control falls back to a CBR constraint exploited by the TMN8 algorithm.
8. Compute the source distortion for the resulting rate  $R_{TBF}$ . If the distortion is above a given threshold  $T$ , then the source is encoded at a constant average rate with the TMN8 rate control; Else the encoder follows the current  $R_{TBF}$  boundaries.
9. Next Frame  $\mapsto$  go to(2).

## 7 Experimentation Results

The first set of experiments aimed at evaluating the gain of the AF adapted rate-control for non-layered H.263+ in a well-provisioned network, with both low and high rates. Fig. 7 depicts its PSNR performance against the TMN8 rate control algorithm, for the low rate experiment. The two testing sequences are News and Foreman (300 frames, 25 frames/s). The source coding rate, when using the TMN8 rate control is chosen to be 180 Kbps in average. The MRED AF1 channel parameters, i.e. dropping probability and minimum and maximum thresholds, have been set respectively to  $D_p = 0.01$ ,  $Thresh_{min} = 12$  KBytes,  $Thresh_{max} = 36$  KBytes. The TBF assured rate  $R_A$  is set to 240Kps. This choice has been led by the fact that in the experiments, losses have been observed if the sending rate was exceeding 75% of the assured rate. This 75% value should allow to guarantee that the probability of dropping in-profile packets will be  $P_{in} = 0$ . The round trip time RTT was emulated to be around 120ms. Competing UDP sources are generated within AF2, AF3, AF4 classes.

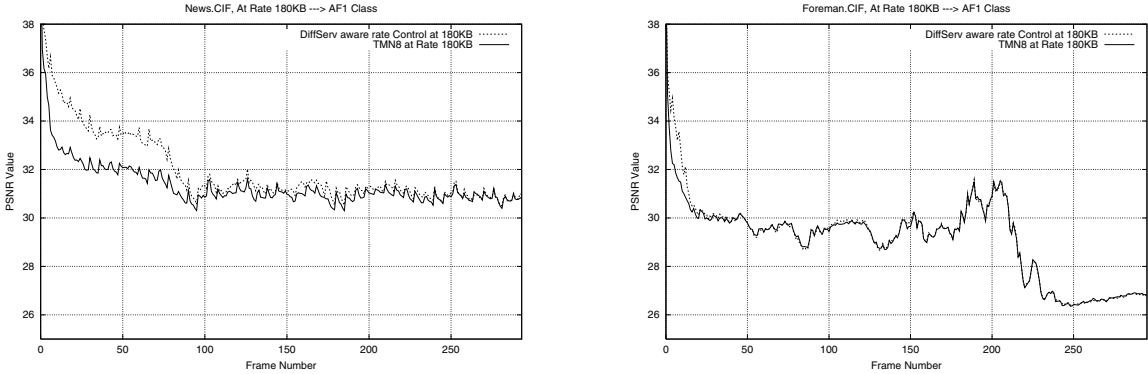


Figure 7: PSNR performance of the AF adapted rate control algorithm

A good isolation between classes has been observed during the experiments. On the curves of Fig. 7 one can observe the flow phase of the TBF algorithm (the first part of the PSNR curve). During this phase, the sender benefits from the available bandwidth by increasing its sending rate to a value which will correspond to the maximum TBF depth  $L_{max}$ , before starting to observe losses. After instant 95s, MRED enters a saturation phase, the increase in the TBF buffer depth  $L$  would not have effect on the source rate any more. TBF has reached its maximum depth. In this case the algorithm falls back into the TMN8 mode with a constant average rate of 180Kbps.

Fig. 8 shows the PSNR obtained with higher rates i.e., 256Kbps for News and 300Kbps for Foreman, with a TBF assured rate set respectively to 340Kbps and 400Kbps. We use the same MRED channel parameters as above. The RTT value is again set to 120ms. The higher degree

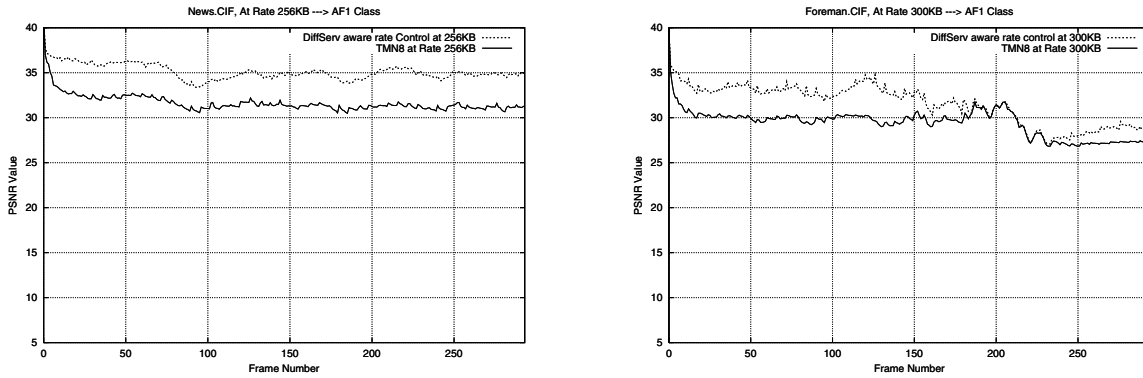


Figure 8: PSNR performance of the AF adapted rate control algorithm

of freedom in terms of bit budget allows to better adapt to the channel conditions, without restrictions induced by the TMN8 buffer occupancy. Increased rate control stability can then be achieved at these higher rates.

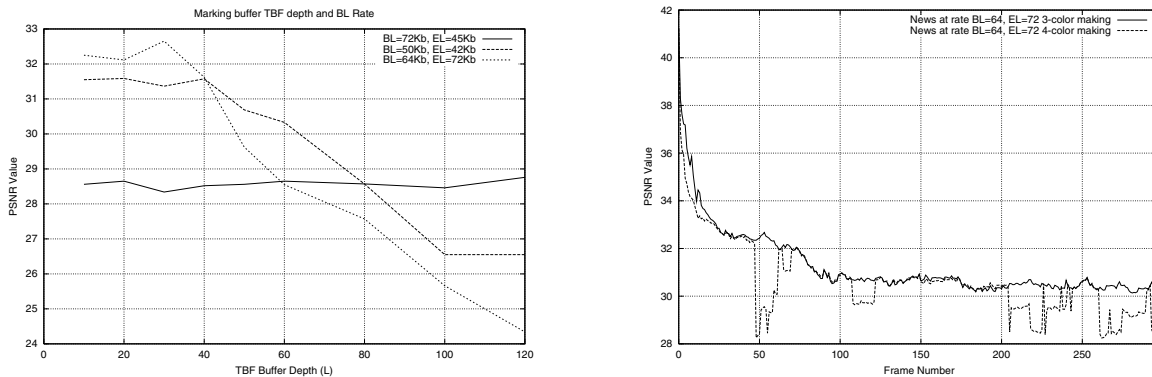


Figure 9: a-Layered Video performance and marking buffer settings. b-PSNR performance with three and four color marking.

The second set of experiments aimed at understanding the impact of the rate control applied to the base layer on the dynamics of layered video traffic. We consider here a two-layer H263+ SNR scalable compliant video encoder. Fig. 9-a shows the overall PSNR performance in terms of the depth  $L$  of the TBF mechanism applied on the base layer and for different repartition of overall rate between base and enhancement layers. It has been observed that first for a value of  $L$  above a maximum value (40 Kbytes on the curve) no shaping is applied any more on the base layer which leads to severe burstiness on the base layer hence to PSNR performance degradations. At the same time, for appropriate values of TBF depth, some bandwidth shares between base and enhancement layers could still lead to bursty base or enhancement streams in turn leading to poor overall performances due to losses. This can be seen on the curve corresponding to the share 72 Kbits/s for the base layer and 45 Kbits/s for the enhancement layer. This emphasizes the need to define a proper setting and mapping for both layering and channel parameters.

Tests have been also carried out to study the effect of the number of marking levels on the quality of the received video stream. The tests have revealed stability for three levels AF1x of marking with simple yet sufficient static mapping between I, P, and B data and the three DSCP values. In the case of three color marking [10], AF11, AF12 and AF13 are assigned respectively

to I, P and B data of the base layer. The enhancement layer is assigned to the AF2 class. In the case of four color marking, the third and the fourth colors are used to differentiate packets containing B data with a different amount of Intra macro-blocks. Fig.9-b shows the PSNR of the received video for the sequence News with respective base and enhancement layers coding rates of 54kbps and 72kbps with the three and four color marking strategies. Higher stability is obtained with the three color marking, due to excessive losses on the fourth color or lowest priority data.

## 8 Conclusion

Tuning the parameters of AF PHB supporting mechanisms, when applied on aggregated flows, in order to satisfy appropriate service differentiation and QoS requirements of bursty video sources is not an easy task. Considering a TBF+RED based implementation of an AF channel, we have first evidenced ranges of values for key parameters of TBF and RED in order to maintain the overall mechanism into a flow phase, avoiding saturation and losses for information of high priority. Given the corresponding setting, an analytical model of the throughput of the AF channel has been validated and exploited into a video source rate control. By tailoring the shaping of outgoing video streams to AF channel throughput, smoother traffic can indeed be obtained, hence maximizing the bandwidth usage within the limit of the negotiated SLA. Applying the rate control on a base layer of a two-layered video revealed the impact of the bandwidth share between base and enhancement layers onto the burstiness of the resulting streams, hence on loss rates and overall performance degradations.

## References

- [1] V. Jacobson, K. Nichols Cisco systems, and K. Poduri Bay Networks, "An expedited forwarding phb," *IETF RFC2598*, June 1999.
- [2] J. Hienanen, F. Baker, W. Wiess, and J. Wroclawski, "Assured forwarding phb group," *IETF RFC2597*, June 1999.
- [3] J. Ribas-Corbera and S. Lei, "Rate Control for Low-Delay Video Communications," Tech. Rep., ITU-Telecommunication Standardization Sector, Portland, June 1997, ITU-T Study Group 16, Video Expert Group.
- [4] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM transactions on Networking*, vol. 1, no. 4, pp. 397–413, August 1993.
- [5] E.Hashem, "Analysis of random drop for gateway congestion control," Tech. Rep., Report LCS TR-465, Laboratory for computer science,MIT, Cambridge,MA, University Press, 1989.
- [6] A. Mankin, "Random drop congestion control," *ACM SIGCOMM*, pp. 1–7, September 1990.
- [7] W. Almesberger, J. Salim, and A. Kuznetsov, "Differentiated services on linux," Tech. Rep., EPFL ICA, CTL Nortel Networks, INR Moscow, June 1999.
- [8] D. Lin and R. Morris Harvard, "Dynamics of random early detection," *Proc.SIGCOMM'97*, September 1997.
- [9] S.Sahu and P.Nain, "On achievable differentiation with token bucket marking for tcp," Tech. Rep. 77, UMASS CMPSCI, 1999.
- [10] O.Medina, J.Bonnin, and L.Toutain, "A Multimedia Color Marker," *IETF*, <http://www.rennes.enst-bretagne.fr/medina/docs/draft-medina-mcmm-00.txt>, April 2000.