A Congestion Control Scheme for Multimedia Traffic in Packet Switching 'Best-Effort' Networks

R. Bolla, M. Marchese, S. Zappatore

Department of Communications, Computer and Systems Science (DIST)
University of Genoa
Via Opera Pia 13, I-16145 Genoa, Italy
{lelus, dama, zap} @ dist.unige.it

Abstract. In this paper a congestion control scheme for multimedia traffic in 'best-effort' networks is presented. The proposed approach is based on the dynamic modification of transmission characteristics, such as average bit rate, frame rate, spatial resolution, frame quality and coding scheme and on the concept of Perceived-Quality of Service (P-QoS) measured by using the Mean Opinion Score (MOS) technique, which is the degree of perceptual relevance. The global control system is organized into two blocks: the "coding agent" and the "congestion controller". This latter detects a congestion situation and decides the bit rate to use for the transmission. The choice is based on some feed-back information from the receiver and is aimed at preventing the future congestion. The "coding agent" chooses the transmission characteristics corresponding to the highest MOS among the ones offering the selected bit/rate. The effectiveness of the strategy has been analyzed in an experimental testbed using TCP/IP suite over an Ethernet LAN and the related results are presented.

1. Introduction

Recently, there has been an increasing interest in multimedia services. Many applications have been developed both in local (electronic documentation and manuals, games, CDROM including dictionaries, cultural exhibitions, ...) and distributed environments (cable TV, Video on Demand (VoD), distance-learning,...). Concerning Networked Multimedia Applications (NMA), which are the object of this work, they are characterized by the presence of a mix of different traffic streams, some of which (e.g., video) can be very bandwidth-demanding. This may not be a problem when a high speed network is available, especially if this latter can grant Quality of Service (QoS), as, for instance, the future B-ISDN. Anyway, due to the fast development of 'best-effort' networks (e.g. Internet) in recent years, the study and the analysis of various techniques (coding, image compression or control schemes, for example) for the transmission of multimedia streams (through networks unable to guarantee a minimum bandwidth) has become more and more important. As far as a LAN environment is concerned, a possible bandwidth limitation (generating a bottleneck) could stem from the use of a network medium (e.g. Ethernet) characterized by low peak bit/rates, compared with a FDDI or ATM LAN, and shared among many users. Furthermore, a more serious problem may be the interconnecting links among LANs. Anyway, at a LAN level, such shared environments are, nowadays, so widely deployed that they are likely to remain in use for some time to come, before being totally replaced by the high capacity ATM LAN switches (which, by the way, also do not provide unlimited resources either). Moreover, also in an ATM environment, a possible use of Available Bit Rate (ABR) service class for multimedia transmission could require the application of similar techniques [1].

In this context, some relevant topics are access and flow control, asynchronous adaptive coding, inter-media and intra-media synchronization and traffic modelling. The main problems in multimedia applications for 'best-effort' networks are the extreme difficulty to devise good source models and the congestion probability due to the difficulty of controlling users entering the network. Therefore, the main aim of control algorithms in a 'best-effort' environment is not to guarantee QoS requirements but to avoid congestion situations [2]. Even the concept of QoS needs more attention than in 'classical' packet networks, where QoS is defined through objective quantities as, for instance, the rate of lost or delayed packets. Since multimedia transmission heavily involves client requirements [3] and human factors [4], it is not simple to define objective metrics, and the introduction of appropriate subjective metrics as Mean Opinion Score (MOS) is strictly necessary [5, 6]. An interesting approach to these topics can be found in [7], where a clear introduction to the transmission of video over "best-effort" packet switching networks is presented.

In this paper, we remind to introduce the Perceived - Quality of Service (P-QoS) general notion, we introduce the concept of "configuration" and the consequent need of statistical tests to evaluate it. The main part of the paper is, then, dedicated to present a feedback control scheme, composed by various operating blocks. A completely decentralized control is considered in this approach. The controlled objects are end-to-end audio-video communication applications which apply a bit-rate control scheme, based on a congestion information "local" to each transmitter-receiver communication. In this formulation, any transmitter-receiver couple acts in autonomous way. The considered protocol environment, as explained in Section 4, is the TCP/IP suite, and the control mechanism is located just above the transport layer (UDP, in this case). Even if the same concept could be partially applied to audio data, the control mechanism has been tested just on video traffic.

The paper is organized as follows. Section 2 describes the overall proposed control scheme, while the definition of the P-QoS is dealt in Section 3. The network congestion control algorithm is explained in Section 4. Some experimental results are discussed in Section 5.

2. The proposed scheme

The overall communication scheme is depicted in Fig. 1. It is divided into different operating blocks, of which the topical ones are the "coding agent" and the "congestion controller".

The "congestion controller" detects a congestion situation and decides the bit rate to use for the transmission. The choice is based on some feed-back information received from the receiver and is aimed at avoiding or reducing the future congestion. The feed-back information concern the packet loss rate, which is the ratio between the useless (lost or delayed) and sent packets, and the jitter. The bit rate B_t is stated every T seconds by the transmitter and this value is maintained for the whole interval. The selected bit-rate is communicated to the "coding agent": this block chooses an audio-video coder with transmission characteristics suitable for the current network load.

Concerning the coder, it has received a great deal of attention in the literature [8] and it is not the object of this paper. The characteristics of the coder used in this approach are mentioned in Section 5 and they should be considered as an example.

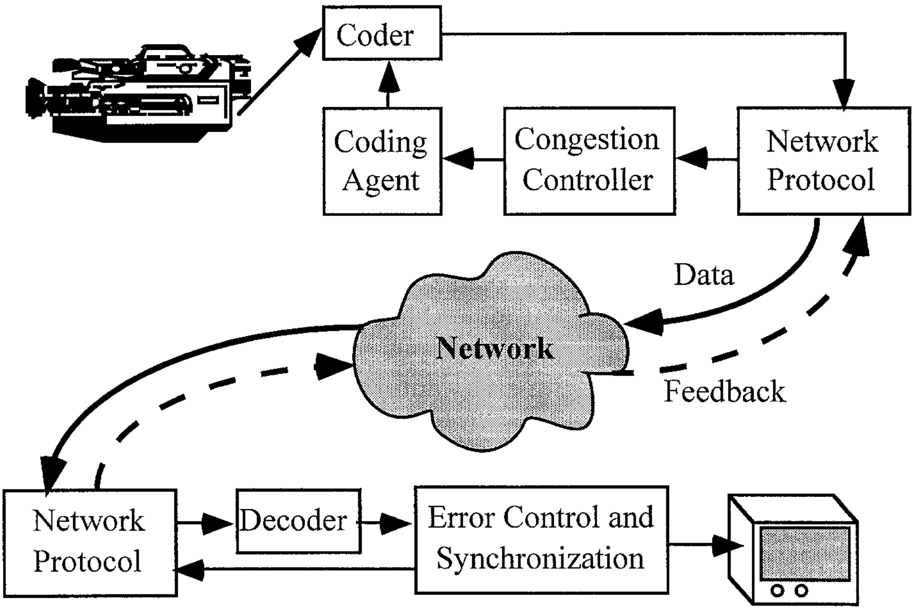

Fig. 1. Overall communication scheme.

3. The P-QoS concept

As already stated, the evaluation of the quality is necessarily subjective in a multimedia service, i.e. the quality strictly depends on the user perception: this is the concept of Perceived Quality of Service (P-QoS).

Different methods of measuring the P-QoS have been considered, for voice transmission [9], for the video signal [4] and for a telecommunication system in general [3, 5]. All the methods mentioned above employ the Mean Opinion Score (MOS) as a quality measure. The estimation of MOS is out of the scope of this paper, but it can be summarized as follows: let us consider a particular video application, a set of representative test sequences and a set of different video transmission characteristics such as coder type, time and spatial resolution and colour depth. A meaningful sample of a population is chosen and asked an opinion (a score) about any mode of the typical utilization conditions for the considered application. Finally, the mean (called Mean Opinion Score) of all reported scores is computed for any mode of transmission. The MOS can be associated to any transmission system and it is a good measure of the real QoS perceived by the user. Fig. 2 shows a possible scale for the MOS.

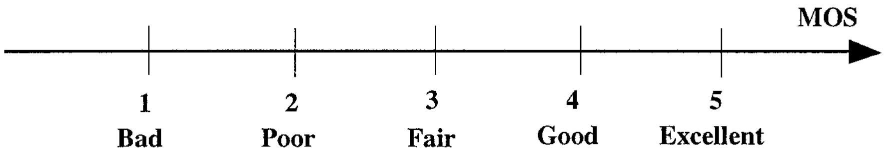

Fig. 2. Scale for the MOS

A deep statistical analysis is needed to get a reliable measure. The MOS values appearing in the rest of the paper have poor statistical relevance; they have to be considered just as an example. In a packet switching video transmission, the objective parameters that directly influence the MOS are summarized in Fig. 3.

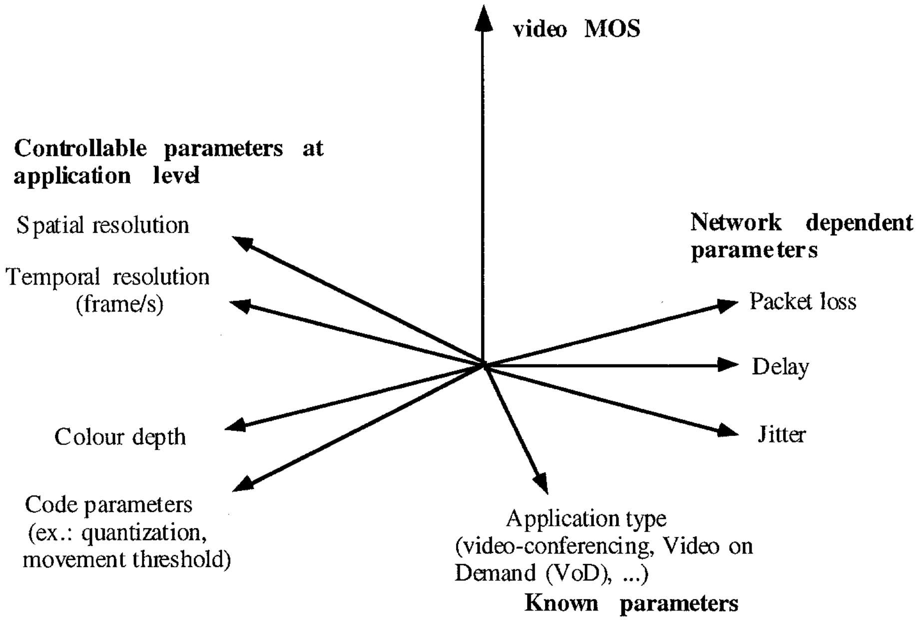

Fig. 3. Parameters influencing the video MOS.

The set of controllable parameters (spatial resolution, temporal resolution, compression mode, ...) along with the associated MOS is called a "configuration". The configuration table can be computed off-line and used by the application.

Due to the very large number of possible configurations, it is not trivial to compute the MOS value for each of them. Since the exhaustive search is not feasible, other methods have to be designed. In this paper a MOS estimation has been performed by using a linear interpolating function for each type of coder. This subject is currently object of research by the authors. On this base it is simple to define the "coding agent" block of Fig. 1. The "coding agent" chooses the best suited configuration, i.e. the configuration with the highest MOS among the feasible ones. The set of feasible configurations depends on the congestion status of the network, estimated by the "congestion control" block described in the next Section. The "congestion control" block estimates, every T seconds, a suited bit rate Bt and the "coding agent" chooses the configuration with highest MOS among the ones which have an outgoing bit rate in the neighbourhood of the Bt. The chosen bit rate is maintained fixed for other T seconds, at least.

4. The congestion control mechanism

This Section will explain the functionality of the "congestion control" block, that is how this block detects the congestion and how it selects the bit rate entering the network. The "congestion control" scheme can be described by using the transition state diagram of Fig. 4, which is composed by the following states.

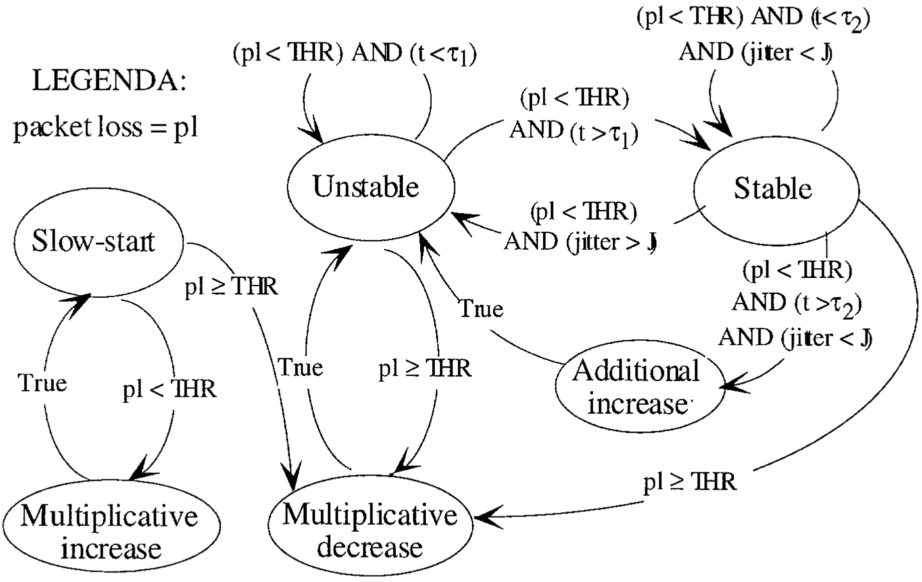

Fig. 4. Congestion control state diagram

4.1 Slow-start.

It is the initial state. The application starts the session by generating the minimum stream of data. The network status is unknown and a high initial rate could produce a resource waste and a high packet loss. The packet loss is periodically measured (at the instants $t_i = i$, 1=1, 2, ...) and it is one of the quantities used to detect the congestion. If the loss is smaller than a fixed threshold (THR), the bit rate is increased according to equation (1), (Multiplicative increase): if B_{t_i} represents the average output bit-rate in the interval $[t_i, t_{i+1})$, the increase takes the form

$$B_{t_i} = k \cdot B_{t_{i-1}}; \quad i = 1, 2, 3, \ldots; k \in \Re; k > 1 \tag{1}$$

where k is a constant heuristically computed as well as the value THR. Otherwise, if the packet loss is bigger than the threshold, the slow-start phase is concluded and the chosen state is the "Multiplicative decrease" one.

4.2 Multiplicative decrease.

The bit rate value is decreased by the following rule:

$$B_{t_i} = d \cdot B_{t_{i-1}}; \quad i = 1, 2, 3, \ldots ; d \in \Re; d > 1 \tag{2}$$

where d is a constant whose value is chosen as in the previous case. The next state is the "Unstable state".

4.3 Unstable state.

At each control instant, if the loss is larger than the threshold, the new state is the "Multiplicative decrease", otherwise the "Unstable state" is maintained until the time spent in that state is larger than τ_1 (a constant, large enough to obtain a feasible measure and to avoid that point measures cause the control mechanism intervention), then the next one is the "Stable state".

4.4 Stable state.

In this state the actual detection of the congestion takes place. Two quantities are used here: the value of the packet loss and the value of the jitter, this last defined as the mean standard deviation of the packet trip time and estimated according to [10]. The transmitter, by using some periodical information about the network jitter estimation and the loss value, decides if the bit rate has to be increased or decreased. In fact, if the loss is smaller than the threshold, if the time is long enough ($t > \tau_2$) to get feasible measures and if the jitter is below a threshold (J) the bit rate is incremented ("Additional increase" state), else, if time is not long enough, the state does not change. If the loss is larger than the threshold the bit rate is decreased ("Multiplicative decrease"), else, if the loss is smaller than the threshold but the jitter is larger than J, the next state is the unstable one.

4.5 Additional increase state.

The bit rate, as mentioned above, is increased. The chosen increase is the additional one, as in (3), to avoid an overestimation of the available bandwidth.

$$B_{t_i} = B_{t_{i-1}} + u; \quad i = 1, 2, 3, \ldots ; u \in \Re \tag{3}$$

where, also in this case, u is chosen by experimenting. The values of the heuristic constants used in the performed tests are specified in the next Section. As it can be seen, the general philosophy of the control system is similar to the TCP one [10].

5. Experimentation and results

A simple video transmission testbed, based on an Apple Macintosh platform, has been implemented to verify the efficiency of the proposed algorithm. The application

QuickTime has been used to acquire and code the video stream. The used configuration table is shown in Tab. 1. The "non-Pareto optimal" configurations have been eliminated. The used protocol stack is the TCP/IP suite with MacTCP driver.

Average bit rate	MOS	Codec	Bit per Pixel	Frame rate	Frame resolution	Frame quality
11	1.5	jpeg	8	1	1	1
13	1.7	jpeg	24	1	1	1
18	1.8	jpeg	24	1	2	1
25	1.9	jpeg	24	1	3	1
30	2.0	jpeg	8	2	2	1
45	2.2	jpeg	8	3	2	1
60	2.4	jpeg	8	4	2	1
175	2.5	jpeg	8	5	1	2
582	2.6	AppleVideo	24	4	2	2
675	2.7	AppleVideo	24	3	2	3
960	2.8	AppleVideo	24	3	3	3
1070	2.9	PICT	8	7	1	-
1200	3.0	PICT	8	5	2	-
1384	3.1	PICT	8	4	3	-
1440	3.2	PICT	8	6	2	-
1730	3.3	PICT	8	5	3	-

Tab. 1. Configuration table.

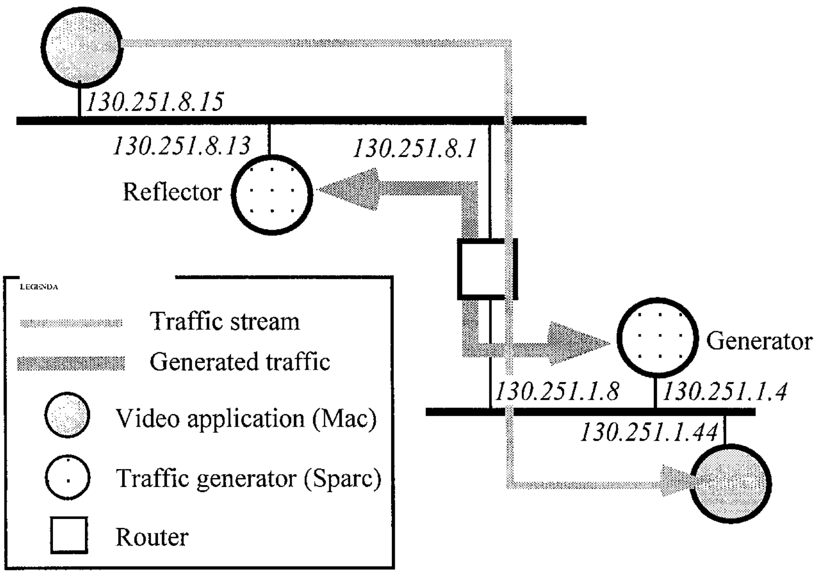

Fig. 5. Testing environment

The experimentation has been performed in a remote environment by using two LANs interconnected by a router. In order to evaluate the proposed control algorithm, a Sun SparcStation 10 has been used to generate different traffic loads, thus disturbing the multimedia packet exchange. The overall testbed is illustrated in Fig. 5.

Some remarks about the traffic generator application are needed to understand the results reported below. Actually, the application is composed of two parts: the "generator" and the "reflector", that are performed on two different computers, as in Fig. 5. The "generator" sends a time-stamped packet stream to the "reflector". The "reflector" receives the packets and sends them back to the "generator. The "generator" receives back the packets and computes the Round Trip Time (RTT - the time between the transmission and the reception) and the jitter.

Each test has been performed starting from a stable situation, bypassing the slow-start phase. The reported values are the results of an average taken over 1 s intervals.

The values of the constants, mentioned in the previous Section, used to obtain the results in the following are reported in Tab. 2.

T	Bit-rate temporal interval	7 s
τ_1	Unstable state temporal constant	20 s
τ_2	Stable state temporal constant	20 s
THR	Packet loss rate threshold	0.1 (10%)
k	Multiplicative increase constant	2
d	Multiplicative decrease constant	1/2
u	Additional increase constant	2 Kb/s

Tab. 2. Constant values

The jitter threshold J is not fixed, being the double of the jitter value when the transition into the stable state is accepted. It means that, if the jitter estimation is larger than twice the jitter value when the transition was accepted, the system goes again into the unstable state.

Figs. 6, 7, and 8 are related to a situation called unloaded in the following: the results have been obtained with few users using the network and with no video applications, by generating a step variable load. This latter is originated by the generator, which measures the traffic as well. In Fig. 6, the traffic load is shown, whereas the jitter and the RTT are shown in Fig. 7 and Fig. 8, respectively.

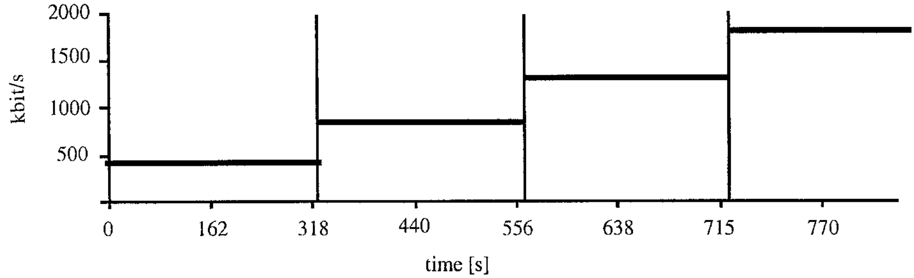

Fig. 6. Traffic load originated by the generator in the unloaded situation.

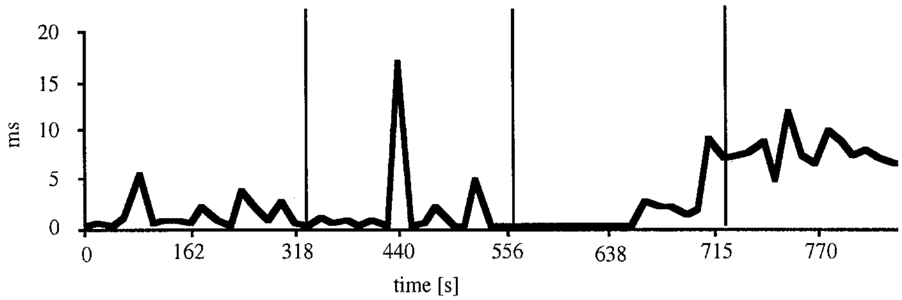

Fig. 7. Jitter in the unloaded situation.

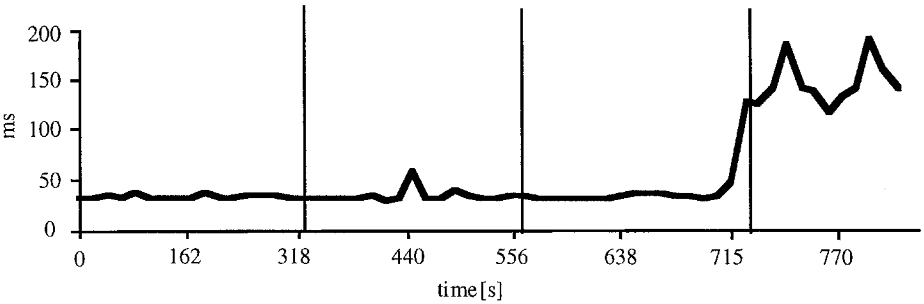

Fig. 8. RTT in the unloaded situation.

The results presented below have been obtained by starting the video application with and without the congestion control mechanism active.

In Fig. 9 is shown the traffic load originated by the generator in the following test situations. In Fig. 10 and Fig. 11 are shown the jitter values measured by the traffic generator with the congestion control activated (Fig. 10) and with no control activation (Fig. 11); whereas the RTT with congestion control and no congestion control are shown in Fig. 12 and Fig. 13, respectively, in the same situation of the previous case.

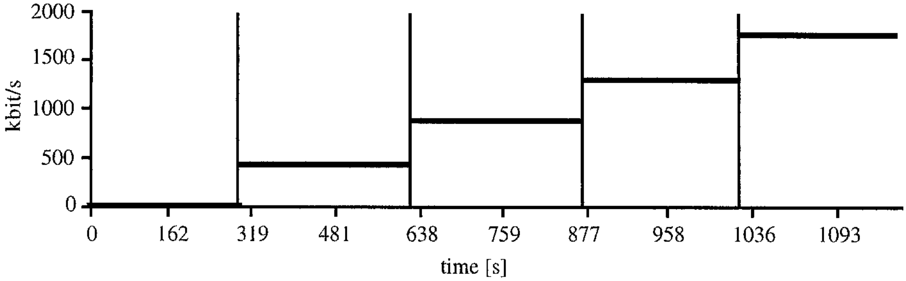

Fig. 9. Traffic load originated by the generator in the test situations with video application.

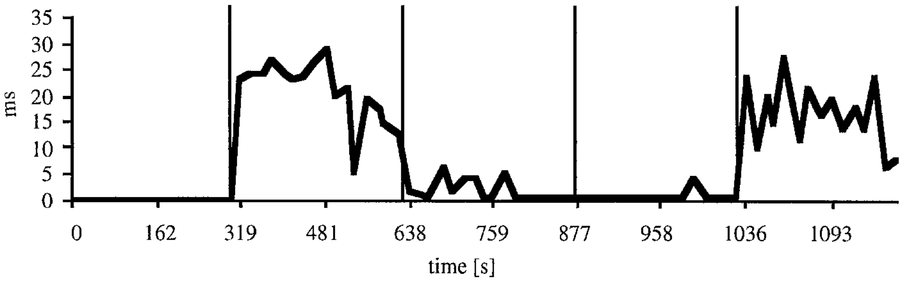

Fig. 10. Jitter measured by the traffic generator with video application and congestion control.

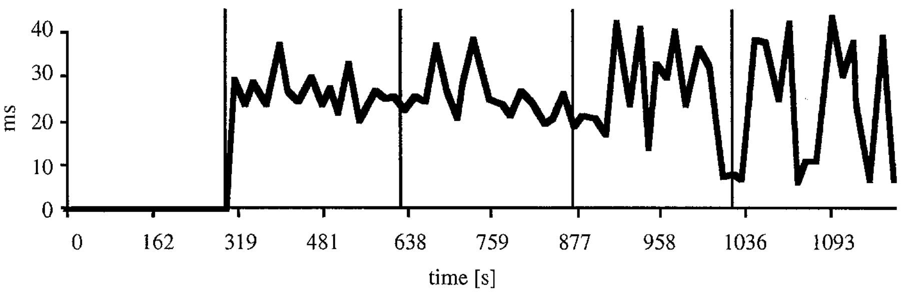

Fig. 10. Jitter measured by the traffic generator with video application and no congestion control.

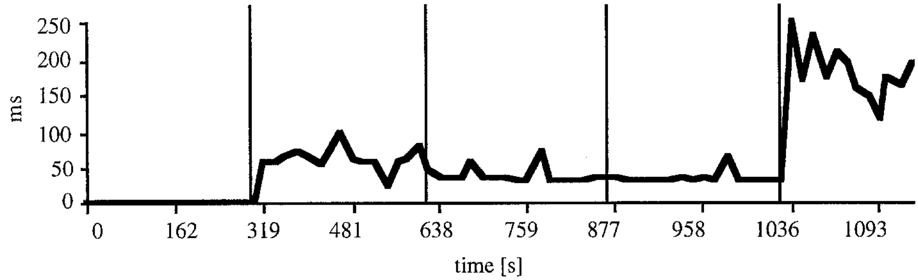

Fig. 12. RTT measured by the traffic generator with video application and congestion control.

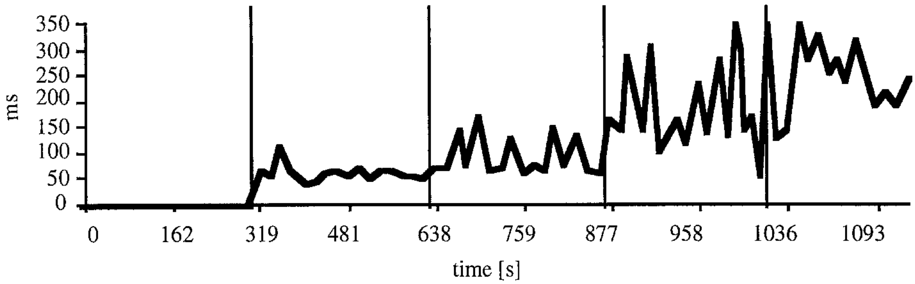

Fig. 13. RTT measured by the traffic generator with video application and no congestion control.

In Fig. 14 and Fig. 15 is shown the bit rate measured by the video application receiver in case of congestion control activated and not activated, respectively. The jitter is depicted in both cases in Fig. 16 (congestion control) and Fig. 17 (no congestion control).

The packet loss measured by the video application is shown in Fig. 18, when the control in active, whereas the same quantity with no congestion control is reported in Fig. 19.

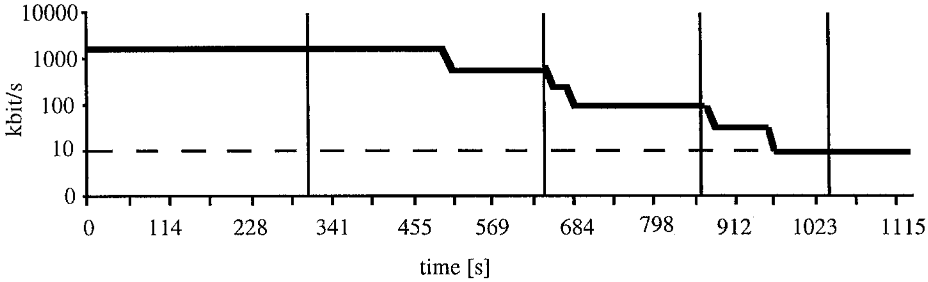

Fig. 14. Output bit rate generated by the video application (with congestion control).

It can be noted the effectiveness of the congestion control both concerning the jitter (Fig. 15, Fig. 16) and the packet loss (Fig. 17, Fig. 18). The values obtained when the control is active are sensibly smaller than those obtained when no control is performed. It is worth noting that, even if the action of the control maintains the bit rate measured by the video application (Fig. 14) smaller than in the case with no control (Fig. 15), the quality perceived by the user is higher, due to the relatively small values of the jitter and the packet loss. In fact, high values of these two quantities mean a very annoying vision from the user point of view.

The control is not only effective for the video application, but it also allows an improvement for the other applications in the network, thus reducing the general network congestion.

The comparison between the case of active control and of no control for the jitter (Figs. 10, 11) and the RTT value (Figs. 12, 13) shows the reduction of the congestion in the network. The value obtained when the control is active is smaller and the difference between the two cases is really outstanding.

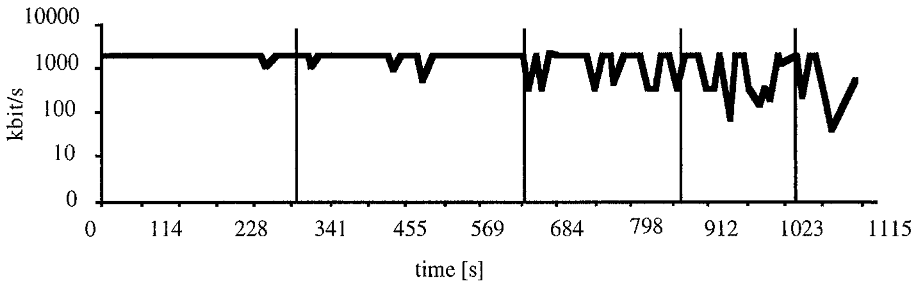

Fig. 15. Bit rate measured by the video application (without congestion control).

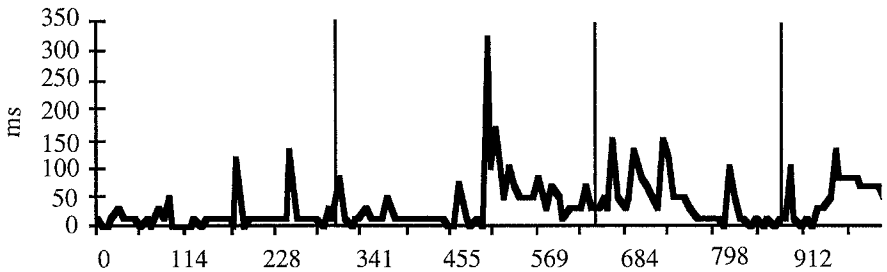

Fig. 16. Jitter measured by the video application (with congestion control).

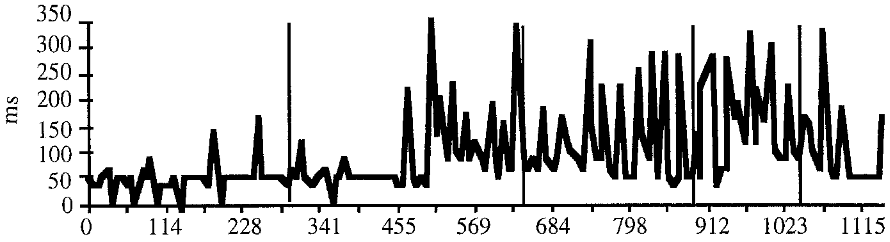

Fig. 17. Jitter measured by the video application (without congestion control).

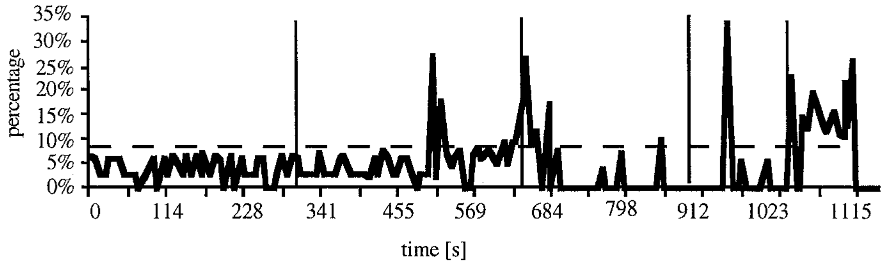

Fig. 18. Packet loss measured by the video application (with congestion control).

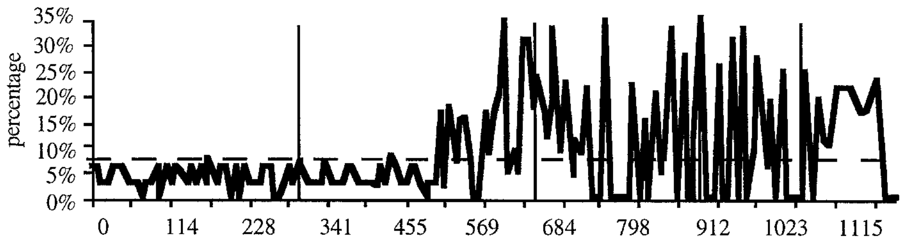

Fig. 19. Packet loss measured by the video application (without congestion control).

6. Conclusions

A congestion control scheme for multimedia traffic in 'best-effort' networks has been presented. After an introduction about the definition of the working scenario and of the general control approach, the general control scheme has been introduced along with the concept of P-QoS. Then a network congestion control scheme organized into operating blocks and based on some feedback information of the network status has been presented.

The algorithm has been analyzed and experimentally tested: some of the obtained results prove the effectiveness of the proposed approach. These results show in particular that the mechanism is not only useful to improve the quality perceived by the user but also to prevent network congestion for any other network application.

7. References

[1] R. Jain et al., "Source behaviour for ATM ABR traffic management: an explanation", *IEEE Communications Magazine*, Vol. 34, No. 11, pp. 50-55, November 1996.

[2] R.Bolla, F.Davoli, M.Marchese, "Quality of Service Management and Control of Multimedia Applications: a Scenario and Two Simple Strategies", *Proceedings of the 20th Local Computer Networks Conference (LCN'95)*., Minneapolis, Minnesota, October 1995.

[3] D.Ferrari, "Client requirements for real-time communication Services", *IEEE Communications Magazine*, Vol. 28, No. 11, pp. 65-72, Nov. 1990.

[4] ITU-T Recommendation F.901, *Usability evaluation of telecommunication services*, 1993.

[5] N.B.Seitz, S.Wolf, S.Voran, R.Bloomfield, "User-oriented measures of telecommunication quality", *IEEE Communications Magazine*, Vol. 32, No. 1, pp. 56-66, January 1994.

[6] N.Jayant, "High quality networking of audio-visual information", *IEEE Communications Magazine*, Vol. 31, No. 9, pp. 84-95, September 1993.

[7] M.Gilge, R.Gusella, "Motion video coding for packet-switching networks - An integrated approach", *Proc. SPIE conference on Visual Communications and Image Processing*, November 1991, Boston, MA.

[8] M.Rabbani, P.W.Jones, *Digital Image Compression Techniques*, SPIE Optical Engineering Press, Bellingham, Washington, USA, 1991.

[9] ITU-T Recommendation F.85, *A method for subjective performance assesment of the quality of speech voice output devices,* 1994.

[10] V.Jacobson, "Congestion avoidance and control", *Proc. ACM SIGCOMM'88*, Stanford, CA, August 1988.