Telecommunication Systems 11 (1999) 187–199

A simulation based approach for network resources dimensioning in video delivery systems

Roberto Canonico, Luigi D'Ardia and Giorgio Ventre

Gruppo di Ricerca sull'Informatica Distribuita, Dipartimento di Informatica e Sistemistica, Università di Napoli Federico II, Napoli, Italy
E-mail: {canonico, dardia, ventre}@grid.unina.it

The availability of high-performance networking technologies and processing power allows creation of new multimedia services and applications in broadband integrated services networks. Such new services will be successful only if delivered to end-users at low cost. At present, for wide-area distributed systems, communication has a relevant impact on the overall costs. As a consequence, adequate methodologies for an efficient dimensioning of the network resources need to be devised. Simulation techniques represent a powerful instrument for the analysis and evaluation of real systems. This paper, moving from the definition of mathematical models for VBR video sources, presents the application of a simulation approach to verify the compatibility of the communication load related to a set of video streams with the data transport capability of an existing network architecture for the delivery of video streams, in the case of CBR communication links.

1. Problem description

As high-performance networks and computing systems of increasing power are today available at low cost, great efforts are focussed on the realization of distributed media servers, to deliver on-demand multimedia information streams from high-capacity storage systems towards a number of client applications, which render the multimedia streams to the users. Realizing such media servers poses a number of challenging problems. The experience gained in a number of wide-area pilot experiments has demonstrated that the deployment on a commercial basis of the existing technologies to provide services like video-on-demand or teleconferencing is mainly hampered by the high cost of communication services. Due to the remarkable costs imposed by the creation of a broadband distribution infrastructure, it is critical to realize an efficient allocation and utilization of the communication resources. This goal can be achieved through a fair compromise between the delivered Quality of Service and a sufficiently high utilization of the available resources.

In this paper we address a typical problem encountered when a distribution network is to be dimensioned to support the transmission of a given mixture of video streams. In an ATM network, for example, this happens when multiple video sources, connected to an ATM switch, inject their traffic over a CBR Virtual Path Connection, with on demand activation of a Virtual Circuit Connection for each video stream [2].

In order to maintain the quality perceived by users at an acceptable level, it is necessary to feed the receiving applications with a continuous data stream. Occasional shortage of resources in any component of the distributed system (network bandwidth, buffering capacity in switching elements, processing power in the end-systems) may result in unacceptable degradation of the information stream presented to users by receiving applications. Additionally, as retransmission of lost data is not viable for time-sensitive multimedia streams, it is necessary to keep the loss ratio as low as possible within the network. This can be accomplished through controlled usage of resources in the system components. For this reason, a number of admission control algorithms have been presented in the literature for different kinds of resources, such as the storage subsystem, the network subsystem and the multimedia terminal end-system. The purpose of the admission control algorithms is to maintain the utilization of resources within the system at a level which is compatible with the Quality of Service expected by end-users. By submitting each service request to an acceptance/rejection decision, it is possible to guarantee an acceptable Quality of Service for each user.

Different approaches have been proposed to solve the admission control problem in multimedia distributed systems [1]. All of them either assume an a priori knowledge of the characteristics of the video sources, or they measure the relevant parameters of the traffic generated by each stream, to determine the admissibility of a new service request with the current state of resource utilization. Both these approaches perform statistical tests based on a simple description of the characteristics of the traffic produced by each source. These traffic characterizations, however, do not reflect the dependence of the traffic patterns from the encoding technique and the content of the multimedia streams. Admission control algorithms have been designed to be implemented in integrated services networks, so a short response time is preferred to a higher utilization of shared resources.

In this paper we present a content-based classification of encoded video sources and suggest a simulation-based approach that can be used to determine the compatibility of a predefined schedule of video sources for a specific video delivery system. A typical scenario where a simulation-based compatibility test might be preferred is the setup over public wide-area networks of long-duration connections, to support the delivery of multiple video streams, which differ in their content. The bursty nature of the traffic produced by encoded video sources suggests to dimension the system resources not just on the basis of the maximum number of video streams simultaneously active, but also taking into account the characteristics of the video streams. Our approach consists of accurately modelling the video sources and of simulating the behaviour of the multiplexer component of the video server. The results of our simulations may be used to dimension in the most efficient way the network resources, in order to deliver a given mix of video streams.

The rest of the paper is organized as follows. In section 2 the peculiarities of the traffic produced by encoded digital video sources are described. Section 3 illustrates a set of models for different kinds of encoded video sources. In section 4 a classification of typical video streams, according to their content, is illustrated. Section 5 presents an application of the simulation approach for a simple video streams multiplexer.

2. Video sources characterization

A video source produces a bit stream by encoding a sequence of digital still images (frames). The statistical characteristics of the bit stream depend on the adopted encoding technique [6]. Most of the currently adopted video coding algorithms produce variable bit rate streams, as the number of bits representing a part of a frame is not constant, depending both on the luminance and chrominance content of the frame and on the spatial and temporal correlation properties of the video stream. The size of the generated data units can be described by an analytical discrete model. For the analysis and the evaluation of distributed multimedia systems we need to model the video sources, in order to use these models in a network simulator. To limit the processing power required by the simulation environment, the chosen model has to be as simple as possible. Several models have been proposed for the correct simulation of encoded digital video sources. These models are usually based either on markovian chains or on autoregressive processes [4]. Sometimes, a combination of the two techniques is used. An analysis of a large set of video traces generated by different multimedia applications has shown a strong dependence of the traffic patterns from the application itself and from the video contents. In [4] a simple mathematical model for digital video sources has been presented, which can be efficiently implemented in a simulation environment. This model is based on an autoregressive process. The extreme versatility of autoregressive processes allows to easily create models for different kinds of video sources, as average and correlation functions are commonly available in simulation packages.

The compression techniques currently adopted in video coding algorithms can be of two types: *direct* and *differential*. Direct coding techniques encode independently each single video frame (spatial compression), while differential techniques encode the differences between adjacent frames (temporal compression). As a consequence, the traffic produced by a video source has a highly variable bit rate, which depends on the characteristics of single frames and on the correlation between adjacent frames. As an example, figure 1 shows two traffic traces. The trace on the left refers to a talk-show TV programme, while the trace on the right refers to a musical video clip. The two bit rates are normalised to their average value. As we can see from the two graphics, the two traffic traces present different characteristics, particularly in terms of bit rate variability.

A video stream can be divided into *scenes*, that is in groups of frames, homogeneous for visual characteristics. A *scene change* occurs when there is a deep alteration in the visual characteristics of the image. The sequence of scene changes obviously depends on the video stream content. Video streams produced by applications with slow dynamics, like videoconferences, are characterized by rare scene changes, while movies or sporting shots present frequent scene changes.

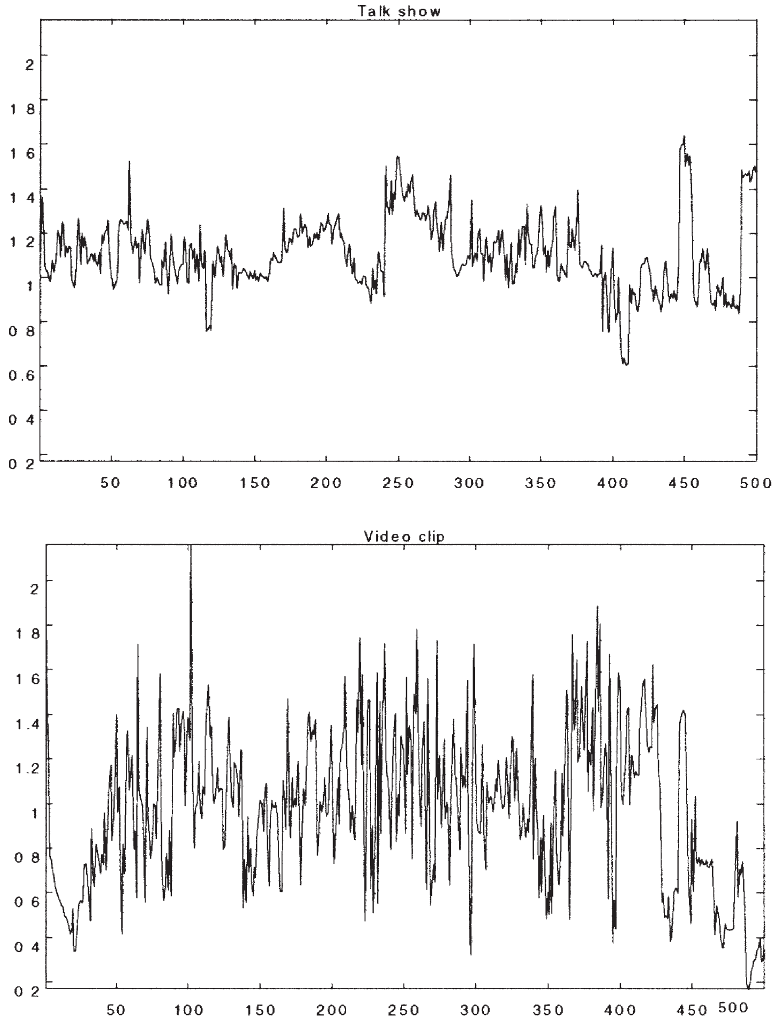
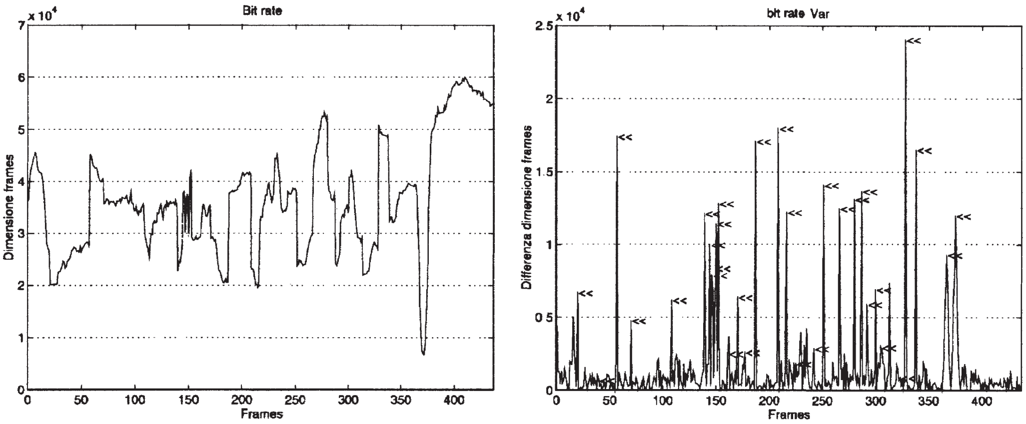


Figure 1. Example video traces (normalized bit rate vs. time in seconds).

When a compression algorithm is used, like, for example, the DCT (Discrete Cosine Transform) used in M-JPEG, a scene change produces a major variation in the bit rate, whose extent depends on the visual characteristics of the scene. The greater the difference between two successive scenes, in terms of brightness and wealth of particulars, the bigger will be the bit rate variation. An empirical method used to establish a scene change consists in estimating the difference in size between two adjacent frames. A scene change occurs when the difference exceeds a threshold value. Measures carried out on different video streams let us estimate such a threshold value as the 15

Figure 2(a) represents a trace produced by a commercial advertisement video stream. It is apparent how variable is the bit rate, so that it is not easy to discriminate



(a) (b)

Figure 2. Scene changes detection.

“real” scene changes. In figure 2(b) the difference in size between consecutive frames is plotted for the same stream. The “<<” symbols indicate a real scene change in the video. It can be noted that, if a sufficiently high threshold value is chosen, scene changes are not mis-detected.

On the basis of these observations, we defined a model for video streams in two subsequent steps. In the first step, we reproduced the bit stream related to a single scene. In the second step we defined the mechanisms for simulating scene changes. In the next section, these two steps are illustrated in more detail.

3. Models for digital video sources

A single scene can be well represented by a model for videoconference streams. Analysing the bit rate probability density function of a videoconference, it appears bell-shaped [10]. A videoconference bit stream can be modelled by a linear autoregressive procedure [10]:

$$X_n = a_0 + \sum_{r=1}^{p} a_r X_{n-r} + \varepsilon_n, \quad n > 0, \tag{1}$$

where $\{X_i\}$ is the sequence of the process outputs, $a_0, a_1, \ldots, a_n$ are real constants and $\{\varepsilon_i\}$ are normal random variables, independent with mean zero, unitary variance and independent from X values. Equation (1) describes the simplest linear autoregressive process, denoted with AR(p), where p is the order of the autoregressive process. To represent a videoconference stream, a second order autoregressive model AR(2) is used. The coefficients a_0, a_1, a_2 can be estimated when the autocorrelation of the source data stream is known.

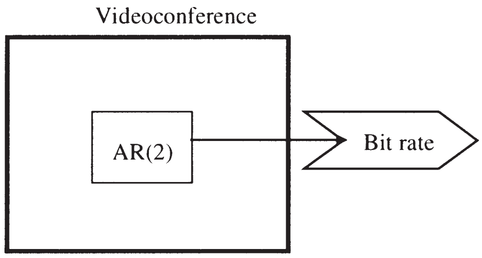


Figure 3. A model for videoconference streams (model 0).

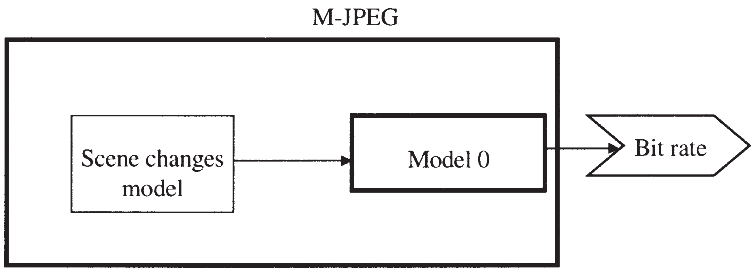


Figure 4. A model for M-JPEG video streams (model 1).

Dividing a video stream into subsequent scenes, the nth frame size is

$$S(n) = S'(n) + f(n), \tag{2}$$

where $S'(n)$ is the average frame size for the scene the nth frame belongs to, and $f(n)$ represents the variation of $S(n)$ from the average value. In a videoconference, no scene changes occur, so $S'(n)$ is a constant value, equal to the stream average frame size. By dynamically changing the $S'(n)$ value, it is possible to model any other direct coding, like M-JPEG. Figure 4 illustrates a model for M-JPEG video streams, where the "scene change" block modifies the average value of the autoregressive process, simulating in this way the scene changes.

The model presented in figure 4 is valid only in the case of direct coding. Differential encoding techniques, like MPEG, produce data streams with a dependance on the difference between adjacent frames. The MPEG video compression algorithm relies on two basic techniques: block- based motion compensation and DCT based compression for the reduction of the spatial redundancy. Three different types of frames are considered in MPEG: I, P and B. I frames are coded with a DCT algorithm, while P and B frames are coded with a motion compensation and interpolation algorithm. To define a model for MPEG video sources we started from an analysis of the statistical properties of several MPEG-encoded video streams. A number of studies on MPEG video streams evidenced the following interesting properties:

- P and B frames produce bit rates scene-independent and independent from the I frames bit rate.

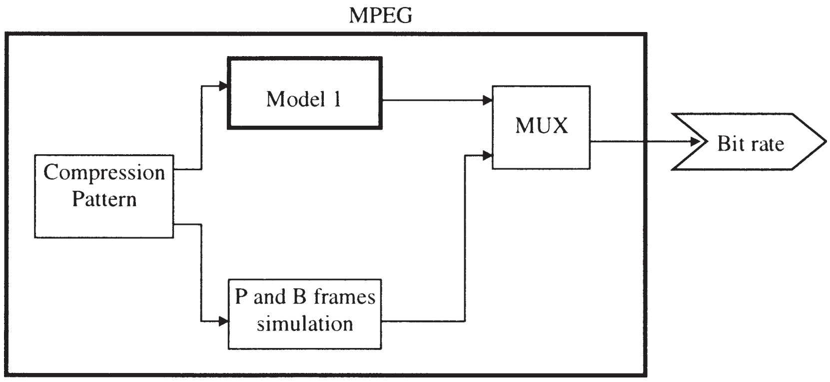


Figure 5. A model for MPEG video streams (model 2).

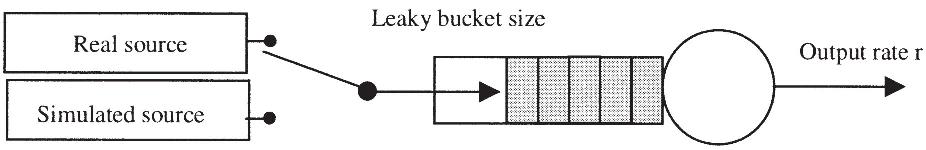


Figure 6. Validation by leaky bucket test.

Table 1
Video models structure.

Model 0	Videoconference	AR process
Model 1	M-JPEG	Model 0 + change scene process
Model 2	MPEG	Model 1 + P and B simulation process

- I frames and P-B frames bit rates are uncorrelated [7].
- P and B frames produce similar traces [12].

Normalizing the B and P frames bit rates to their average values, we obtain two sequences, characterized by similar values of the statistical parameters. On the basis of these observations, it is possible to affirm:

- The P and B frames are strictly connected and they must be treated in the same way.
- The P and B frames produce a sequence whose bit rate is independent from scene changes.
- The I frames are independent from P and B frames and depend on video scenes.

In figure 5 we present a block diagram for an MPEG video streams model, presented in [4], which is in accord with the properties mentioned above.

The MPEG model was implemented in a simulator, according to the layered structure of table 1.

In [4] a set of experiments are described, which validated the models for M-JPEG and MPEG video sources. These experiments compared the packet loss rate for a leaky bucket fed with the stream produced by a real source (M-JPEG or MPEG encoded) and by a simulated source.

4. MPEG sources classification

To better generalize the results of our analysis, we investigated possible classification criteria for video streams. For this reason we analysed a set of 21 different video traces, available via anonymous ftp. These traces refer to MPEG-encoded video streams, with the same encoding parameters.

For all of the streams, the (σ/μ) parameter was calculated for both the I frames stream and the P-B frames stream, where μ and σ are the bit rate average and variance. Representing on the axis of a cartesian plane the $(\sigma/\mu)_{\mathrm{I}}$ and $(\sigma/\mu)_{\mathrm{PB}}$ parameters, we get the diagram of figure 7, where we notice that:

(1) The video streams position on the diagram depends on the content.

(2) None of the video streams falls in the bottom-right area of the diagram (high values of $(\sigma/\mu)_{\mathrm{PB}}$ and low values of $(\sigma/\mu)_{\mathrm{I}}$). In fact, it is unlikely that a strong variability in size of P-B frames is not accompanied by high values of $(\sigma/\mu)_{\mathrm{I}}$.

(3) Sporting video streams have values of $(\sigma/\mu)_{\mathrm{PB}}$ as low as talk-shows. The sporting shots have much movement, but just in a small part of the image; consider, for example, a shot during a soccer match where only a small part of the image is subject to movements.

In figure 8 and in table 2, we have divided the video sources into three classes, depending on their content.

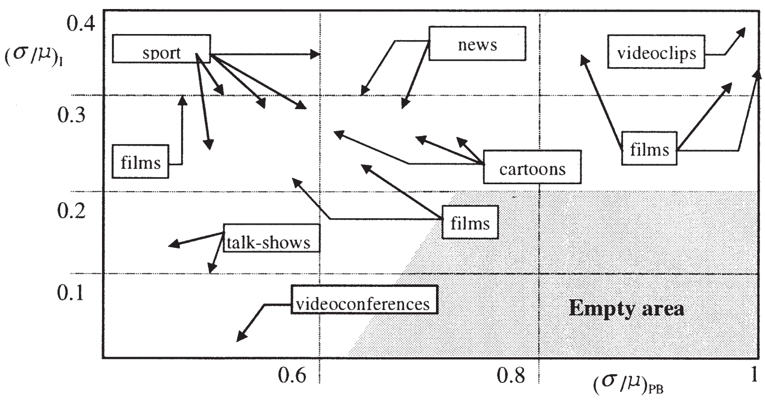


Figure 7. Video streams representation on a bidimensional diagram.

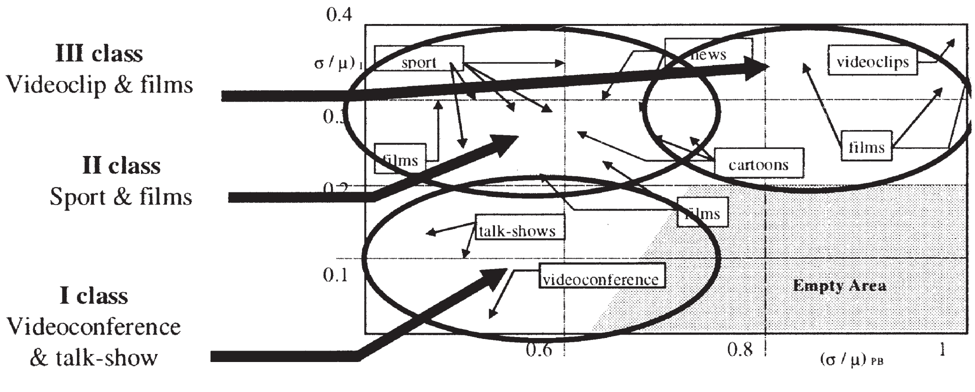


Figure 8. Video streams classification.

Table 2
Content based classification.

	$(\sigma/\mu)_{\mathrm{I}}$	$(\sigma/\mu)_{\mathrm{PB}}$	Video stream content
I class	0−0.2	0−0.7	Videoconferences and talk-shows
II class	>0.2	0−0.7	Sport and cartoons
III class	>0.2	>0.7	Videoclips and movies

Table 3
Maximum and minimum values of normalized bit rate.

Movies	Min	Max
I class	0.7792	1.3602
II class	0.6558	1.5564
III class	0.6751	1.7431

When we aggregate a number of video streams of the same class, the aggregated streams differ in a more noticeable way. Table 3 shows the minimum and maximum normalized bit rate produced when 5 video traces, belonging to the same class, are multiplexed.

The possibility of simulating sources on the basis of their contents, can be used to investigate the most efficient strategies for network resources allocation. In the following section we illustrate a case study where the simulation approach is applied, and the suggested classification is used to assemble mixes of video sources which represent real combinations of video streams.

5. An application scenario

In this section we illustrate how the classification, on the base of their content, of a set of encoded video sources, and the availability of simple models of the sources can

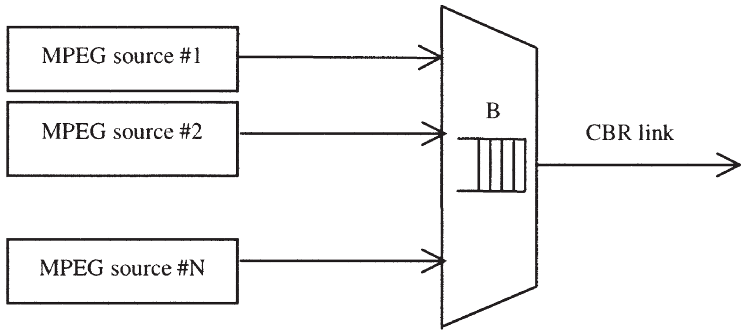


Figure 9. Architecture of the simulated system.

be applied to solve a dimensioning problem for a video delivery system. In order to test the compatibility of a given set of MPEG video streams with an assigned video streams multiplexer, two sets of simulations were performed. The architecture of the video streams multiplexer is depicted in figure 9. A number N of MPEG sources, modeled as described in section 3, with the same average rate $r = 4$ Mbps, and given $(\sigma/\mu)_{\mathrm{I}}$ and $(\sigma/\mu)_{\mathrm{PB}}$, are multiplexed over an output link. In order to accomodate occasional peaks in the bit rate of the aggregated traffic, the streams multiplexer contains a FIFO buffer.

In the first set of experiments, the buffer size B was assumed to be infinite and the N sources were identical. The purpose of these experiments was to evaluate the peak bit rate for different combinations of the $(\sigma/\mu)_{\mathrm{I}}$ and $(\sigma/\mu)_{\mathrm{PB}}$ parameters. We define the *excess bandwidth* as the difference between the peak bandwidth detected during a simulation and the long-term average bandwidth, normalized to the long-term average bandwidth. Table 4 reports the excess bandwidth for different combinations of $N = 6$ identical sources. Table 5 reports the same results when $N = 12$ identical sources were multiplexed. The shaded areas in the tables correspond to sources falling in the "empty area" of figure 8. In both the tables, $(\sigma/\mu)_{\mathrm{I}}$ is constant in each column, while $(\sigma/\mu)_{\mathrm{PB}}$ is constant in each row. The results in tables 4 and 5 suggest that the excess bandwidth depends on the characteristics of the MPEG source, and that, increasing the number of sources, the excess bandwidth slightly decreases.

A second set of simulations was carried out, by assigning a finite size B = 64 Mbit to the buffer in the multiplexer. The output link bandwidth was set to R = 34 Mbps, as for E3 links. On the basis of the classification of MPEG sources that we illustrated in section 4, we selected three different simulated MPEG sources, with the same average rate $r = 4$ Mbps, each of which represented a class of video streams.

Different combinations of sources were chosen, in order to verify their compatibility with the multiplexing system, i.e., we pulled out the combinations that did not lead to buffer overflow. The results of the simulations lead us to the conclusion that all the combinations with less than 7 active sources are compatible with the system.

Table 4
Excess bandwidth for $N = 6$.

	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.40	14.00%	14.49%	15.42%	17.33%	16.93%	19.87%	26.31%	22.53%
0.45	15.51%	15.87%	16.58%	17.60%	20.58%	24.89%	22.80%	23.11%
0.50	16.27%	16.80%	17.91%	18.40%	20.00%	24.67%	21.96%	25.60%
0.55	19.11%	18.31%	19.42%	20.40%	21.64%	22.31%	24.40%	27.64%
0.60	19.56%	20.04%	20.84%	21.91%	22.36%	24.27%	25.51%	26.58%
0.65	21.60%	21.60%	22.58%	21.91%	23.69%	24.98%	24.67%	24.89%
0.70	23.20%	22.80%	23.73%	23.16%	24.58%	26.13%	28.89%	33.02%
0.75	24.44%	24.58%	25.38%	25.29%	26.22%	28.22%	28.67%	33.11%
0.80	26.36%	26.04%	27.24%	27.11%	26.27%	28.80%	29.51%	30.04%
0.85	27.33%	28.13%	27.78%	28.04%	29.24%	31.20%	31.82%	34.49%
0.90	28.98%	28.84%	29.47%	29.91%	30.53%	31.60%	33.38%	34.76%

Table 5
Excess bandwidth for $N = 12$.

	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.40	9.50%	10.93%	11.95%	12.33%	11.50%	14.98%	20.38%	16.03%
0.45	10.64%	12.95%	11.92%	12.65%	14.12%	20.32%	16.39%	16.61%
0.50	11.62%	12.48%	12.79%	13.24%	14.28%	18.32%	15.68%	18.43%
0.55	14.41%	14.18%	13.82%	13.85%	16.32%	17.28%	17.36%	18.77%
0.60	15.96%	14.41%	14.98%	15.03%	18.25%	17.44%	18.33%	18.23%
0.65	16.04%	15.43%	16.25%	15.65%	17.59%	17.84%	17.75%	17.77%
0.70	17.97%	16.22%	16.11%	17.46%	19.04%	18.59%	19.61%	24.90%
0.75	17.57%	17.66%	17.41%	20.64%	18.85%	20.28%	19.67%	27.03%
0.80	18.82%	18.75%	19.45%	20.13%	18.76%	20.73%	21.07%	22.31%
0.85	19.44%	19.10%	20.95%	21.72%	20.80%	21.18%	24.00%	26.71%
0.90	20.82%	19.79%	24.05%	21.50%	21.94%	21.68%	27.25%	24.98%

Table 6
Parameters for a mix of MPEG video sources.

Source	$(\sigma/\mu)_{\mathrm{I}}$	$(\sigma/\mu)_{\mathrm{PB}}$
A	0.15	0.50
B	0.30	0.50
C	0.30	0.80

Table 7
Compatible combinations.

A	B	C
8	0	0
7	1	0
7	0	1
6	2	0
6	1	1

Subsequently, we tested all the combinations with 8 active sources. Only the combinations listed in table 7 never led to buffer overflow. A more detailed analysis of the simulation results, showed that for the combinations not listed in table 7, a much larger buffer in the multiplexer is required.

6. Conclusions

A proper dimensioning of the communication resources is crucial when a multimedia distribution infrastructure is to be built upon wide-area broadband networks. This task is achieved through a trade-off between Quality-of-Service and resources utilization, as a consequence of the characteristics of the traffic produced by encoded video streams.

In this paper we present the application of a simulation based approach to verify the compatibility of an assigned schedule of video streams with a given video delivery system. This approach, relying on simple models for encoded video sources, consists in simulating the behaviour of the target system when loaded by typical mixes of video streams.

In systems like Video-on-Demand servers, a classification of the video streams on the basis of their content, as suggested in section 4, may be helpful to verify the compatibility of the existing network infrastructure with a preassigned mix of video streams. Furthermore, strategies for the most efficient utilization of the available bandwidth can be devised when a number of different video streams have to be delivered from a main storage center towards remote distribution centers. The proposed approach, of course, is applicable only when sufficient time is available to system designers to run the necessary simulations. However, we feel that in systems where the characteristics of the load remain unchanged over a long period of time, this approach can be helpful to dimension the network infrastructure in the most efficient way.

References

[1] R. Canonico, S.P. Romano and G. Ventre, Admission control policies in integrated services networks, Dipartimento di Informatica e Sistemistica, Università di Napoli Federico II, Italy (1998).

[2] P. Crocetti and L. Fratta, Video on demand service based on ATM virtual private networks, in: *Proceedings of ECMAST '96*, Louvain-la-Neuve, Belgium (1996).

[3] L. D'Ardia, Analisi e simulazione di sorgenti di traffico video digitale, Tesi di Laurea, Dipartimento di Informatica e Sistemistica, Università di Napoli Federico II, Italy (1997).

[4] L. D'Ardia, B. Fadini and G. Ventre, Analysis, classification and simulation of multimedia traffic sources, in: *Proceedings of DCCN '97*, Distributed Computer Communication Networks, Tel Aviv (1997).

[5] R. Garg, Characterization of video traffic, Technical Report, TR-95-007, International Computer Science Institute, Berkeley, CA (1995).

[6] D.P. Heyman and T.V. Lakshman, Source models for VBR brodcast-video traffic, IEEE/ACM Transactions on Networking 4(1) (1996) 40–48.

[7] M. Krunz and S.K. Tripathi, Modeling bit rate variations in MPEG sources, Technical Report CS-TR-3573, Department of Computer Science, University of Maryland at College Park (1996). Available at `http://www.cs.umd.edu/TRs/ authors/Satish_K_Tripathi.html`.

[8] A.A. Lazar, G. Pacifici and D.E. Pendarakis, Modeling video sources for real-time scheduling, ACM/Springer, Multimedia Systems 1(5) (1994) 253–266.

[9] D. LeGall, MPEG: a video compression standard for multimedia applications, Communications of the ACM 34(4) (1991) 46–58.

[10] B. Maglaris, Performance models of statistical multiplexer in packet video communications, IEEE Trans. on Communications (June 1988) 834–844.

[11] P. Pancha and M. El Zarki, Leaky bucket access control for VBR MPEG video control, in: *Proceedings of IEEE Infocom '95*, Boston, MA (1995).

[12] O. Rose, Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems, in: *Proceedings of the 20th Annual Conference on Local Computer Networks*, Minneapolis, MN (1995) pp. 397–406.