

A NEW MULTICAST TECHNIQUE FOR VIDEO TRANSMISSION OVER ABR SERVICES IN ATM NETWORKS

Santiago Felici
Andres Lopez-Penaranda
Juan Perez

Institut de Robotica. Universitat de Valencia
Poligono La Coma (Paterna) 46071 Valencia
Ph: +34 96 398 35 63 Fax: +34 96 398 35 50
Santiago.Felici@uv.es

Jorge Martinez
Dept. Comunicaciones. Universidad Politecnica Valencia
jmartinez@dcom.upv.es

Abstract Multicast techniques are the only way to simultaneously provide flows of information from one source to several destinations. The intention of this paper is to study and to evaluate different multicast techniques using a video coder based on an adaptive video compression algorithm with subband coding and a best effort network service like ATM with the Available Bit Rate (ABR) service. This video transmission can adapt faster and easily to changing network conditions. In this way, we present an evaluation process for a determined network configuration. Thereafter we discuss the results obtained by simulation and propose for this video transmission a trade-off between these multicast techniques, in order to obtain the best perceptual video quality.

Keywords: ATM-ABR services, multicast, QoS, adaptive video compression algorithm, multiresolution

1. INTRODUCTION

Different multicast techniques and different network technologies are being analyzed to determine which one offers better performance for multimedia traffic. This study is more relevant when the network offers best effort services because in this scenario it is more restrictive to

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35522-1_37](https://doi.org/10.1007/978-0-387-35522-1_37)

maintain a certain Quality of Service (QoS) using the available network resources. We focus the study on ATM with Available Bit Rate (ABR) service.

The ABR class of service was initially conceived to support data traffic. Its service model is based on the *best-effort* paradigm but enhanced by some specific characteristics: *fair sharing of the available resources among the contending ABR connections* and *a closed-loop feedback mechanism with Resource Management Cells (RM) with each destination*. Nevertheless in a multicast tree, when different connections over the same source are running simultaneously, this closed-loop feedback with each destination becomes a problem, because each destination is providing different information to the source. The switches within the *multicast connection* (or *multicast tree*) have to manage different RM cells to the same source, what is called a multicast congestion control.

The intention of this paper is to evaluate different multicast techniques for ATM-ABR, using a subband based video coder, and finally to propose a suitable multicast technique for this kind of information exchange. The rest of the paper is structured as follows: in Section 2 is explained the network configuration for the evaluation process of these multicast algorithms; Section 3 describes the operation of the adaptive video coder based on subband coding; in Section 4 different multicast algorithms over ATM-ABR are evaluated, trying to compare them; in Section 5, we propose a new algorithm and in Section 6 we evaluate the performance of the proposed technique, providing some numerical results. Finally, Section 7 presents the conclusions and ideas for future work.

2. DEFINITION OF A NETWORK CONFIGURATION FOR THE EVALUATION PROCESS

This section explains the basic assumption for the evaluation of the different multicast techniques. It is necessary to use a network configuration which let us stress the video sequence. The network configuration used in the experiment, is given in Figure 1. It has 4 switches which are multicast capable. Source A is a multicast one and source B is an unicast one working as a greedy source (what means that it uses as much bandwidth as it available for them). The *multicast connection* (or *multicast tree*) has three leaves, one in the second (A1), third (A2) and fourth (A3) switch (from left to right). The links L1 and L3 are 50 km long and the link L2 is 100 km long. The access link is 0.2 km long. The propagation delay is 5 μ sec/km. The ABR sources are explained in [1], and are specified by the ABR source parameters: PCR (Peak Cell

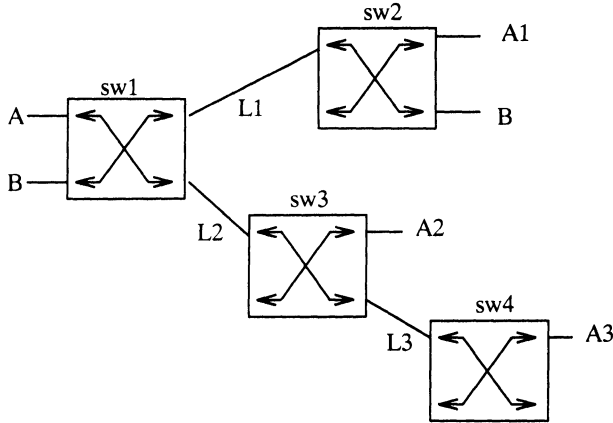


Figure 1 Network configuration

Rate)= 23.85 cells/msec, MCR (Minimum Cell Rate)= 0.8516 cells/msec and ICR (Initial Cell Rate)=2.24 cells/msec. The rest of parameters used for these ABR sources are: Trm=10 msec, ADTF=10000 msec, RIF=0.0625, TCR=0.8516 cells/msec, CDF=0, RDF=0.0625, Mrm=2 cells, Nrm=32 cells, TBE=0, Crm=1000 cells and FRTT=0.

The links have 10 Mbps of bandwidth, apart from link L2 for which the bandwidth is changed. Because the intention of this paper is to evaluate different multicast congestion controls, we compare the ACR (Allowed Cell Rate) of the sources when different changes of the bandwidth are introduced in the configuration. The bandwidth change in link L2 from 10 to 3 Mbps at 150 msec and from 3 to 10 Mbps at 300 msec.

3. ADAPTIVE VIDEO COMPRESSION ALGORITHM AND SUBBAND CODING

Video-based applications that are rate adaptive can obtain substantial benefits over best effort network services, as can be seen in [2]. In ABR connections, these benefits can be summarized in the following three aspects. First, these applications typically require some guarantee on bandwidth, for example a minimum encoding rate for a video stream, but can take advantage of spare bandwidth. This can be supported by an ABR connection using a Minimum Cell Rate (MCR) at the connection set up. Second, when explicit rate feedback is used and the ABR connections supporting these applications are multiplexed on a dedicated queue at the switches, the cell transfer delay is more predictable because the congestion control mechanism keeps the queues almost empty.

And third, the feedback mechanism keeps each source informed of the available bandwidth it has at his disposal.

A video compression algorithm is based on three steps: a decomposition process, a quantization process and finally a entropy coding process. But if adaptive performance is required, each process requires a suitable design. Adaptability means multiresolution, which can be implemented using a subband decomposition or subband coding. A subband decomposition is a process where the information is decomposed in subbands at different levels of resolution; a video signal can be decomposed in a 3D domain, see reference [3].

Each subband has different resolution levels of the original video and if we add all subbands in a reverse decomposition process, we obtain the original video. Obviously, depending on the video information of each subband, not all subbands have the same importance in the human visual system (HVS), because the HVS has different perceptual responses to these subbands. This perceptual priority will determine the order in which subbands are going to be transmitted.

3.1. OPERATION OF THE VIDEO CODER

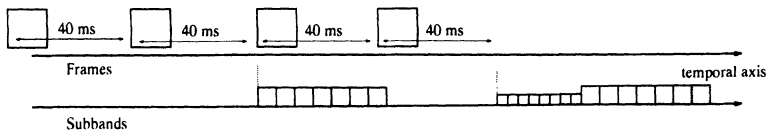


Figure 2 Subband generation using 3D Wavelet Transform with two resolution levels. Different frames are processed every 40 ms

In a system with two resolution levels as explained in [2], a set of 4 frames ($4 \times 40 = 160$ ms) for a video sequence of $25 \frac{\text{frame}}{\text{sec}}$, are needed to perform a complete 3D subband decomposition. This represents a trade off between the decorrelation ratio and the number of frames that need to be stored at the coder. The process can be observed in Figure 2.

For example, let assume that our system is going to perform the decomposition of 4 frames, that we label as frames 1, 2, 3 and 4. The system uses the pair of frames 1-2 and the pair 3-4 to obtain the first resolution level. This process generates 8 subbands from each pair of frames, but because we use the pair of subbands with the lowest resolution (fewer frame details) of the original pair of frames (1-2 and 3-4) to generate the second resolution level, then only 7 subbands remain at the first resolution level. By this process we obtain 8 additional subbands from the second resolution level. Therefore, at the end of the decom-

position process we obtain $7+7+8$ subbands. A full explanation of this subband based video coder can be found in [2].

Once each subband is available, the coder creates an information unit per subband, which contains the necessary information to reconstruct each individual subband. The information unit is called Packet Data Unit (PDU). The order in which the different PDUs are transferred, will define their priority. This order is determined by their perceptual weight as said below.

4. OVERVIEW OF DIFFERENT MULTICAST TECHNIQUES

In this section, we are going to describe the five multicast algorithms published in ATM Forum, which will be used for a multicast video transmission with the video coder of Section 3. Our purpose is to compare these algorithms by the ACR of the sources and how they distribute the available bandwidth between the different sources at the switches. In order to evaluate these parameters, we select the configuration of Figure 1

The multicast algorithms from the literature like [4] are:

- **Fast Indication (FI):** this algorithm proposes that the source transmits at the minimum available bandwidth of all the c branches of the *multicast connection*. When a FRM (Forward RM) cell is received, the switch changes it to a BRM (Backward RM) cell and fills the ER (Explicit Rate), NI (No Increase ACR at source) and CI (Congestion Indication) fields with two classes of information: *external information*, which arrives to the switch by the BRM cells and *internal information*, like queue length and Fair Share calculated in the switch. The minimum between the ER calculated both with the *internal information* and with the *external information*, is written in the BRM cell which is sent. This algorithm has a very fast response but a big consolidation noise (fluctuations).
- **Wait For All (WFA):** this algorithm eliminates the consolidation noise, because it wait one BRM cell for each branch of the *multicast connection* to send a BRM cell to the source. Now, the information is more reliable, because we wait for the complete information of the entire connection which is given feedback to the source. But we have to wait a period of time (called consolidation time) to feedback. Thus, this algorithm has a slow response, because we have to wait the BRM cell from the slowest leaf of the *multicast connection* to send a BRM cell to the source

- **Fast Overload Indication (FOI):** it avoids the consolidation time of the WFA algorithm during the overload period. On the one hand, if there is overload, the switch sends a BRM cell with the information currently it has, like in the FI algorithm. On the other hand, during the steady state the switch works as the WFA algorithm. This algorithm has a problem: the $\frac{BRM}{FRM}$ ratio is greater than one, and the desirable $\frac{BRM}{FRM}$ ratio is 1 [4]. However, this algorithm has a faster response than the WFA algorithm and it has less consolidation noise than the FI algorithm
- **RM Ratio Control (RMRC):** in the FOI algorithm the $\frac{BRM}{FRM}$ ratio is larger than one. To avoid this situation, the RMRC algorithm proposes to control this ratio. During the overload, BRM cells are sent like in the FI algorithm, but during the steady period we do not send BRM cells like in the WFA algorithm. By stead we recover the excess of BRM cells sent in the overload period. This mechanism does not guarantee a $\frac{BRM}{FRM}$ ratio equal to one, but it guarantees a $\frac{BRM}{FRM}$ ratio to be lower than in the FOI algorithm
- **Memory Enhanced (ME):** it is not exactly an algorithm, but a new mechanism to improve the multicast algorithms in order to reduce the consolidation noise. With the ME mechanism, the switch has information of each branch in the *multicast connection*, not only for the whole connection. This mechanism avoid fluctuations around the operating point, because we have more information about the state of each branch of the *multicast connection*

The above algorithms are mechanisms which decide when a BRM cell has to be sent to the source. Nevertheless a unicast algorithm is also necessary to calculate the portion of the available bandwidth for each connection (ER). In this case, the unicast algorithms are CAPAC [5] and ERICA [6]. In summary, the unicast algorithm calculate the ER written in the BRM cells, but these cells are decided to be sent by the multicast algorithm.

The simulations results can be seen in the following figures. Notice that multicast source A is stressed by link L2 meanwhile, the unicast source B uses the excess of bandwidth. In Figure 3 the ACR of the sources are shown, with important oscillations at 150 and 300 msec (with bandwidth changes, see Section 2). It is remarkable that this algorithm has a fast response as can be observed by the slope of the ACR at these times. The unicast algorithm used in these simulations is CAPAC.

By comparison, we simulated the WFA algorithm. The results can be seen in Figure 4. The oscilations are smaller than in the FI algo-

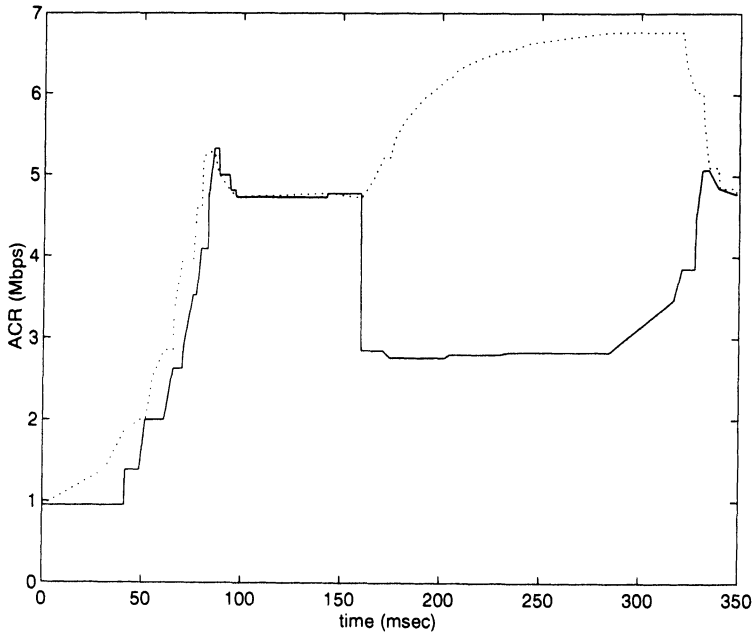


Figure 3 ACR of source A (continuous line) and source B (dashed line) with the Fast Indication algorithm using CAPAC

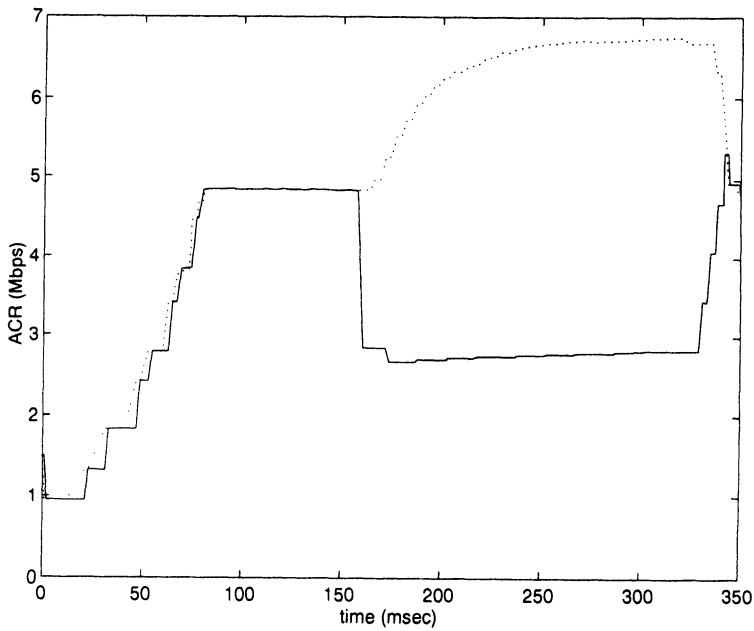


Figure 4 ACR of source A (continuous line) and source B (dashed line) with the Wait For All algorithm using CAPAC

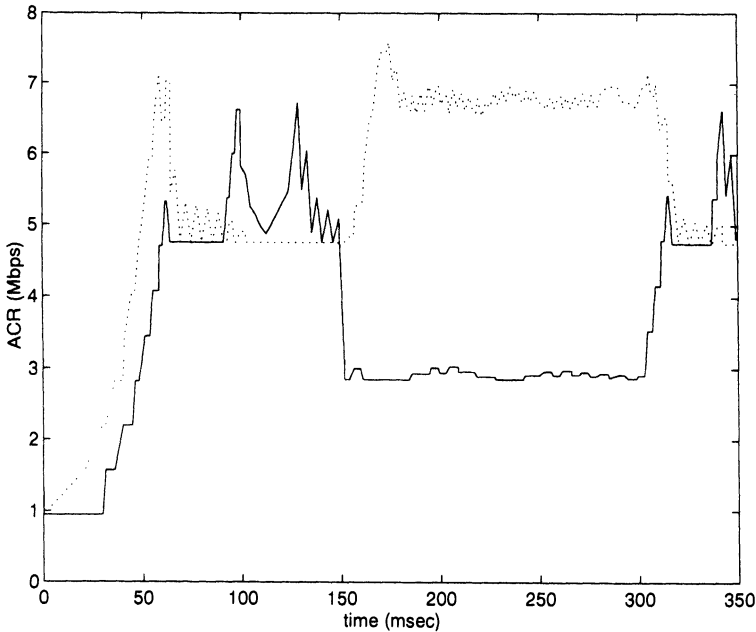


Figure 5 ACR of source A (continuous line) and source B (dashed line) with the Fast Indication algorithm using ERICA

rithm, but the response to these changes is slower. Finally, using the FI algorithm as a multicast algorithm with the ERICA as the unicast algorithm, it has a faster response, with steeper slope than in Figure 3, as can be seen in Figure 5. Notice that oscillations are not relevant for the video coder, because its time scale is greater than these oscillations. These oscillations are produced by the multicast connection.

In conclusion, we see that these multicast algorithms have been designed for data traffic. They need to adapt to the worst situation, the bottleneck of the *multicast connection*, transmitting at the minimum bandwidth. Because video application are rate adaptive and a multiresolution process can be used as explained in Section 3, a better choice than a minimum criteria is proposed as presented in next section, trying to trade-off between minimum and maximum values of ER.

5. PROPOSED MULTICAST ALGORITHM

Previous multicast algorithms try to fit the bottleneck bandwidth in the *multicast connection*. Operating in this way, we force the rest of the destinations to work at the video quality determined by this bottleneck. A maximum value would be unrealistic thus as always, a trade-off be-

tween minimum and maximum is the best option to assign the available bandwidth to the video source.

It is interesting to determine this value through a probability distribution function of the explicit rates, given by each leaf in the *multicast connection*, independently of the distance between source and each destination. This should be a complex but necessary task to implement, however the switch has to do that in an easy way. A valid approximation is presented in this paper.

Our proposed multicast algorithm, called *Trade-Off in a Fair Share* (TOFS) is based on the number of hops from each destination to the switch. The usage of the number of hops in the RM cells, suppose to declare a new field. This is not a real problem because RM cell data are always nearly empty.

When a particular switch within a *multicast connection* receives several BRM cells, for each branch, it calculates the FS (Fair Share) by the minimum value between MER (Medium Explicit Rate, containing *external information*) and the FS given by the unicast algorithm (*internal information*), like the FI algorithm. The calculated FS is the value which corresponds to the FS available for each branch, whereby each branch has a different number of destinations connected. For instance, it can have either one destination or another switch connected to more destinations. To take the number of destinations into account, we compute the number of hops (*nhops*) associated to each branch, weighing each FS with this values.

For the selection of an unicast algorithm to calculate the FS, we have observed in the Figures 3, 4 and 5 that the better response is given by the FI algorithm with the CAPAC congestion control. Nevertheless, because the ERICA algorithm is well known and more oftenly used, we will choose ERICA for the unicast algorithm in the TOFS algorithm.

The proposed TOFS multicast algorithm is resumed by the next expression:

$$TOFS = \sum_i FS_i \frac{nhops_i}{nhops_{total}} \quad (1)$$

where $FS_i = \min(MER, FS_i)$ and $nhops_{total} = \sum_i nhops_i$.

Because the subband video coder is adaptive, it easily achieves a suitable working compression point with a properly bit allocation, where the perceptual distortion is improved.

The proposed multicast algorithm is based on the FI algorithm, but with the ER given by the TOFS algorithm and calculating the number of hops associated to each switch as $nhops = \frac{nhops_{total}}{i}$, being i the number of branches of that switch. These few modifications let us introduce a more realistic scenario, where the video source can take profit and

adjust to a number of different destination, not only the worst case or bottleneck. Also, because this algorithm is a modification of FI, it keeps its properties, like the $\frac{BRM}{FRM}$ ratio nearly to one.

6. RESULTS AND DISCUSSION

First of all it is important to notice that the operation of the explained algorithm matches the ERICA behavior, as it should be expected, when there is one leaf at the multicast tree. On the other hand, as it had been said, the video source does not operate at the ACR given by the bottleneck but a tradeoff within the overall multicast tree.

In Figure 6 the ACR of different sources is shown in order to evaluate and to compare the behavior of TOFS algorithm against the previous multicast algorithms. The multicast source gets more bandwidth than in the previous multicast algorithms.

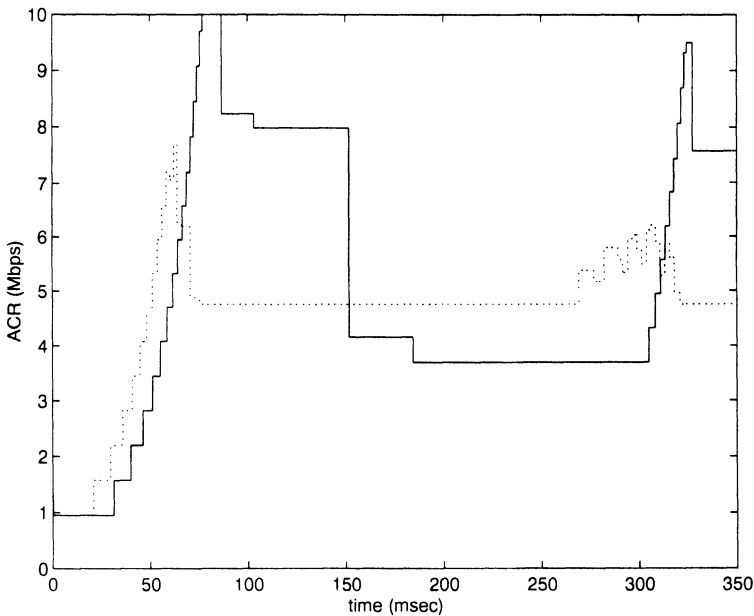


Figure 6 ACR of source A and source B with the TOFS algorithm

A more exhaustive evaluation with several video sequences and a new methodology to get more reliable measures in a subband video coder can be found in [7].

7. CONCLUSIONS AND FUTURE WORK

In conclusion, while ABR services over ATM have been designed for data traffic, a number of studies show that video transmission can take profit from these best effort services, even in a *multicast connection*. Furthermore, we have shown that for adaptive video transmissions, in this case based on subband coding, working at the bandwidth given by the bottleneck of the *multicast connection* is not the best choice and then an intermediate solution is better. The proposed solution in this paper is called TOFS algorithm. Further studies, will be carried out using CAPAC as unicast algorithm.

In future work, the same could be done for IP networks, because multicast techniques although the technology of ATM and IP differs show certain similarities. Over IP, we have similar mechanisms to RM cells using Real Time Protocol and Real Time Control Protocol (RFC1889 and RFC1990), but in this case, we should need to consider the QoS using Differentiated services(DS)(RFC2475).

References

- [1] S.Fahmy R. Goyal R. Jain, S. Kalyanaraman and S. Kim. Source Behavior for ATM ABR Traffic Management: An Explanation. *IEEE Communications Magazine*, 34(11):50–57, November 1996.
- [2] S. Felici and J. Martinez. A video Compression Algorithm for ATM Networks with ABR services, using visual criteria. *Lectures Notes in Computer Science: 6th Workshop Interactive Distributed Multimedia Systems and Telecommunications Services Spriger- Verlag*, 1718:245–258, October 1999.
- [3] Gilbert Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, USA, 1996.
- [4] Rohit Goyal Bobby Vandalore Shivkumar Kalyanaraman Sastri Kota Pradeep Samudra Raj Jain, Sonia Fahmy. Feedback consolidation algorithms for abr point-to-multipoint connections. *ATM Forum Document/97-0615*, Julio 1997.
- [5] J.R. Vidal J. Martinez and L. Guijarro. A Low Complexity Congestion Control Algorithm for the ABR Class of Service. In *International Distributed Multimedia Systems IDMS98, Oslo (Norway)*, pages 219–230, September 1998.
- [6] Shivkumar Kalyanaraman Rohit Goyal Ram Viswanathan Sonia Fahmy, Raj Jain. Erica switch algorithm: A complete description. *ATM Forum Document/96-1172*, Agosto 1996.
- [7] S. Felici and J. Martinez. A statistical method to evaluate adaptive video applications based on subband codign over best effort networks, particularly with abr services. *Proc. of Access to the Internet in the next century, EUNICE99*, pages 65–70, 1999.