Evaluation of different Multicast Techniques for Video Transmission based on Subband Coding over ABR services in ATM networks

S.FELICI, A.LOPEZ-PENARANDA [(1)], J.MARTINEZ [(1)] and J.PEREZ
Institut de Robotica. Universitat de Valencia
Poligono La Coma (Paterna) 46071 Valencia
SPAIN, Ph: +34 96 398 35 63 Fax: +34 96 398 35 50 ma
[(1)] Dept. Comunicaciones. Universidad Politecnica Valencia

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

Key-Words: 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 maintain a certain Quality of Service (QoS) using the available network resources. We focuses the study in 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 in a problem, because each destination is providing different information to the source. Then, 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 for this a subband based video coder, and finally after that to propose a suitable multicast technique for this kind of information. 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 gives an explanation for 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; after that in section 5, we propose a new one 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

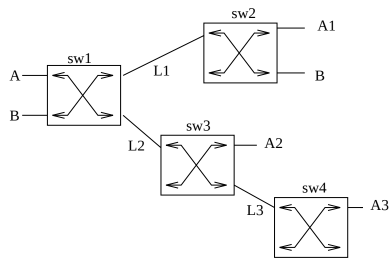

Figure 1: *Network configuration*

This section explains the basis over which is intended to carry out 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, can be observed in figure 1 and it has 4 switches, which are multicast capable. The source A is a multicast one and source B is an unicast one working as a greedy source[1]. 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 switches 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 μsec/Km. The ABR sources are explained in [1], and using its notation, the ABR source parameters used are: PCR(Peak Cell Rate)= 23.85 cells/msec, MCR(Minimum Cell Rate)= 0.8516 cells/msec and ICR(Initial Cell Rate)=2.24 cells/msec. More details are given at the bottom of this page[2].

The links have 10 Mbps of bandwidth, unless different changes of the bandwidth required in link L2. Because the intention of this paper is just to evaluate different multicast congestion control, then we compare the ACR (Allowed Cell Rate)of the sources when different changes of the bandwidth are introduced in the configuration. These changes of bandwidth are in link L2 from 10 to 3 Mbps at 150 msec and finally from 3 to 10 Mbps at 300 msec.

[1]Greedy sources use as much bandwidth as it available for them

[2]The rest of parameters used for these ABR sources are: Trm=10 mseg, 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, FRTT=0

3 Adaptive video compression algorithm and subband coding

Over best effort network services, video-based applications that are rate adaptive can obtain substantial benefits 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 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 their 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, and multiresolution 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; for a video signal can be decomposed in a 3D domain, see reference [3].

Each subband has different resolution level 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 subband have same importance from the human visual system (HVS) point of view, because human has different perceptual responses to these subbands. This perceptual priority will determine the order in which subband are going to be transmitted.

3.1 Operation of the video coder

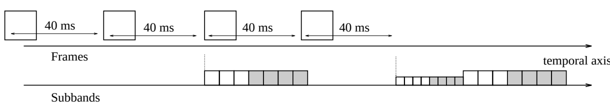

Figure 2: *Subband generation using 3D Wavelet Transform with two resolution levels. Different frames are processed every 40 ms*

In a system with two resolution levels as explained in [2], using a video sequence of 25 $\frac{frame}{sec}$, then a set of 4 frames (4x40=160 ms) 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 lowest resolution (fewer frame details) from the original pair of frames (1-2 and 3-4) to generate the second resolution level, then only 7 subbands remain at first resolution level. By this process we obtain 8 additional subbands from second resolution level. Therefore, at the end of the decomposition 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 distribute the available bandwidth between the different sources at the switches. In order to evaluate these parameters, we select the configuration of figure 1 because it tests these algorithms, using different changes of bandwidth at their links.

The multicast algorithms we can find in the literature like [4], are:

- **Fast Indication (FI)**: this algorithm proposes that the source transmits at the minimum available bandwidth of all 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 minimun 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 information of the entire connection to give feedback to the source. But we have to wait a period of time (called consolidation time) to feedback. Then, this algorithm has a slow response, because we have to wait the BRM cell from the farthest leaf of the *multicast connection* to send a BRM cell to the source

- **Fast Overload Indication (FOI)**: it removes the consolidation time of the WFA algorithm during the overload period. On the one hand, if there is overload, the switch send a BRM cell with the information it has, like the FI algorithm. On the other hand, during the steady state the switch works as the WFA algorithm. This algorithm has a problem: the BRM/FRM ratio is greater than one[3]. 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 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, nevertheless during the steady period we do not send BRM cells like in the WFA algorithm, but we recover the excess of BRM cells sent in the overload period. This mechanism does not guarantee a BRM/FRM ratio equal to one, but it guarantees a BRM/FRM ratio lower than 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 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 and then, the unicast source B uses the excess of bandwidth. In figure 3 are shown the ACR of the sources, with important oscillations at 150 and 300 msec (with bandwidth changes, see section 2). It is remarkable, 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.

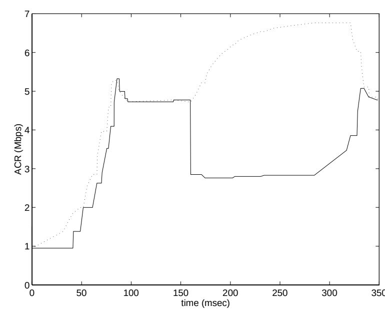

Figure 3: *ACR of source A (continuous line) and source B (discontuous line) with Fast Indication algorithm using CAPAC*

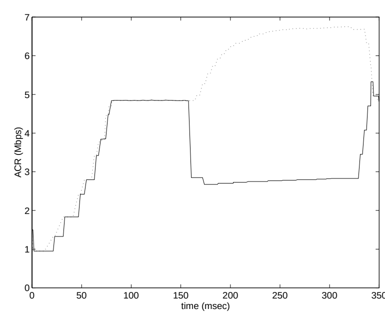

Figure 4: *ACR of source A (continuous line) and source B (discontuous line) with Wait For All algorithm using CAPAC*

By comparison, we simulated the WFA algorithm, and the results can be seen in figure 4. The oscillations are smaller than in the FI algorithm, but the response to these changes is slower. Finally, using FI algorithm as a multicast algorithm with the ERICA as the unicast algorithm, it has a faster response, with grater 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 have to say that these multi-

[3] The desirable BRM/FRM ratio is 1 [4]

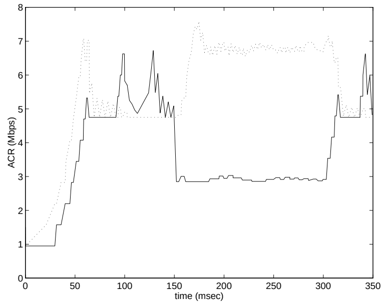

Figure 5: *ACR of source A (continuous line) and source B (discontuous line) with Fast Indication algorithm using ERICA*

cast algorithms have been designed for data traffic, because they need to adapt 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 the proposal presented in next section, trying to take a trade-off between minimum and maximum values of ER.

5 Proposed multicast algorithm

Previous multicast algorithms try to fit the bottleneck in the *multicast connection*. Nevertheless operating in this way, we force the rest of destinations to work at the worst video quality determined by this bottleneck. Maybe a maximum value should be unrealistic and as always, a trade-off between min. and max. 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 task to implement but necessary, 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, but there is no problem because RM cell data are always nearly empty.

When a determined 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, but each branch has connected a different number of destinations. For instance, it can has either one destination or another switch, which will connect to more destinations. To take into account the number of destinations, we compute the number of hops (*nhops*) associated to each branch, weighing up each FS with this values.

About the discussion of an unicast algorithm to calculate the FS, we can 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 the ERICA for the unicast algorithm in the TOFS algorithm.

Next expressions resume the proposed TOFS multicast algorithm

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

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

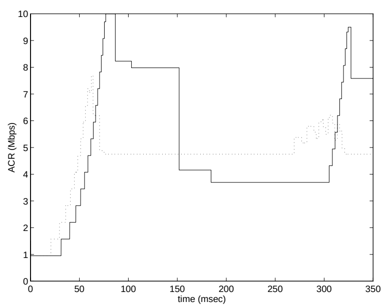

Figure 6: *ACR of source A and source B with TOFS algorithm*

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

As conclusion, meanwhile ABR services over ATM have been designed for data traffic, a number of studies show that video transmission can take profit of 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.

As future work, the same could be done for IP networks, because multicast techniques although differ from ATM and IP technology, they show certain similarities. Over IP, we have a 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 study 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*, 34:50–57, 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 *Internation Distributed Multimedia Systems IDMS98, Oslo (Norway)*, pages 219–230, September 1998.

[6] S.Kalyanaraman R.Goyal S.Fahmy, R.Jain and R.Viswanathan. 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 appplications based on subband codign over best effort networks, particularly with abr services. In *Proc. of Access to the Internet in the next century, EUNICE99*, pages 65–70, 1999.