

IMPACT OF SCALABILITY IN VIDEO TRANSMISSION IN PROMOTION-CAPABLE DIFFERENTIATED SERVICES NETWORKS

Eren Gurses, Gozde Bozdagi Akar

Middle East Technical Univ
Dept. of Electrical and Electronics Eng.
Ankara, Turkey

Nail Akar

Bilkent University
Dept. of Electrical and Electronics Eng.
Ankara, Turkey

ABSTRACT

Transmission of high quality video over the Internet faces many challenges including unpredictable packet loss characteristics of the current Internet and the heterogeneity of receivers in terms of their bandwidth and processing capabilities. To address these challenges, we propose an architecture in this paper that is based on the temporally scalable and error resilient video coding mode of the H.263+ codec. In this architecture, the video frames will be transported over a new generation IP network that supports differentiated services (Diffserv). We also propose a novel Two Rate Three Color Promotion-Capable Marker (trTCPCM) to be used at the edge of the diffserv network. Our simulation study demonstrates that an average of 30 dB can be achieved in case of highly congested links.

1. INTRODUCTION

Transmission of high quality video over the Internet is now becoming a reality due to progresses in video compression, networking technologies, efficient video coders/decoders and increasing interest in applications such as video on demand, videophone, and videoconferencing. To fulfill different receiver requirements by using one common bitstream, techniques which can simultaneously support a variety of bitrates are needed while maintaining end-to-end quality. Coding video in a scalable manner partially solves this problem by offering different rates to different users. For maintaining end-to-end quality, two QoS (Quality of Service) architectures have been proposed by the IETF (Internet Engineering Task Force): the integrated services (IntServ) with the resource reservation protocol (RSVP) and the differentiated services (Diffserv). Diffserv provides a less complicated and scalable solution compared to Intserv, which fits very well to the structure of scalable video coding. Recently, several studies have been done on transmitting scalable video (MPEG-2, H.263+, MPEG-4) over Diffserv networks. In [1], Markopoulou and Hang address the issue of transmission of scalable video (H.263+) in contexts where packet drops, rather than packet delays, are the primary determinant of application performance. However, in this work only SNR scalability is used and there is no policing al-

gorithm involved at the edge to check the conformance of incoming packets. In [2], Shin et. al. use a relative priority index to represent the relative preference of each packet in terms of loss and delay. Instead of using scalable video, their work is based on full scale video.

In this paper, we propose a new architecture for transmitting H.263+ video over a Diffserv network. Our first contribution in this paper is based on a new approach for layering of the bitstream. We propose that interframes selected by the reference picture selection mode of H.263+ are transmitted in the base layer as opposed to the enhancement layer. Our simulation results using this approach demonstrate better bandwidth utilization and error resilience in comparison with the following two cases: i) non-scalable coding ii) temporal scalability without reference picture selection. The second contribution of this paper is the proposal of a new Two Rate Three Color Promotion-Capable Marker (trTCPCM) to be used at the edge of the Diffserv network. As opposed to conventional markers that check the conformance of incoming packets and that only demote the color of the packet in case of non-conformance, this new marker can promote as well as demote, the color of an incoming packet. In this study, we refer to the differentiated services networks using trTCPCM at the edge as promotion-capable differentiated services networks.

2. TEMPORAL SCALABILITY

Many scalable video-coding techniques have been proposed over the past few years for real-time Internet applications by several video compression standards such as MPEG-2/4 and H.263/263+ [3]. The types of scalability which are defined in these standards can be categorized as *temporal*, *spatial*, *SNR*, and *object* (only for MPEG4) scalability. All these types of scalable video consist of a Base Layer (BL) which is the minimum amount of data needed for decoding the video stream and one or more Enhancement Layers (EL). Both the base layer and the enhancement layer can be composed of I-P-B pictures which are the three generic picture

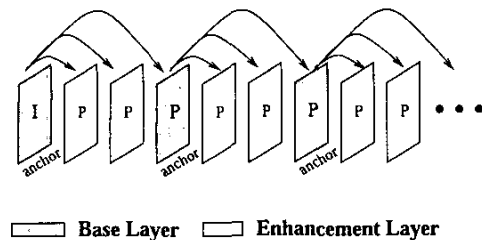


Fig. 1. Temporal Scalable video coding

types used in the above-mentioned standards. Other than the temporal scalability, SNR scalability is also widely used especially in video transmission over a Diffserv network [1], [2], [?]. In SNR scalability studies, BL is formed of I, P, and B pictures with a coarser quality. One of the drawbacks of this approach is that when one of the Enhancement Layer-P frames (*EP*) is lost, the following *EP*'s quality will degrade. Another scalability structure that is suitable for the Diffserv architecture is *Fine-Granular-Scalability* (FGS) [6].

In FGS, there is no temporal relation among the frames in the EL [6]. Since in FGS EL is formed of bitplane blocks which are DCT coded, bandwidth may be utilized more efficiently. However because of lack of temporal relation, increase in bitrate occurs especially in cases where the BL bitrate is chosen to be small as compared to the total rate.

In order to solve the above-mentioned problems, we used the *Reference Picture Selection* mode of H.263+ (Annex N)[3] as in Figure 1 in this work in order to achieve scalability. In this figure, reference pictures are shown as anchor frames. This is a simpler version of the temporal scalability mode of H.263+ (Annex O) [3], with backward prediction disabled. Since in a Diffserv architecture BL and EL can be marked differently, choosing the frames of EL by *Reference Picture Selection* is expected to decrease the degradation in PSNR in case of inter frame losses.

3. TRTCPCM

Diffserv is essentially a *priority dropping* mechanism which defines different service classes [4] for applications with different QoS requirements. An end-to-end service differentiation is obtained by concatenation of per-domain services and Service Level Agreements (SLAs) between adjoining domains along the path from source to destination (Figure 2). Per domain services are realized by traffic conditioning such as classification, metering, policing, shaping at the edge and simple differentiated forwarding mechanisms at the core of the network. Two of the more popular proposed forwarding mechanisms are Expedited Forwarding (EF) and Assured Forwarding (AF) Per Hop Behaviors (PHB). Since AF may enable service offerings at lesser cost

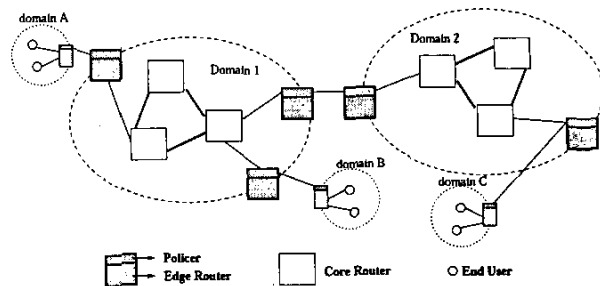


Fig. 2. A Typical Diffserv Topology

than EF for audio, video, Web and other applications, we used AF PHB (RFC 2597) for transmitting scalable H.263+ bitstream in this work.

The AF PHB defines four AF (Assured Forwarding) classes: AF1, AF2, AF3, and AF4. Each class is assigned a specific amount of buffer space and bandwidth. Within each AF class, one can specify three drop precedence values: 1, 2, and 3. In the notation AF_ny , n represents the AF class number (1, 2, 3, or 4) and y represents the drop precedence value (1, 2, or 3) within the AF_n class. In instances of network congestion, if packets in a particular AF class (for example, AF1) need to be dropped, those packets will be dropped according to the following guideline:

$$P(AF11) < P(AF12) < P(AF13),$$

where $P(AFny)$ is the drop probability of the subclass $AFny$.

For traffic conditioning, a Two Rate Three Color Marker (trTCM), is commonly used in Diffserv architectures [8]. The trTCM meters an IP packet stream and marks its packets based on two rates, Peak Information Rate (PIR) and Committed Information Rate (CIR), and their associated burst sizes Peak Burst Size (PBS) and Committed Burst Size (CBS). The trTCPCM we propose in this study is an extension of the trTCM. The trTCM does not have the capability to promote the drop precedence of a packet whereas our trTCPCM can, while making sure that SLAs will not be violated. Specifically, trTCPCM consists of two token buckets p and c , where $T_p(t)$ and $T_c(t)$ are the token counts of the token buckets p and c , respectively, at time t . PIR and CIR are the filling rates and CBS and PBS are the depths of the token buckets p and c , respectively. We assume that initially both buckets are full. Table 1 presents the marking algorithm we propose that is capable of promoting lower drop precedence packets to higher drop precedence if needed. The main idea behind this marker is, if the 'actual AF11 rate' (i.e., bitrate of the AF11 marked VBR source) is sufficiently below the committed rate (i.e., CIR), we promote some of the AF13 packets to AF11 in order to get a better treatment from the network. This promotional packets can especially be very useful when the contract rates and the actual rates are dif-

Table 1. trTCPCM Algorithm

<pre> if (codePt == Red){ if ($T_p(t) - B > P_t$) { if ($T_c(t) - B > C_t$){ codePt = Green $T_p = T_p - B$ $T_c = T_c - B$ }else{ codePt = Yellow $T_p = T_p - B$ } } else { codePt = Red } else if (codePt == Yellow){ if ($T_p(t) - B > 0$) { if ($T_c(t) - B > C_t$) { codePt = Green $T_p = T_p - B$ $T_c = T_c - B$ } } } </pre>	<pre> } else { codePt = Yellow $T_p = T_p - B$ } } else { codePt = Red } else if (codePt == Green){ if ($T_p(t) - B < 0$) codePt = Red else if ($T_c(t) - B < 0$) { $T_p = T_p - B$ codePt = Yellow } else { codePt = Green $T_p = T_p - B$ $T_c = T_c - B$ } } </pre>
---	--

ferent. In Table 1, red, yellow, and green packets denote packets marked as AF13, AF12, and AF11, respectively. The incoming packet to be marked is assumed to be of size B and P_t and C_t are the threshold values used in the algorithm. and selected as $P_t = 0.7 \times PBS$, $C_t = 0.7 \times CBS$.

4. PERFORMANCE STUDY

We use ns-2 in this simulation study [10]. In all the simulations carried out in this section, the topology shown in Fig 3 is used. 11 traffic sources are connected to the Diffserv domain through the edge router $edge_1$, and two traffic sinks, $dest_1$, $dest_2$ are connected via edge router $edge_2$. Out of the 11 sources, one is CBR (Constant Bit Rate) and the remaining 10 sources are VBR(video) traffic generators. In order to model the AF13 low priority background traffic, the node with label s_1 is used in CBR mode. Video source with label s_2 , is the tagged source to monitor the associated flow and to calculate the PSNR (Peak Signal-to-Noise Ratio). Nodes with labels $s_3..s_{11}$, are the remaining 9 video sources, which may be activated/deactivated throughout the simulations. Each video source is transmitting the 'foreman' sequence with 400 frames coded at 25 frames/sec. All video sources start transmission of the same video from random starting points relative to the tagged source s_2 , in order to prevent synchronization. All video sources mark their BL with AF11 and the EL with AF13. However, these markings can change while the video packets traverse the prosed marker at the edge according to the policing rules dictated as in Table I. The link between $core$ and $edge_2$ is the bottleneck link with a capacity of $0.5Mbps$, which approximately corresponds to the sum of CIRs of 9 video sources, according to the CIRs given in Table 2. Other links in this simulation have a capacity of 1Mbps. Video sources send intra

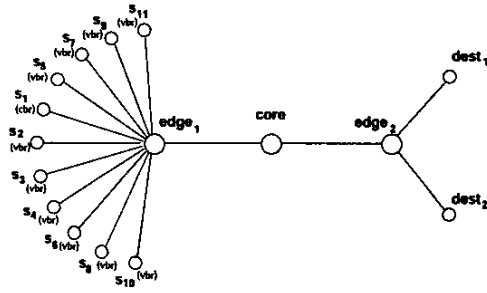


Fig. 3. Diffserv network topology

Table 2. Information about the traffic sources

simulation name	source label	type	rate (kbps)	BL (kbps)	CIR (kbps)	CBS (byte)	PIR (kbps)	PBS (byte)
(valid in all)	s_1	CBR	200	0	0	0	0	0
no-layer	$s_2 - s_{11}$	VBR	86.826	0	56	8000	110	4000
intras	$s_2 - s_{11}$	VBR	86.826	19.845	56	8000	110	4000
rps	$s_2 - s_{11}$	VBR	105.357	50.8635	56	8000	110	4000

frames for each 25 frames, and an anchor frame in every 3 frames. Simulation "rps" uses the Reference Picture Selection mode described in Section II, in order to obtain temporal scalability. The intra and anchor frames constitute the BL at 50.8635 kbps. In simulation "intras", sources mark only the intra frames(25 fps) with AF11, which corresponds to the BL at 19.845 kbps whereas in simulation "no-layer" there is no base layer definition. Therefore all packets for the "no-layer" case are premarked as AF13. In Figure 4, the above simulations are carried out for 7 active video sources (s_2-s_8). Figure 4 depict the per frame PSNR of the received video for the three different scalability modes. In "intras" type of temporal scalability simulation, intra frames are protected with the priority dropping mechanism, however an inter frame loss still may cause a sharp degradation in the PSNR until the arrival of the next intra-frame. In "rps" type

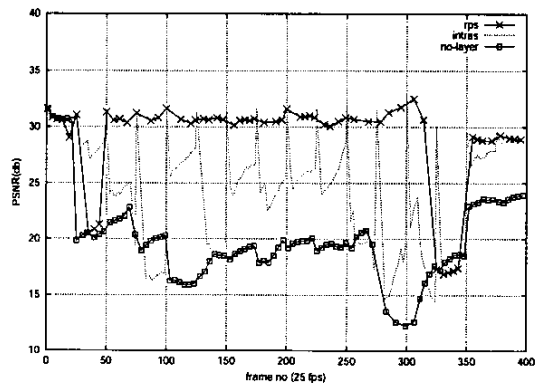


Fig. 4. PSNR plot for received frames

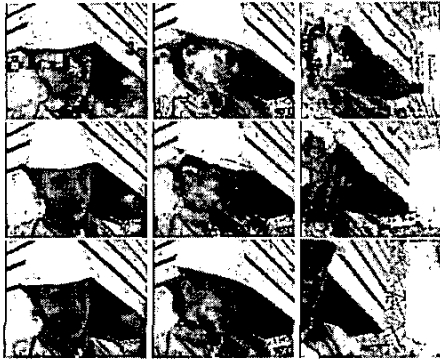


Fig. 5. Frames 253, 265, 287 from simulations “no-layer”, “intras” and “rps” respectively

Table 3. Comparison of trTCM and trTCPCM markers using ‘rps’ simulation at nodes s_2 , $edge_1$, $dest_1$

	Code Point	s_2 (kbps)	$edge_1$ (kbps)	$dest_1$ (kbps)
trTCM	AF11	50.8635	50.8635	50.1835
	AF12	0	0	0
	AF13	54.494	54.494	16.710
	Total:	105.3575		66.8935
trTCPCM	AF11	50.8635	58.078	55.499
	AF12	0	34.3695	19.8645
	AF13	54.494	12.910	0.424
	Total:	105.3575		75.7875

of temporal scalability simulation, the degradation in PSNR in case of inter frame losses is prevented by the special coding method based on reference picture selection. There is no drop in the PSNR unless an intra or an anchor frame is lost. Since the bottleneck link is highly congested in all simulations, a very small percentage of AF12 and AF13 packets can be transmitted without loss. Since in “rps” the anchor frames are also marked as AF11, the degradation in PSNR in case of inter frame losses is prevented and better quality is achieved. In Figure 5 the snapshots of the received video are also given for demonstrating the visual quality of “rps” versus “no-layer” and “intras”.

From the above results, it can be concluded that transmission of “scalable video”, by putting the BL into AF11 and EL into AF13 packets, over a diffserv network, enables the receiver to get a temporally scaled-down video based on the congestion in the network. In this simulation, we also studied the benefit of using trTCPCM at the edge with respect to the conventional trTCM. The contracted CIR of each video source is 56 Kbps which is slightly larger than the mean bitrate of AF11 marked packets of the tagged video source which is 50.8635 Kbps (Table 3). The trTCPCM promotes some of the AF13 packets to AF12 or AF11 while still conforming to the token bucket constraints. Demotion of some packets also took place in trTCPCM. In contrast, the trTCM only demoted the color of the packets With

trTCPCM, more packets are colored with AF11 compared to trTCM which explains the better throughput observed at the destination in Table 3.

5. CONCLUSION

In this paper, we developed a simulation-based framework to evaluate the performance of H.263+ video over a Diffserv network using different modes of scalability. Temporal scalability with the reference picture selection mode is shown to provide better results in terms of both PSNR and subjective video quality when compared to non-scalable coding (“no-layer”) and temporal scalability without the reference picture selection mode (“intras” mode). Future work needs to be done to compare SNR scalability with rps-based temporal scalability for the transmission of video over a diffserv network, however our preliminary results favor the latter in terms of subjective video quality. A hybrid temporal and various FGS scalability mode of operation appears to be promising and is also left as future work. One other contribution of this paper is the introduction of a novel policing algorithm, namely the trTCPCM, that is also capable of promoting packets which was absent in the trTCM policing algorithm. Promotion capability will be critical in occasions where the actual rates and the contracted rates using assured services are different.

6. REFERENCES

- [1] A. Markopoulou, S. Han, “Transmitting Scalable Video over a DiffServ Network,” Final Project, Stanford Univ., 2001.
- [2] J. Shin *et al.* “Quality-of-Service Mapping Mechanism for Packet Video in Differentiated Services Network”, IEEE Trans on Multimedia, vol.3, no.2, June 2001, pp. 219-231.
- [3] ITU-T Recommendation H.263+, “Video Coding for Low Bit Rate Communication”, 1998.
- [4] J. Heinanen, F. Baker, W. Weiss, J. Wroclawski, “Assured Forwarding PHB Group”, RFC 2597, June 1999.
- [5] S. Blake *et al.* “An Architecture for Differentiated Services”, RFC 2475, Dec 1998.
- [6] H.M. Radha *et al.* “The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming Over IP”, IEEE Trans. on Multimedia, vol.3, no.1 Mar 2001.
- [7] G. Côté *et al.* “H.263+: Video Coding at Low Bit Rates”, IEEE CSVT, vol.8, no.7, Nov 1998.
- [8] J.Heinanen, R. Guerin, “A Two Rate Three Color Marker”, RFC 2698, Sep 1998.
- [9] F. Wang *et al.* “A Random Early Demotion and Promotion Marker for Assured Services”, IEEE JSAC, vol.18, no.12, Dec 2000.
- [10] “Network Simulator, ns-2”, <http://www.isi.edu/nsnam/ns/>