P2P GROUP COMMUNICATION USING SCALABLE VIDEO CODING

Yago Sánchez[1], Thomas Schierl[2,1], Cornelius Hellge[1], and Thomas Wiegand[1,2]

[1]Image Communication Chair
Department of Telecommunication Systems
Technical University of Berlin
Einsteinufer 17, 10587 Berlin, Germany

[2]Image Processing Department
Fraunhofer Institute for Telecommunications
Heinrich Hertz Institute
Einsteinufer 37, 10587 Berlin, Germany

ABSTRACT

P2P-streaming has become of high interest in the last years, since it reduces the load on expensive servers, due to the participation of receivers in the media transmission. In this paper, P2P content delivery is shown to be a promising technique for video group communication, for which the main requirement is low-delay. Combining low-delay encoding and low-delay P2P Application Layer Multicast makes it possible to fulfill the delay constraints for interactive group communication applications. For such an application, congestion is a considerable problem, since it causes packet loss or late arrival of the packets, degrading the quality of the service. The results presented in this paper show how rate adaptation in combination with the Scalable Video Coding (SVC) helps to overcome problems in the network, providing a better solution than when non-adaptive single layer coding is transmitted.

Index Terms— Peer-to-Peer, interactive, video communication, SVC, rate adaptation, inter-peer congestion

1. INTRODUCTION

In the last few years P2P infrastructures have been increasingly used for streaming live video content over the Internet. Previous studies [1][2] have already pointed out that relying on these overlay structures increases the latency of the system, since the packets have to pass through intermediate nodes. Moreover, it was also argued that overlay structures are intrinsically unreliable since network problems such as congestion not only affect a particular node, but all nodes to which this node should forward packets.

These aspects are of a big importance in group communication, since delays have to be kept small and therefore, the time to provide error control is quite limited. In spite of the increased latency, P2P content delivery is still a good solution for group communication for a reasonable number of participants. This number is restricted by the physical layer characteristics, such as propagation delay in the links and throughput of the peers. In fact, in [3] it is shown that P2P streaming can also be used for lower delay scenarios, if appropriate P2P topologies and transport mechanisms are deployed, such as congestion-aware multi-tree delivery. Since the approach in [3] is based on H.264/AVC single layer coding (SLC) and adaptation using temporal scalability, the resulting video signal shows gaps in the displayed video sequence when transmission errors cannot be corrected.

Other non-P2P based communication approaches such as [4] have already shown the potential of SVC [5] in conversational video scenarios. The application in [4] relies on a central node, which is responsible for multiplexing different incoming video signals and forwarding them to the different users. As the number of participants increases, e.g. in interactive communication, P2P structures may help to reduce the load in this central node.

However, P2P structures are inherently unreliable. They may suffer from link outages or link congestion, which may affect not only a unique participant but a larger amount of users. Congestion control for interactive communication services is an especially important issue to take into account, since link congestion not only results in packet losses but also in the late arrival of packets. Late delivered packets which are useless since their play-out time has already passed, result thus in a waste of the available throughput, already reduced by congestion.

The main focus of this work is to improve the media delivery during periods of time under congestion, by applying a rate adaptation based on the priority of the data, i.e. sending lower and more important layers first and then higher layers if possible for the Scalable Video Coding.

The remainder of the paper is organized as follows. In section 2, the low-delay P2P communication is described. Section 3 explains the SVC rate adaptation used to overcome periods during which the network is congested. In section 4 the performed simulations and results are presented and the paper is concluded in section 5.

2. LOW-DELAY P2P COMMUNICATION

It is well known that one of the main factors which contribute to give a perception of real-time applications is the end-to-end delay. In order to operate at reasonable delay, the P2P structure may be built as a tree-based Application Layer Multicast (ALM) structure [2][3][6], which implies the use of push-based media delivery between nodes. Other approaches, pull-based mesh P2P architectures [7], introduce higher delays due to overlay construction and data delivery [8], which is not appropriate for the selected use-case. Figure 1 shows the distribution of the video of the presenter peer among the P2P overlay. A multiple-tree based overlay as in [3][6] is selected to increase the reliability though path diversity, where a peer is placed in all trees as shown in Figure 1 with the circled peer. Thus the loss of connectivity in a tree leads to a partial loss of the data, instead of the whole content.

The one way end-to-end delay consist of the so called worst-case-delay, i.e. the maximum delay between the source and the slowest receiver, and the *frame-transmission-delay*, which is the time required by the source to upload the biggest encoded frame. Additionally, the *coding delay* should be taken into account, which

is the delay imposed by the coding structure. Note that, SVC low-delay encoding using hierarchical P-frames has been selected to limit the aforementioned delays and to meet the recommendation for interactive communications, so that the low-delay P2P communication can be achieved.
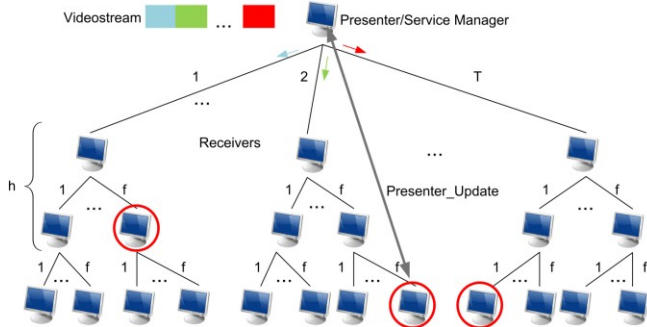

Figure 1: P2P ALM trees and speaker change example

Eq. (1) shows the worst-case-delay d_{max}, which consists of the propagation delay and the queuing delay. The former corresponds to the delay in the links and the latter to the delay incurred at relay nodes to forward packets to all dependent nodes (fanout). These delays depend on the number of fanout (f), the one-link-propagation-delay (T_p), upload throughput (B), maximum packet size (L) and the depth of the trees, illustrated in Figure 1 by the levels of inner nodes (h). Note that the delay at the source for uploading each packet is already included in the frame-transmission-delay (cf. [9]).

$$\overbrace{d \max}^{worst-case-delay} = \overbrace{(h+1)*T_p}^{propagatin-delay} + \overbrace{h*f*\frac{L}{B}}^{queuing-delay} \quad (1)$$

One aspect of group communication is that participants can decide to speak at any point when the current presenter has finished speaking. Therefore, a change of Presenter/Service Manager should occur without implying any disturbance on the communication service and within a small period of time. This could be accomplished by the Presenter_Update message shown in Figure 1, which consists of two handshakes. The receiver/s should send a reply to the current Presenter/Service Manager to become the new Presenter and this should choose one of them and inform all the affected peers about the structure update.

Another important issue to take into account is that P2P structures may suffer from link congestion or temporary link outages. Such problems in the network require measures to overcome the problems, as well as some error control to provide a reasonable quality during these periods of time. Retransmission requests would require an additional transmission delay not tolerable for a group communication service, since the system already imposes almost the highest acceptable delay for the service. For that reason, the use of Forward Error Correction codes (FEC) will be considered in the following as the sole error control method.

3. SVC RATE-SHAPING FOR LOW DELAY

As a consequence of congestion in the network, packet losses may occur, as well as late arrival of the packets, which has a big impact on the quality of the service for a low delay media delivery application. In order to tackle the problem, a fast response should be taken, adapting the sending rate to ensure that at least the most important packets are received.
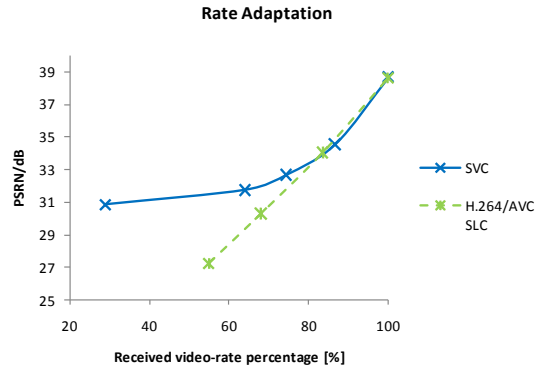

Figure 2: Example of the quality for rate adaptation based on temporal scalability for H.264/AVC SLC and SVC

Due to the strict delay constraints which have to be fulfilled, additional restrictions exist on how the FEC is to be applied, since Source Blocks (SB) over multiple frames would result in an additional decoding delay. P2P architectures impose an already high transmission delay, due to the latency of the system caused by the intermediate relay nodes. To minimize any further delays, it may only be possible to apply FEC on a per video frame basis, which means that the SB size is equal to a frame. As a consequence, the effectiveness of this error control methods is reduced and a fast response to problems in the system is required, since FEC codes may not be enough to overcome these problems.

This work proposes a rate-shaping based on the importance of the packets in order to overcome periods under congestion. The video delivery can clearly be improved by the use of SVC, adapting the sent video content by sending the packets from lower layers first and then the rest if possible. The SVC NAL unit header's priority indicator field [5] can be used to perform this packet prioritization. In this case, a lower quality or spatial representation of the original data could be guaranteed, whereas in case of encoding the data with H.264/AVC SLC, packet drops result in a representation of the data at a lower frame rate. Furthermore, rate adaptation based on the temporal scalability feature of H.264/AVC SLC is less efficient than spatial or SNR scalability, in terms of R/D values, as depicted in Figure 2. The first point for SVC in Figure 2, corresponding to a 25 % of the video rate, represents the base layer. The additional operation points are obtained adding temporal layers of the enhancement layer. For the PSNR calculation of the H.264/AVC SLC adaptation based on temporal scalability freeze frame concealment is considered for the failing frames. Exemplarily, the operation point with lowest video rate for single layer coding corresponds to a frame rate of about 7 fps, while with SVC a higher rate reduction can be obtained having a video representation with still 30 fps, considered to be much better.

A congestion state comprises two different phases: the *detection phase* and the *rate-adaptation phase*. For the first phase, random losses and late arrivals are considered depending on the degree of congestion. In other words, depending on this degree of congestion, a rate reduction is selected and a given amount of packets, which corresponds to this reduction in the throughput, are discarded due to packet loss or because their play-out time has already passed. After the *detection phase*, i.e. time in which packet loss or late packet arrival is detected by the sender, the *rate-adaptation phase* takes place until the congestion is over. The congestion detection is performed based on acknowledgement

packets (ACK) and Round Trip Time (RTT) calculation at the side of the sender, as shown in Figure 3. Failing ACKs and big increases in the RTT will indicate congestion in the network, and senders should start the *rate-adaptation phase*. During this phase the sender adapts the sent media content based on the packet importance and the estimated available throughput, and the mentioned rate-shaping is performed.
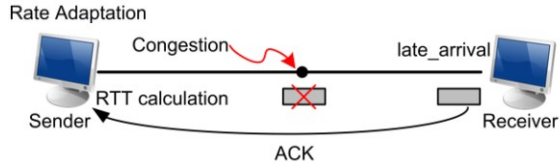

Figure 3: Rate Adaptation and congestion detection

4. SIMULATION RESULTS

The aforementioned system has been simulated in the network simulator NS-2. Two different simulations have been carried out. First, the media delivery with single layer coding has been simulated, where the additional available throughput as result of the gain in efficiency compared to SVC has been applied for FEC, and second, the media delivery with SVC, where only the base layer has been protected. The media streams of about 30 seconds are a concatenation of the test sequences Carphone, Foreman and Mother&Daughter at 30 fps using low delay SVC coding (Scalable Baseline Profile, JSVM9.17) and low delay H.264/AVC SLC coding (Baseline Profile, JSVM9.17). The video characteristics and FEC protection rates are summarized in Table 1 and Table 2 respectively. The selected code rates for the FEC correspond to a whole data rate of 562 and 615 kbps. For the SVC stream, a 10 % and 20 % of redundancy is selected to reach these rates. However in case of H.264/AVC SLC, over-provisioning values of 20 % and 30 % have been simulated respectively, due to the higher coding efficiency than SVC, i.e. SVC introduces an overhead of 11.3 % in the encoded sequences.

	Video rate [kbps]	avg. PSNR [dB]
H.264/AVC SLC	459.90	38.62
SVC- base layer	143.85	30.86
SVC- base + enh. layer	511.75	38.64

Table 1: Characteristic of media streams

Forward Error Correction (FEC)	
FEC code	Raptor
Source block (SB) length	SB/frame(33.3 ms @ 30 fps)
Code rate base layer (SVC)	0.82 (10 %) and 0.66 (20 %)
Code rate enh. layer (SVC)	1
Code rate (H.264/AVC SLC)	0.85 (20 %) and 0.78 (30 %)

Table 2: Forward Error Codes, selected protection

An interactive group communication comprising 100 participants has been simulated. Each participant has a 2 Mbit/s uplink throughput and a RTT of 50 ms to each other participant, which meets the constraint that the whole transmission delay from the presenter to any other peer is kept under 150 ms, for a system with 10 trees following eq. (1).
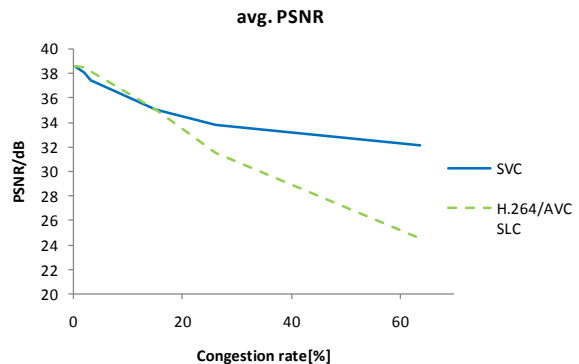

Figure 4: avg. PSNR vs. congestion rate for 10 % over-provisioning for SVC and 20 % for H.264/AVC SLC

The congestion was simulated over all links, which are considered independent. During the simulation some links suffer from congestion with a given probability over the time. This probability follows a geometric distribution with a simulation time step of 100 ms, which means that every 100 ms a link may change from normal state to congested state with a given probability. Once in congested state, they stay in this state between 2 and 5 seconds following a uniform distribution.

The results presented from Figure 4 to Figure 7 show the quality of the received video content for different congestion rates r_c, i.e. time in congestion state over the whole time, as an average value over all the receivers. The simulated congestion rate values are r_c = [0.175 %, 0.3 %, 1.75 %, 3 %, 15 % 26 %, 63.6 %], which correspond to the probabilities of changing to congestion state of [0.005 %, 0.01 %, 0.05 %, 0.1 %, 0.5 %, 1 %, 5 %]. The relationship between congestion rate and this probability is described in eq. (2) from the expected values of a uniform and a geometric distribution.

$$r_c = \frac{E\{t_{congestion}\}}{E\{t_{congestion}\} + E\{t_{nocongestion}\}} \qquad (2)$$

During these congestion states, a rate reduction is simulated, where the received data rate is randomly selected from a uniform distribution between 40 % and 90 % of the nominal video+FEC rate. Depending on this rate, the random losses and rate adaptation at each affected forwarding peer, described in Section 3, are carried out.

In Figure 4 and Figure 5 the results are depicted for a data rate of 562 kbps, which corresponds to an over-provisioning value of 10 % for SVC and 20 % for H.264/AVC SLC. For congestion rate values above 15 %, the rate adaptation in combination with SVC shows an improvement in PSNR when compared to H.264/AVC SLC, as shown in Figure 4. For smaller values, there is a slight difference between both methods, where the PSNR is a little bit higher for H.264/AVC SLC. Conversely, higher values of congestion rate show a bigger difference, where SVC outperforms H.264/AVC SLC by up to 8 dB in PSNR.

This enhancement is even more noticeable when looking at the average frame rate as shown in Figure 5. The performance of the proposed rate adaptation method in combination with SVC is

noticeably a better solution. It can be clearly seen how SVC keeps the average frame rate almost at the highest value, considered really important for low-delay interactive communication applications, whereas the approach with H.264/AVC SLC performs quite poorly with a big reduction in the continuity of the video content.
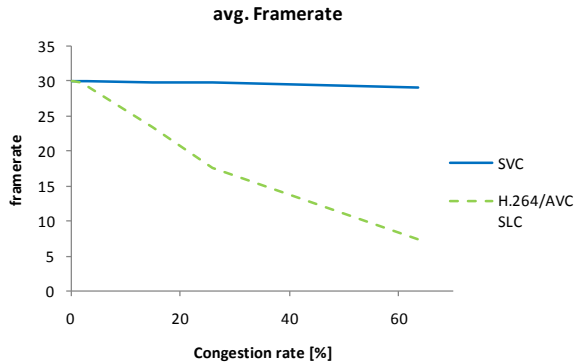
avg. Framerate

Figure 5: avg. frame rate vs. congestion rate for 10 % over-provisioning for SVC and 20 % for H.264/AVC SLC

The results for an over-provisioning value of 20 % for SVC and 30 % for H.264/AVC SLC, which correspond to a data rate of 615 kbps, are shown in Figure 6 and Figure 7. An increase in the range of values of congestion rates, where H.264/AVC SLC has a higher PSNR than SVC is noticeable, as well as an increase in this difference, which can reach up to 2 dB, as shown in Figure 6. However, if focusing on the frame rate in Figure 7 the rate adaptation based on SVC again clearly outperforms the H.264/AVC SLC approach.
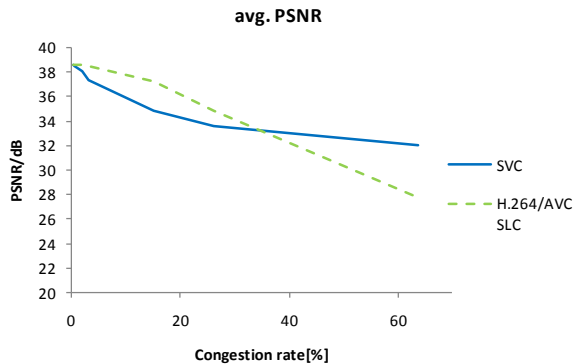
avg. PSNR

Figure 6: avg. PSNR vs. congestion rate for 20 % over-provisioning for SVC and 30 % for H.264/AVC SLC

5. CONCLUSION

The video delivery using P2P overlays in combination with SVC outperforms the video transmission with H.264/AVC SLC for group communication. A P2P Application Layer Multicast delivery system, which fulfills the delay constraints for interactive communication, is presented showing the benefits of performing a rate adaptation for group communication services in combination with SVC.

Since the P2P structure itself imposes a high transmission delay, the effectiveness of error correction methods is limited and,

therefore, a rate adaptation has to be performed in periods of time under congestion. In spite of the additional restrictions and constraints for low-delay communication services, P2P group communication shows to be a compelling and cost-attractive approach, where combination with SVC media delivery ensures a quality of service. The significant quality enhancement obtained by the transmission of SVC in comparison to H.264/AVC single-layer coding, in terms of PSNR and frame rate, encourages the implementation of the proposed system with SVC.
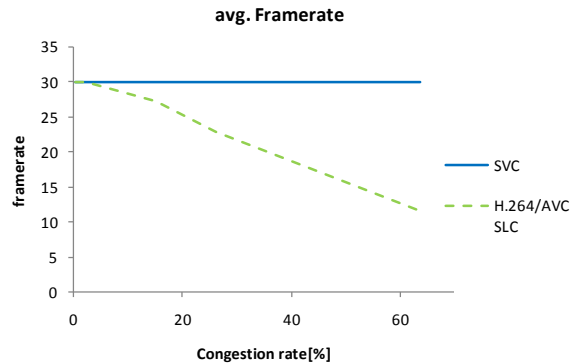
avg. Framerate

Figure 7: avg. framerate vs. congestion rate for 20 % over-provisioning for SVC and 30 % for H.264/AVC SLC

6. REFERENCES

[1] J. Liu, S.G. Rao, B. Li, and H. Zhang, "Opportunities and Challenges of Peer-to-Peer Internet Video Broadcast", in proceedings of the IEEE, vil. 96, no. 1, pp 11-24, January 2008.

[2] V. N. Padmanabhan, H. J. Wang and P.A. Chou, "Resilient peer-to-peer streaming", in ICNP, 2003.

[3] E. Setton, P. Baccichet, and B. Girod, "Peer-to-Peer Live Multicast: A Video Perspective", in Proc. of IEEE, 2008.

[4] A. Eleftheriadis, M. R. Civanlar, and O. Shapiro, "Multipoint videoconferencing with scalable video coding," J. Zhejiang Univ. Sci. A, vol.7, no. 5, pp. 696–705, May 2006.

[5] H. Schwarz, D. Marpe, and T. Wiegand, „Overview of the scalable video coding extension of the H.264/AVC standard", IEEE Trans. Circuits Syst. Video Technol, vol. 17, no. 9, Sep. 2007.

[6] M. Castro, P. Druschel, A. M. Kermarrec, A.R.A Nandi, and A. Singh, "SplitStreaming High-bandwidth content distribution in a cooperative environment", in Proceedings of ACM SOSP, 2003.

[7] X. Zhang, J. Liu, B. Li and T.-S. P. Yum, "DONet/CoolStreaming: A data-driven overlay network for live media streaming", in Proc. INFOCOM'05, Miami, FL, USA, March 2005.

[8] M. Zhang, Q. Zhang, L. Sun, S. Yang, „Understanding the Power if Pull-Base Streaming Protocol: Can We Do Better?", in IEEE Journal on Selected Areas in Communications 25(9): 1678-1694 (2007).

[9] J. Noh, A. Mavlankar, P.Baccichet, and B. Girod, "Reducing End-to-End Transmission Delay in P2P streaming systems Using Multiple Trees with Moderate Outdegree", In Proc. International Conference on Multimedia and Expo (ICME) 2008