FEASIBILITY OF MULTI-VIEW VIDEO STREAMING OVER P2P NETWORKS

Engin Kurutepe, Thomas Sikora

Communication Systems Group, TU Berlin,
EN-1, Einsteinufer 17, 10587 Berlin, Germany
{kurutepe,sikora}@nue.tu-berlin.de

ABSTRACT

We propose to stream multi-view video over a multi-tree peer-to-peer (P2P) network using the NUEPMuT protocol. Each view of the multi-view video is streamed over an independent P2P streaming tree and each peer only contributes upload capacity in a single tree, in order to limit the adverse effects of ungraceful peer departures. Additionally, we investigate the feasibility of using the proposed P2P networking architecture, NUEPMuT, for the streaming of multi-view video content with the currently available Internet connection bandwidths.

Index Terms— Multi-View Video, Peer-to-peer networks, Video Streaming

1. INTRODUCTION

There is constantly growing interest towards 3-D entertainment applications from both content providers and end users. It is being expected that about 50% of all theaters will be 3-D capable by the end of 2009. Given the interest in 3-D, it will be unsurprising to see 3-D content migrate from theaters to home entertainment systems, just like movies migrated from movie theaters to our TV-sets. Therefore, there is a need for new technologies, which would enable transmission of new 3-D content to end users. In line with the convergence of telephony and TV communications to IP-networks, the authors expect IP-only delivery of 3-D content in the near future.

Multi-view representations provide a good compromise between the amount of data and rendering complexity and they can be compressed using simulcast encoding, where each view is encoded as an independent conventional video stream. However, more efficient compression can be achieved by exploiting the spatial redundancy between cameras, in addition to the temporal redundancy in a camera stream [1]. This approach is currently being standardized under JVT as an amendment to H.264/AVC [2]. Even after state of the art compression, multi-view representations are very data intensive: 38dB PSNR at about 5 Mbps is a common operating point for a 704×480, 30fps, 8 camera sequence with MVC encoding. High bit-rate of multi-view representations is an important challenge for both the content providers and end users. At the end user side, there has been work on selective streaming [3], such that only the necessary parts of the multi-view video are streamed according to the viewer's head position. This kind of selective streaming work would benefit greatly from scalability features of the MVV stream which have been proposed in [4] and [5]. On the other hand the problem at the server-side is that given a fixed bandwidth the number of viewers which can be served simultaneously is reduced drastically due to much increased bitrate.

There has been a wealth of research on Peer-to-Peer (P2P) video streaming, and it has been shown that P2P architectures can greatly increase the number of simultaneously served viewers. The underlying idea beneath the P2P architectures is that the peers share their disk and bandwidth resources with the P2P network by intelligently forwarding the packets they receive, to other peers. This allows for the bandwidth load to be lifted from the server and distributed among peers. Although P2P streaming has been shown to be feasible [6] and cost-effective [7], there are two important intrinsic issues facing any P2P approach: Reliability of end systems and upload/download bandwidth asymmetry for most peers. Reliability of end systems is lower than dedicated routers and when an end system ungracefully leaves the network, all its children are starved until a new parent is found. Upload/Download bandwidth asymmetry affects how the total capacity of the P2P network grows as new peers join the network. As many of the domestic cable and DSL connections offer less bandwidth for upload than for download, many peers can lack the upload bandwidth to contribute into the network, even if they have enough download bandwidth to watch the content, thus have no choice but becoming free-riders in the network. It has been shown [8] [9] [10] that multiple description coding (MDC) can help for both issues. MDC involves dividing a video stream into several independent descriptions, which are then streamed over different paths in a P2P setting, making use of path diversity to ensure at least some of the descriptions arrive at their destinations in the case of peer departures. Additionally, since each of these descriptions has a lower bitrate in comparison to the original stream, a peer which has not enough upload bandwidth for the

This work is supported by EC within FP6 by 3DTV Network of Excellence under Grant 511568.

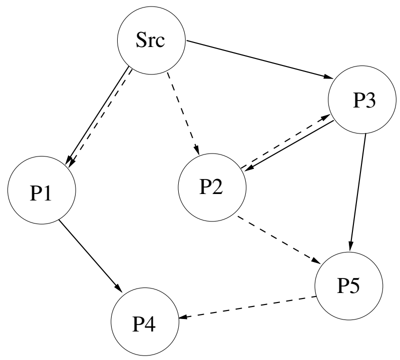

Fig. 1. A simple multi-tree network with two independent trees, where each peer contributes capacity to a single tree.

original stream, could contribute capacity in the P2P network for some of the MDC descriptions.

The parallelism between multi-view video and multiple description coding of conventional video is noteworthy. Whereas separate descriptions are streamed over separate trees in P2P streaming of conventional video, we propose to consider each view in an MVC encoded multi-view video as a separate description and stream each view over an independent P2P streaming tree to get similar benefits as in conventional video streaming. In this paper, we will investigate how feasible such a multi-tree approach is on currently available Internet connection bandwidths, and present results of simulations done with real-life bandwidth distributions.

The rest of this paper is organized as follows: In Section 2 we describe the details of the proposed approach. Section 3 defines the experimental setup and presents our results. And finally in Section 4 we list our conclusion and outlook for the future.

2. SYSTEM DESCRIPTION

In this section, we describe a simple multi-tree P2P live streaming protocol which was used to examine how P2P streaming works for multi-view videos. But before we proceed to the details of our protocol, the interaction between different multi-view encoding schemes and multi-tree streaming should be discussed properly.

Simulcast encoding regards each view as an independent video stream and encodes using H.264/AVC, obviously the effects of packet losses in simulcast encoded streams are contained within the same stream and do not propagate to neighboring views. Additionally, we also consider two flavors of multi-view coding (MVC). First one with complex spatial prediction structure where there are spatial references between all neighboring frames, and the second simplified prediction structure, where there are only spatial references between frames at anchor time instances at the beginning of each group of pictures (GOP). It was shown in [11] that this simplified prediction structure enables much less complex decoding and simplified random access within the stream with a negligible loss of compression efficiency. Unlike simulcast encoding, the effects of packet losses are likely to affect neighboring views in both MVC prediction schemes, but only if an anchor frame is lost in the simplified spatial prediction scheme, i.e. the effects of losses of B-frames between anchor frames are contained within the same view, whereas in the complex prediction structure any packet loss affects the quality of other views. The resilience against packet losses in other views is a very important property for the proposed system, which streams different views over independent P2P trees.

2.1. The P2P protocol: NUEPMuT

NUEPMuT stands for NUE (Abbreviation of Communication Systems Group in German) P2P Multi-Tree. The tree management is currently handled by the source node, since the processing and bandwidth overheads are negligible compared to the load of multi-view streaming.

2.1.1. Join Process

A new peer joins the network using the six-way handshake process: first a JOIN packet containing the total maximum capacity the joining peer is willing to contribute to the network is sent to the source node. The source replies for each tree with a list of fertile peers and indicates if capacity is requested in that particular tree. The source allocates the total capacity of the new peer to the tree, which currently has the minimum total available capacity. Therefore, each peer is fertile only in a single tree and acts as a free-rider in the other trees as seen Fig 1. This makes sure that an ungraceful peer departure disrupts only a single tree and has been shown to increase stability [6], reducing the number of lost packets and quality impact. Then the new peers sends PING packets to each peer in the lists provided by the source, which in turn reply with their depth in the tree and current available capacity. Finally, the new peer sends ATTACH packets to the peers with minimum tree depth in each tree, where the peer with the lowest RTT measurement is preferred between peers with the same depth. Due to the independency of the trees, for each tree, the selection is always made among disjoint sets of potential parents. If the ATTACH request is rejected, the new peer selects the next suitable candidate for that tree, or retries to join that tree again if no other candidates are available. An ATTACH acknowledgement also contains the list of the ancestors, which is used to prevent loops.

2.1.2. Tree Management

A peer must constantly check its parent and children to detect ungraceful departures. NUEPMuT has a heartbeat frequency, with which each node sends out a small packet to its parent. Therefore, a parent can detect dead children which do not send their heartbeat packets. Conversely, a parent replies to each heartbeat packet it receives from its children. A parent which does not reply to a heartbeat packet is assumed to be dead and triggers a new join process. During these rejoins, in order to prevent loops in the tree, attach requests are denied if a rejoining peers is in the ancestor list of a potential parent.

Additionally, peers inform the source when they don't have available capacity left due to a new join, or when capacity becomes available due to graceful or ungraceful departures. This information is used in the join process to select the tree with minimum total available capacity and to maintain a list of fertile peers.

2.2. Feasibility Analysis

In order for multi-tree P2P streaming of multi-view videos to be possible, the total amount of upload capacity in the P2P network should be larger than the total bit-rate of all streaming requests:

$$\sum_{<N>} c_i \geq \sum_{<N>} r_i, \quad (1)$$

where c_i and r_i are the provided bandwidth and requested bit-rate for the ith node, respectively. Assuming each peer requests to stream the complete multi-view representation, i.e. no selective streaming is performed, r_i is equal to the bit-rate for the MVV bitstream for all nodes. So by dividing both sides of the Eq. 1 by N, the total count of nodes, we get:

$$\frac{\sum_{<N>} c_i}{N} \geq R_{MVV}, \quad (2)$$

where R_{MVV} is the total bit-rate of the MVV bitstream. Eq. 2 says in plain words that the average opload bandwidth of the peers in the network should be greater than or equal to the bit-rate of the MVV bitstream. Otherwise, either the quality of service suffers for all peers, when all new peers are admitted, or only as many peers are admitted as the available capacity can support, in order to maintain the quality of service for the existing peers in the network.

3. RESULTS

In this paper, we present the initial results of our feasibility study. The P2P streaming protocol described in this paper was implemented as a new protocol in the NS-2 network simulator. We have used a moderately large transit-stub topology of 690 network nodes and realistically modeled link delays. 213 P2P agents were randomly placed on these 690 nodes. In order to to accurately simulate the bottle-neck quantity in such a P2P streaming system, the distribution of the up-link bandwidth available to these P2P agents, was modeled as described in [6]. Sripanidkulchai et al. measured the actual upload bandwidths of computers connected to the internet using variety of methods. Although, we employed the measurements published by them in our simulations, it should also be noted that [6], although the most recent and almost unique work of its kind, was published in 2004 and in parallel with Moore's law the network bandwidth must have increased in the mean time as well. In all simulations the multi-view video server is assumed to have a relatively low outbound bandwidth for a server at 5Mbps in order to emphasize the distribution of the bandwidth load on the peers. At 1Mbps multi-view bitrate, this corresponds to only 5 simultaneous viewers in a unicast streaming environment, whereas the proposed system was able to consistently support around 25 peers in our simulations with a 5Mbps server and 1Mbps multi-view content.

We have modeled the arrival and departure of new peers to the overlay network as exponential random variables, with mean inter-event times of μ_J and μ_D respectively. In the case, there is no enough capacity available in the overlay network, a new peer is rejected, in order to preserve a quality of service for the peers already in the network.

Table 1. AAFC with respect to peer churn

		μ_J(sec)		
		1.0	**2.0**	**5.0**
μ_D(sec)	**60**	0.610	1.05	2.84
	120	0.205	0.718	1.10

We have investigated the average available free capacity (AAFC) in the overlay network. This metric is averaged over time and over overlay trees and denotes average number of additional outbound streams a node could support at any given time. Obviously, for a node which is uploading at maximum capacity available capacity is zero. However, since the AAFC metric is averaged over time and all peers, it is always positive. But, the lower the AAFC metric, the more new join requests are likely to be rejected due to the lack of free capacity in the overlay network. Table 1 shows how AAFC metric changes with different join and departure behaviors of the peers, also known as churn, μ_J and μ_D are given in seconds and four streams were streamed over four independent P2P delivery trees.

Additionally, we have also investigated how the available capacity in the overlay network changes as the number of views in the multi-view representation and the number of P2P trees increase. First, we assumed that the quality and therefore the bitrate of each view stays constant at 250kbps, which means that the total bandwidth and load on the overlay network increase linearly as the number of views increase. Table 2 shows how the AAFC metric drops drastically with the increasing load of 4, 8 and 16 views. This means that the number of peers not allowed into the overlay tree also increase drastically with the network load.

However, if the total bitrate of the representation is kept

Table 2. AAFC in the P2P network with respect to number of views (Fixed bitrate per view)

	# of views		
	4	**8**	**16**
AAFC	1.10	0.321	0.104

constant while the number of views in the representation is increased, the decoded video quality is bound to decrease. However, we leave the effects of quality decrease outside the scope of this paper. On the other hand, this increase has relatively little effect on the AAFC metric of the overlay network, as it can be seen from Table 3, where the total bitrate was fixed at 1Mbps.

Table 3. AAFC in the P2P network with respect to number of views (Fixed total bitrate)

	# of views		
	4	**8**	**16**
AAFC	1.10	1.42	1.29

4. CONCLUSIONS

In this paper, we have presented a novel P2P streaming system for multi-view videos and investigated its feasibility with the help of the average available free capacity (AAFC) metric. Although, we have fixed the outbound bandwidth of the multi-view video server at a low value, we could still observe that the streaming of multi-view content was not only possible but also appears to be feasible, due to the largely increased number of possible viewers, which can be supported with a given server bandwidth.

We consider these presented initial results as a worst case analysis. The NUEPMuT protocol is still under development, has much room for improvement and needs to be investigated further. We also think that, even if very thorough and complete, which the results presented in this paper are definitely not, the presentation of only protocol level results does not provide enough information on the real performance of a streaming system and definitely need to be extended with objective and subjective quality tests, which will be performed and their results presented in future publications.

5. REFERENCES

[1] Karsten Mueller, Philipp Merkle, Heiko Schwarz, Tobias Hinz, Aljoscha Smolic, Tobias Oelbaum, and Thomas Wiegand, "Multi-view video coding based on H.264/AVC using hierarchical B-frames," in *Picture Coding Symposium 2006*. PCS, 2006.

[2] A. Vetro, P. Pandit, H. Kimata, and A. Smolic, "Joint Draft 3.0 on Multiview Video Coding," *Joint Video Team, Doc. JVT-W209*, 2007.

[3] Engin Kurutepe, M. Reha Civanlar, and A. Murat Tekalp, "Client-driven selective streaming of multi-view video for interactive 3DTV," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1558 – 1565, November 2007.

[4] Nukhet Ozbek and Murat Tekalp, "Scalable multi-view video coding for interactive 3DTV," in *Proceedings of IEEE ICME 2006*. IEEE, 2006, pp. 213–216.

[5] Michael Drose, Carsten Clemens, and Thomas Sikora, "Extending single-view scalable video coding to multi-view based on h.264/avc," in *Proceedings of IEEE ICIP 2006*. IEEE, September 2006, pp. 2977–2980.

[6] K. Sripanidkulchai, A. Ganjam, B. Maggs, and H. Zhang, "The feasibility of supporting large-scale live streaming applications with dynamic application end-points," *Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, pp. 107–120, 2004.

[7] Eric Setton and John Apostolopoulos, "Towards quality of service for peer-to-peer video multicast," *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 5, pp. 81 – 84, Sept. 16 2007-Oct. 19 2007.

[8] M. Castro, P. Druschel, A.M. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, "SplitStream: high-bandwidth multicast in cooperative environments," *Proceedings of the nineteenth ACM symposium on Operating systems principles*, pp. 298–313, 2003.

[9] E. Setton, P. Baccichet, and B. Girod, "Peer-to-peer live multicast: A video perspective," *Proceedings of the IEEE*, vol. 96, no. 1, pp. 25–38, Jan. 2008.

[10] E. Akyol, A. M. Tekalp, and M. R. Civanlar, "Adaptive Peer-To-Peer Video Streaming with Optimized Flexible Multiple Description Coding," *Image Processing, 2006 IEEE International Conference on*, pp. 725–728, 2006.

[11] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "Efficient Prediction Structures for Multiview Video Coding," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 11, pp. 1461–1473, 2007.