

Scalable Video Coding in Content-Aware Networks: Research Challenges and Open Issues

Christian Timmerer¹, Michael Grafl¹, Hermann Hellwagner¹, Daniel Negru², Eugen Borcoci³, Daniele Renzi⁴, Anne-Lore Mevel⁵, and Alex Chernilov⁶

¹Klagenfurt University, Universitätsstrasse 65-67, A-9020 Klagenfurt, Austria, {firstname.lastname}@itec.uni-klu.ac.at

²CNRS LaBRI Lab. – University of Bordeaux 1, 351 Cours de la Libération, 33400 Talence, France, daniel.negru@labri.fr

³University "Politehnica" of Bucharest (UPB), 1-3, Iuliu Maniu Ave., 061071, Bucharest 6, Romania, Eugen.Borcoci@elcom.pub.ro

⁴bSoft ltd, 156 via Velini, 62100 Macerata, Italy, daniele@bsoft.info

⁵Thomson Grass Valley France, 40, rue de Bray, 35510 Cesson-Sevigné, France, Anne-lore.Mevel@grassvalley.com

⁶Optibase, 7 Shenkar St., P.O.Box 2170, Herzlia, 46120 Israel, Alexc@optibase.com

Abstract The demand for access to advanced, distributed media resources is nowadays omnipresent due to the availability of Internet connectivity almost anywhere, anytime, and with a huge amount of different devices. This calls for rethinking of the current Internet architecture by making the network aware of which content is actually transported. This paper introduces Scalable Video Coding (SVC) as a tool for Content-Aware Networks (CANs) which is currently researched as part of the EU FP7 ALICANTE project. The architecture of ALICANTE with respect to SVC and CAN is reviewed, use cases are described, and, finally, research challenges and open issues are discussed.

1 Introduction

In recent years the number of contents, devices, users, and means to communicate over the Internet has grown rapidly and with that the heterogeneity of all the involved entities. Many issues can be associated with that which are generally referred to as ongoing

research in the area of the Future Internet (FI) [1]. One project in this area is the European research FP7 Integrated Project “Media Ecosystem Deployment Through Ubiquitous Content-Aware Network Environments” (ALICANTE) [2] which proposes a novel concept towards the deployment of a new networked *Media Ecosystem*. The proposed solution is based on a flexible cooperation between providers, operators and end-users, finally enabling every user (1) to access the offered multimedia services in various contexts, and (2) to share and deliver her/his own audiovisual content dynamically, seamlessly, and transparently to other users.

Towards this goal, ALICANTE’s advanced concept provides *content-awareness* to the network environment, *context-awareness* (network/user) to the service environment, and *adapted services/content* to the end-user for her/his best service experience possible, taking the role of a consumer and/or producer. By *environment*, it is understood a generic and comprehensive name to emphasize a grouping of functions defined around the same functional goal and possibly spanning, vertically, one or more several architectural (sub-)layers. This name is used to characterize its broader scope with respect to the term *layer*.

The ALICANTE architecture introduces two novel virtual layers on top of the traditional network layer, i.e., a Content-Aware Network layer (CAN) for network packet processing and a Home-Box (HB) layer for the actual content adaptation and delivery. Furthermore, Scalable Video Coding (SVC) is heavily employed for the efficient, bandwidth-saving delivery of media resources across heterogeneous environments (cf. Section 2). Technical use cases that will benefit from this architecture are outlined in Section 3 and Section 4 details the research challenges and open issues to be addressed in the course of the project. Finally, the paper is concluded in Section 5.

2 ALICANTE: MediA Ecosystem Deployment Through Ubiquitous Content-Aware Network Environments

2.1 Overview and System Architecture

The ALICANTE architecture promotes advanced concepts such as content-awareness to the network environment, user context-awareness to the service environment, and adapted services/content to the end-user for his/her best service experience while being a consumer and/or producer.

Two novel virtual layers are proposed on top of the traditional network layer as depicted in Figure 1: the *Content-Aware Network (CAN) layer* for network packet processing and a *Home-Box (HB) layer* for the actual content adaptation and delivery.

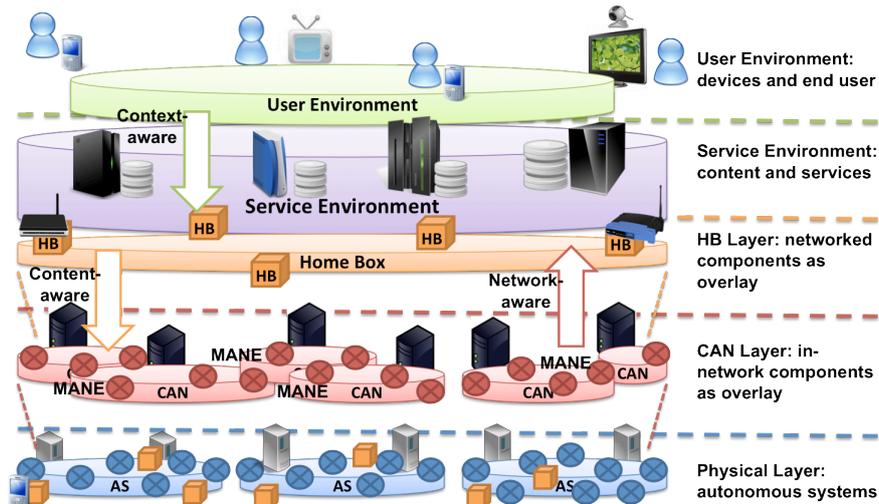


Figure 1. ALICANTE concept and system architecture.

Innovative components instantiating the CAN are called *Media-Aware Network Elements (MANE)*. They are actually CAN-enabled routers and associated managers, offering together content-aware and context-aware Quality of Service/Experience, security, and monitoring features, in cooperation with the other elements of the ecosystem.

The upper layer, i.e., the *Service Environment*, uses information delivered by the CAN layer and enforces network-aware application

procedures, in addition to user context-aware ones. The novel proposed *Home-Box (HB)* entity is a physical and logical entity located at end-users' premises and gathering context, content, and network information essential for realizing the big picture. Associated with the architecture there exists an open, metadata-driven, interoperable middleware for the adaptation of advanced, distributed media resources to the users' preferences and heterogeneous contexts enabling an improved Quality of Experience. The adaptation will be deployed at both the HB and CAN layers making use of scalable media resources as outlined in the next section.

For more detailed information the interested user is referred to [3].

2.2 Scalable Video Coding and Content-Aware Networks

The adaptation relies on Scalable Video Coding (SVC) [4]. SVC follows a layered coding scheme comprising a base layer and one or more enhancement layers with various dimensions. Three basic scalable coding modes are supported, namely spatial scalability, temporal scalability, and Signal to Noise Ratio (SNR) scalability, which can be combined into a single coded bit stream:

- **Spatial (picture size) scalability.** A video is encoded at multiple spatial resolutions. By exploiting the correlation between different representations of the same content with different spatial resolutions, the data and decoded samples of lower resolutions can be used to predict data or samples of higher resolutions in order to reduce the bit rate to code the higher resolutions.
- **Temporal (frame rate) scalability.** The motion compensation dependencies are structured so that complete pictures (i.e., their associated packets) can be dropped from the bit stream. Note that temporal scalability is already enabled by AVC and that SVC has only provided supplemental enhancement information to improve its usage.
- **SNR/Quality/Fidelity scalability.** A video is encoded at a single spatial resolution but at different qualities. The data and decoded samples of lower qualities can be used to predict data or

samples of higher qualities in order to reduce the bit rate to code the higher qualities.

The adaptation deployed at the CAN layer will be performed in a Media-Aware Network Element (MANE) [5]. MANEs, which receive feedback messages about the terminal capabilities and channel conditions, can remove the non-required parts from a scalable bit stream before forwarding it. Thus, the loss of important transmission units due to congestion can be avoided and the overall error resilience of the video transmission service can be substantially improved.

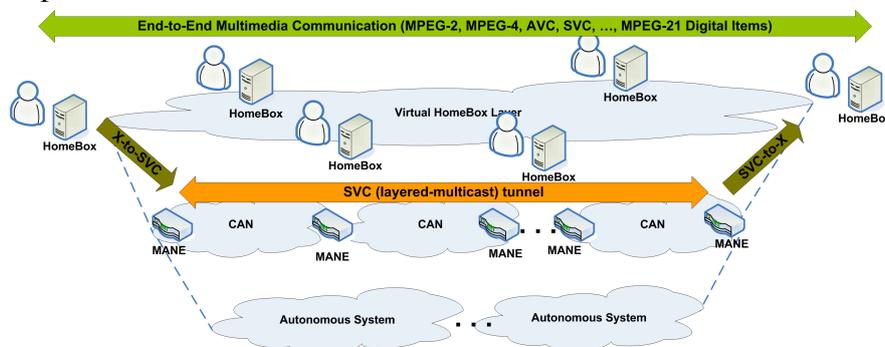


Figure 2. Concept of SVC (layered-multicast) tunnel.

Design options of in-network adaptation of SVC have been described in previous work [6] and first measurements of SVC-based adaptation in an off-the-shelf WiFi router have been reported in [7]. More complex adaptation operations that will be required to create scalable media resources, such as transcoding [8] of media resources which have increased memory or CPU requirements, will be performed at the edge nodes only, i.e., in the Home-Boxes. Therefore, the ALICANTE project will develop an SVC (layered-multicast) tunnel, as depicted in Figure 2, inspired by IPv6 over IPv4 tunnels. That is, within the CAN layer only scalable media resources – such as SVC – are delivered adopting a layered-multicast approach [9] which allows the adaptation of scalable media resources by the MANEs implementing the concept of distributed adaptation. At the border to the user, i.e., the Home-Box, adaptation modules are deployed enabling device-independent access to the SVC-encoded content by providing X-to-SVC and SVC-to-X transcoding/rewriting

functions with $X=\{\text{MPEG-2, MPEG-4 Visual, MPEG-4 AVC, etc.}\}$. An advantage of this approach is the reduction of the load on the network (i.e., no duplicates), making it free for (other) data (e.g., more enhancement layers). However, multiple adaptations may introduce challenges that have not (yet) been addressed in their full complexity (cf. Section 4).

3 Use Cases

In order to evaluate the concept of SVC in the context of CANs/HBs, several use cases have been defined which are briefly introduced in the subsequent sections.

3.1 Multicast/Broadcast

In this scenario, multiple users are consuming the same content from a single provider (e.g., live transmission of sport events). The users may have different terminals with certain capabilities as depicted in Figure 3. The ALICANTE infrastructure is simplified in Figure 3 to highlight the interesting parts for this scenario (i.e., the HBs and the MANEs). Note that the SVC layers depicted in the figure are only examples and that SVC streams in ALICANTE may comprise temporal, spatial, and quality (SNR) scalability with multiple layers. The properties and numbers of SVC layers will be determined by the HB at the SP/CP side based on several parameters (e.g., diversity of terminal types, expected network fluctuations, size overhead for additional layers, available resources for SVC encoding/transcoding, etc.) which are known a priori or dynamically collected through a monitoring system operating across all network layers.

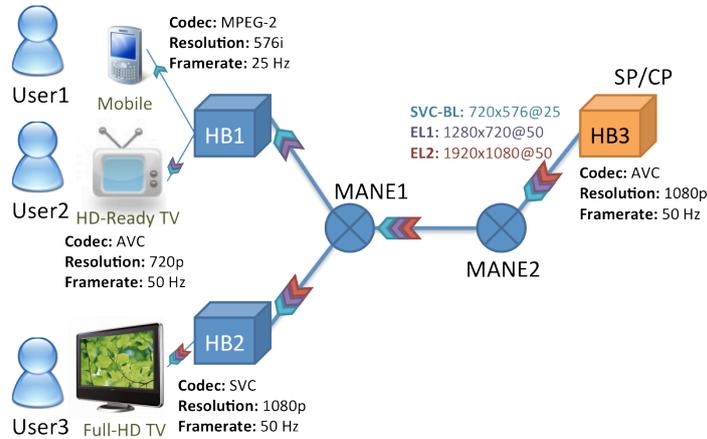


Figure 3. Multicast/broadcast with SVC adaptation.

3.2 Home-Box Sharing

In this scenario, a user consumes content through a foreign (shared) HB, e.g., the user accesses the content/service to which she/he has subscribed while being abroad (e.g., business trip, vacation). Figure 4 depicts a user consuming content at two different locations on two different terminals, connected to different HBs. Note that the user might as well use her/his mobile phone to consume content through HB2.

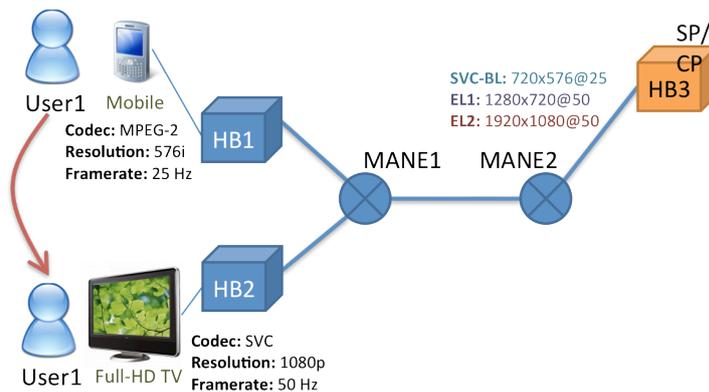


Figure 4. Home-Box sharing.

3.3 Video Conferencing

This scenario consists of an n:m video conferencing session (e.g., in family meetings, office meetings, etc.) as depicted in Figure 5. The media distribution is handled over a multicast shared bi-directional non-homogeneous tree in the ALICANTE network. In such a way only the minimum amount of network resources are spent, while assuring to the end user a maximum on quality.

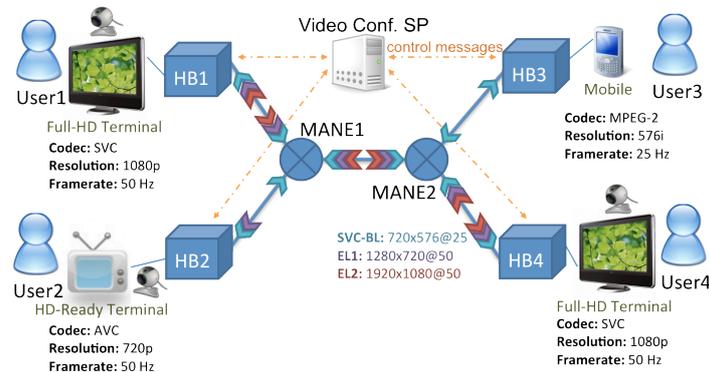


Figure 5. Video conferencing.

3.4 Peer-to-Peer Media Streaming

The HBs operate in peer-to-peer (P2P) mode within the ALICANTE ecosystem as illustrated in Figure 6. The MANEs, through which the P2P traffic flows, act as proxy caches which intercept requests for content pieces issued by HBs and aggregate them including the capabilities of requesting terminal. Furthermore, content pieces are only forwarded if the requesting terminals can decode them. Therefore, unnecessary traffic is reduced to a minimum freeing up the network resources for other data (e.g., additional enhancement layers).

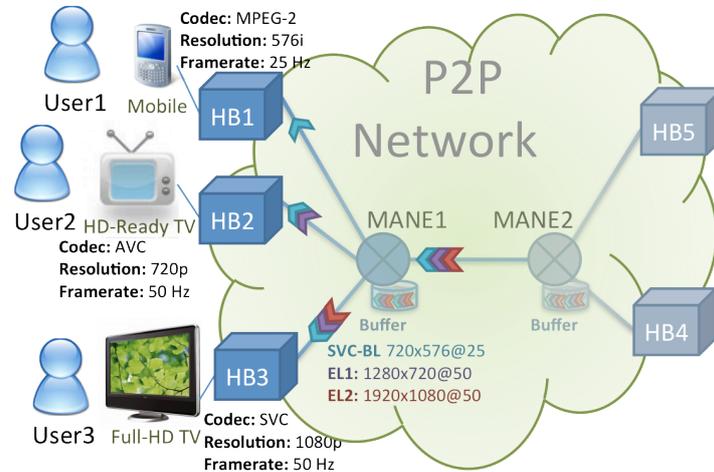


Figure 6. P2P media streaming.

4 Research Challenges and Open Issues

In this section we would like to point out some research challenges and open issues with respect to utilizing Scalable Video Coding within Content-Aware Networks.

Distributed adaptation decision-taking framework. Due to the fact that many, possibly heterogeneous entities are involved – in the production, ingestion, distribution and consumption stages – there is a need to develop a framework for distributed adaptation decision-taking. That is, finding the optimal decision regarding to the adaptation of the content for a single entity (i.e., HB, MANE) within a network of various entities in the delivery system. Note that the decision-taking is needed at the request and during the delivery of the multimedia content as (network) conditions might change.

Distributed adaptation at HB and CAN layers. The actual adaptation at both layers needs to be done efficiently, based on several criteria, in order to obtain low (end-to-end) delay, minimum quality degradation, and assuring scalability in terms of the number of sessions that can be handled in parallel.

Efficient, scalable SVC tunneling and signaling thereof. The approach of tunneling the content within SVC streams in the (core) network opens up a number of issues due to the SVC adaptation

within the MANEs, the SVC transcoding/rewriting within the HBs, and the associated signaling thereof. The issues range from efficiency and scalability to quality degradations.

The impact on the Quality of Service/Experience (QoS/QoE).

As there may be many adaptations happening during the delivery of the content, the impact on the Quality of Service/Experience needs to be studied in order to find the best trade-off for the use cases in questions. While for the QoS many objective measures are available, the QoE is highly subjective and requires tests involving end users; these tests are time consuming and costly. In any case, a good test-bed is needed for both objective and subjective tests for the evaluation of the QoS and QoE respectively. The possible mappings between QoS and QoE will be considered in this work also.

5 Conclusions and Future Work

In this paper we have introduced the usage of scalable video coding in content-aware networks for various use cases described in the paper. In particular, SVC is a promising tool for making the network aware of the actual content being delivered, i.e., when it comes to technical properties such as bit rate, frame rate, and spatial resolution. Furthermore, it allows for efficient and easy-to-use in-network adaptation due to the inherent structure of SVC.

The use cases described in the paper indicate the advantages of using SVC and in-network adaptation and we highlight research challenges and open issues. However, as this work is in its early stage it lacks of validation results for the scenarios and solutions proposed which remains part of our future work.

Acknowledgments This work is supported in part by the European Commission in the context of the ALICANTE project (FP7-ICT-248652). <http://www.ict-alicante.eu/>

References

1. Tselentis, G. (et.al.): Towards the Future Internet - Emerging Trends from European Research, IOSPress (2010)
2. ALICANTE Web site. <http://www.ict-alicante.eu/>. Accessed 21 June. 2010

3. Borcoci, E., Negru, D., Timmerer, C.: A Novel Architecture for Multimedia Distribution Based on Content-Aware Networking, In: Proc. The Third International Conference on Communication Theory, Reliability, and Quality of Service (CTRQ2010), Athens/Glyfada, Greece (2010)
4. Schwarz, H., Marpe, D., Wiegand, T.: Overview of the Scalable Video Coding Extension of the H.264/AVC Standard, *IEEE Transactions on Circuits and Systems for Video Technology*, 17(9), 1103–1120 (2007)
5. Wenger, S., Wang, Y.-K., Schierl, T.: Transport and Signaling of SVC in IP Networks, *IEEE Transactions on Circuits and Systems for Video Technology*, 17(9), 1164–1173 (2007)
6. Kuschnig, R., Kofler, I., Ransburg, M., Hellwagner, H.: Design options and comparison of in-network H.264/SVC adaptation, *Journal of Visual Communication and Image Representation*, 19(8), 529–542 (2008)
7. Kofler, I., Prangl, M., Kuschnig, R., Hellwagner, H.: An H.264/SVC-based adaptation proxy on a WiFi router, In: 18th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV2008), Braunschweig, Germany (2008)
8. Shen, B., Tan, W.-T., Huve, F.: Dynamic Video Transcoding in Mobile Environments, *IEEE MultiMedia*, 15(1), 42–51 (2008)
9. McCanne, S., Jacobson, V., Vetterli, M.: Receiver-driven layered multicast, In: SIGCOMM (1996)