

Combined Adaptation and Caching of MPEG-4 SVC in Streaming Scenarios

Michael Mackay, David Hutchison
Computing Department, InfoLab21
Lancaster University
LA1 4WA, Lancaster, United Kingdom
{m.mackay, dh}@comp.lancs.ac.uk

Michael Ransburg, Hermann Hellwagner
Klagenfurt University
Multimedia Communication Group
Universitätsstraße 65–67, Klagenfurt, Austria
{firstname.lastname}@itec.uni-klu.ac.at

Abstract

A key objective of the European Project ENTHRONE II is the ability to optimise the delivery of multimedia content to a wide group of heterogeneous users. One example of this is in the cooperative deployment of adaptation and caching functionality in the edge network. This hybrid approach makes it possible not only to store content locally, thus minimising the cost incurred through subsequent requests, but also to better serve heterogeneous groups of users by dynamically adapting the content to suit a wide range of terminal devices. In this paper, we describe and evaluate how the cooperative deployment of MPEG-21-based adaptation and caching of MPEG-4 SVC can result in improvements both in the quality of the content received at the user terminal and the resources consumed during the delivery.

1 Introduction

A key aim throughout the second phase of the ENTHRONE Project¹, has been to optimise the content delivery mechanism both to tailor the service received by the user and to minimise the resources consumed by the provider. This is a particular challenge in heterogeneous usage environments [1]. Several approaches to meet this challenge have been identified but of particular interest has been the potential to deploy these in a cooperative manner such that the properties of one complement those of the other.

This paper describes and evaluates the cooperative use of dynamic content adaptation and caching in edge networks. We believe that this combination has significant advantages not only to improve the Quality of Experience (QoE) for users, in terms of session initiation times and content suitability/robustness, but also to reduce the resources required during delivery for the provider through lower bandwidth consumption [2]. We describe how this functionality can

be introduced with minimal overheads to the delivery process and show how this results in a more robust system, e.g., capable of adapting to unreliable network conditions. Our work is based on the MPEG-4 Scalable Video Codec (SVC) [3] as it is a useful underlying technology, both for dynamic content adaptation and caching techniques, through its layering approach.

The remainder of this paper is organised as follows. Section 2 provides some related work and Section 3 describes the application scenario. Section 4 describes the ENTHRONE approach to dynamic adaptation and caching and how this can be used cooperatively. Finally, Section 5 presents the evaluation and analysis we have undertaken to demonstrate the advantages of our approach and we end with conclusions and outlook.

2 Related Work

MPEG-21-based adaptation of SVC [4] involves the intelligent, dynamic removal of certain layers of the SVC bitstream based on the current usage environment (e.g., network bandwidth and display resolution), a high-level generic Bitstream Syntax Description (gBSD) and a description of the available adaptation options. It consists of the following steps. 1) Based on the adaptation options and usage environment, an Adaptation Decision-Taking Engine (ADTE) decides which layers to drop in order to meet the predefined QoS parameters of the session. 2) The bitstream is adapted according to this decision by dynamically disregarding SVC layers at the adaptation node. This is performed in a codec-agnostic way by interpreting the gBSD instead of inspecting SVC layers directly. Thus, this mechanism supports any type of scalable media which is properly described by a gBSD. The adaptation mechanism requires the presence of a monitoring infrastructure in the network and the receiving client to measure the content delivery parameters and trigger adaptation as necessary.

Multimedia Caching [5] describes several stream caching strategies which have been identified for both

¹ENTHRONE (IST-038463), <http://www.ist-enthrono.org>

homogeneous and heterogeneous clients. Heterogeneous clients may possess different characteristics and so may have differing requirements for the same media content, hence the use of layered coding through MPEG-4 SVC. In this way a cache can store the minimal set of SVC layers required by all connected terminals, thus reducing the bandwidth consumed between the server and the caching node at the expense of storage space. Additionally, prefix caching could be used, for example, to ensure that the first X minutes of an object is stored to minimise playback initiation times. The exact parameters involved in these tradeoffs need to be fine-tuned during operation based on usage of the system.

To the best of our knowledge our work represents the first attempt to combine the above mechanisms, i.e., multimedia caching and MPEG-21-based adaptation.

3 Application Scenario

We define our application scenario as being based on an enterprise-class network, which we assume to be a large university network. The key properties that we identify here are as follows:

- Managed network environment
- High capacity up-link to the Internet
- Generally over-provisioned network resources
- Numerous access technologies employed
- Potentially geographically dispersed subnets
- 2000+ heterogeneous end hosts on many subnets

In this case, we assume the university network to include: a high capacity campus 'core', a number of large subnets representing academic departments, a large public wireless infrastructure for staff and students and perhaps a residential students' network. The university may also manage connectivity to local schools or colleges through a variety of technologies. In terms of network capacity, we assume the network core to be overprovisioned by design with ample capacity in wired connections in departmental networks. The wireless network, however, may operate under high loads during peak times and the 'external' links to other institutions may be restricted to xDSL or other limited-bandwidth technologies. In the context of content delivery, we assume our adaptation and caching point will be located close to the network ingress/egress gateway, allowing it to serve all the users on the university network. The adaptation node will lie along the content delivery path, thus negating the need for specific routing, with the caching node either collocated with it (in the same device) or in close proximity as shown in Figure 1. A range of user terminals will be supported in this scenario ranging from large public displays and powerful desktop computers to PDAs (and other tablet devices) and mobile telephones with wireless connectivity.

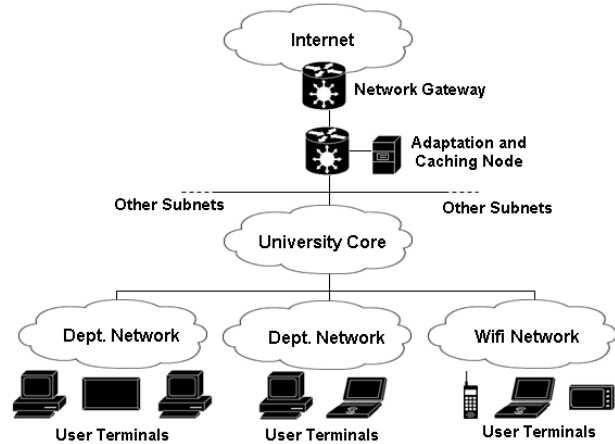


Figure 1. Use Case Deployment Scenario

4 ENTHRONE-enabled Content Delivery

The content delivery management architecture defined within ENTHRONE is called the Integrated Management System (EIMS). It is composed of a number of sub-systems distributed between the entities responsible for managing content creation, advertisement and end-to-end delivery. We focus on the latter process within this paper.

4.1 EIMS Management Architecture

The ENTHRONE architecture identifies four key entities in the content delivery process: the original Content Source (CS), a number of intermediate Network Providers (NPs), at least one adaptation and caching node, and the user terminals. We define the CS as a VoD Server, provided by a Content Provider (CP), which acts as an endpoint of the process. This is separated from the destination (the user terminal) by one or more QoS-enabled Autonomous Systems (AS), each administered by a NP. Along the path there will also be an intermediate entity offering both an adaptation node, which we call the adaptation Television and Multimedia processor (aTVM), and Caching Node that coordinates the delivery and provides adaptation and caching functionality.

Finally, we define a Service Provider (SP) entity which advertises the content, administers the aTVM/Caching Nodes and maintains service agreements with the intermediate NPs and user. The SP also acts as the central EIMS management entity which configures and reserves the resources involved in content delivery before 'handing off' management of the delivery process to the aTVM node. The EIMS sub-systems at the aTVM then initiate the process (including the cache), monitor content delivery and provide adaptation capabilities if necessary. This architecture and flow of management control is shown in Figure 2.

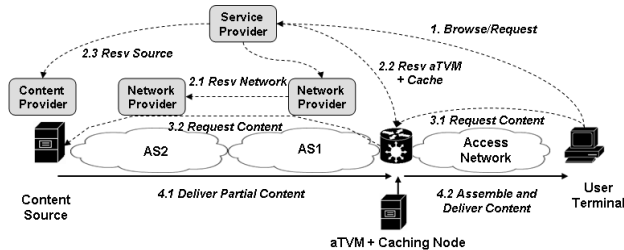


Figure 2. EIMS-based Content Delivery

4.2 Dynamic Content Adaptation

Dynamic adaptation is controlled through the EIMS subsystems at the aTVM. In summary, once an initial content variation is selected (by the SP) and delivery is initiated by the user, the EIMS will configure and invoke monitoring entities along the delivery path to measure the quality of the delivery process. If a monitor detects a situation where the delivery parameters are falling (either due to network contention or some other event), the aTVM can choose to dynamically reduce the content quality to ease the load on the network. This is done by the EIMS Adaptation Manager (AM). The AM incorporates the ADTE and instructs the aTVM to perform an adaptation on a specific content stream passing through it.

Adaptation may be done in a number of ways but, based on the MPEG-21 SVC descriptions which we discuss in Section 2, it involves dynamically dropping specific bit-stream segments, i.e., Network Abstraction Layer Units (NALUs) [3] for SVC, from the content stream. At a simple level, the aTVM may interpret the gBSD and thereby identify packets (containing NALUs) belonging to the stream in question and consequently drop those relating to layers that have been decided by the ADTE to be removed. Based on the gBSD, this can also be performed on an encrypted stream, since the aTVM does not need to inspect the NALUs directly.

4.3 Multimedia Caching

The operation of the Caching Node is transparent to the user terminal and other parts of the ENTHRONE architecture as it will be managed by the SP and interact exclusively with the aTVM collocated with it. The SP administers Caching Nodes by determining the cache replacement policy (LFU, LRU, etc) and the specifics of the caching mechanism used (which layers/how much of the content to cache, etc). The SP does not, however, explicitly provision Caching Nodes with content as they dynamically provision themselves based on what is being delivered to users via the aTVM. In summary, during delivery, the EIMS subsystems

at the aTVM poll the Caching Node to determine if the content is required and, if so, it will forward a copy to the node where it is cached. This involves monitoring the popularity of content in the cache against that being delivered to determine what should be provisioned based on the parameters provided by the SP.

Where the aTVM is delivering content that the Caching Node holds a copy of, the EIMS subsystems in the aTVM coordinate the delivery such that the node is used to source the cached content while the remainder is delivered via the traditional CS. The process involved here however depends largely on the caching method used. Where base-layer caching is used for example, the aTVM must coordinate delivery from the cache and source simultaneously and in parallel such that the frames in each stream arrive together where they can be assembled and forwarded to the user. In contrast, if prefix caching is used, the aTVM would first invoke the cache and then coordinate delivery from the source as the cached content ends. Other caching techniques have been explored within the context of the project.

5 Evaluation and Discussion

In this section we quantitatively analyse the benefits of integrating adaptation and caching of SVC content in streaming scenarios as described in Section 4. In particular we compare:

- SVC delivery without adaptation and caching
- SVC delivery with adaptation
- SVC delivery with adaptation and caching

To this extent, for each scenario we assume a variable number of terminals (with differing usage environments) which all consume the same content with variable start times. Due to space restrictions, we do not evaluate the other advantages of our approach (lower initiation times through prefix caching and improved robustness through dynamic adaptation) here and will do so in further work.

Table 1 provides an overview of the test content. We use four different SVC contents which differ in their scalability dimensions in order to represent a comprehensive test set. Specifically, *SVC 1* provides scalability in the temporal domain by offering six temporal layers while *SVC 2* offers three spatial layers and *SVC 3* offers three quality layers. Finally *SVC 4* offers full scalability with three spatial, five temporal and three quality layers. All of these test contents have an equivalent frame rate (30) and AU count (3000).

As shown in Figure 3, five terminals consume the *SVC 4* content and, due to each terminals' usage environment, the content is requested at different qualities (QCIF to 4CIF in the figure). From left to right, first the QCIF quality is requested, then additionally a CIF quality is requested, and so on, until finally the highest quality (4CIF) is requested.

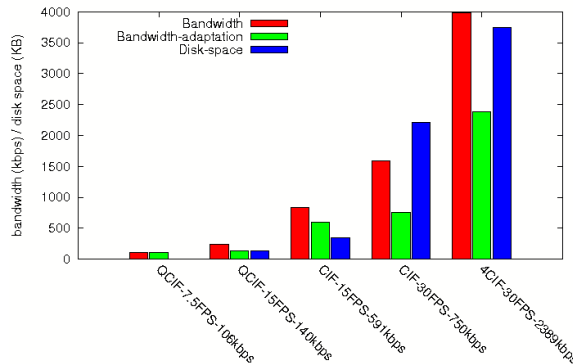
Table 1. Characteristics of Test Data

	SVC 1	SVC 2	SVC 3	SVC 4
Media Size [kB]	7907	27682	18790	29195
Nr. of AUs	3000	3000	3000	3000
Average AU Size [kB]	2.64	9.23	6.26	9.73
Nr. of NALUs	6005	12007	12007	27017
Nr. of VCL NALUs	3000	9000	9000	24000
Nr. of Prefix NALUs	3000	3000	3000	3000
Nr. of PS/SEI NALUs	5	7	7	17
Nr. of Packets	3005	9007	9007	24017
Frame Rate [fps]	30	30	30	30
GOP size	32	16	16	16
Spatial Layers	1	3	1	3
Temporal Layers	6	5	5	5
Quality Layers	1	1	3	3

For each terminal we show the bandwidth and caching disk space requirements for the aTVM/Caching Node at the network gateway (as introduced in Section 3) which correspond to the different delivery scenarios: 1) no adaptation/caching, 2) adaptation, 3) adaptation and caching. Since the benefits of caching become apparent when the same content is consumed with different starting times, we assume at least a 10s time difference between each of the five content requests.

The results of these measurements are shown in Figure 3. For each requested quality, the first bar shows the bandwidth needed without adaptation or caching. This corresponds to delivering a separate stream for each terminal. The second bar shows the bandwidth needed between the server and aTVM/Caching Node, assuming the same start times for all delivered content variations. That is, for the first request only the QCIF-7.5fps quality is delivered, for the second request this is replaced by the QCIF-15fps quality (from which the first quality can be retrieved at the aTVM by performing MPEG-21-based adaptation), and so on. However, adaptation becomes ineffective with differing start times for the content delivery. In such a case, caching needs to be performed in addition to the adaptation. Consequently, the third bar shows the disk space requirements for the cache in order to sustain the lower bandwidth requirements shown in the second bar. Subsequently, we repeated these tests for the three other SVC contents and the results (not displayed due to space restrictions) were generally similar, i.e., required network resources can be considerably reduced at the cost of additional complexity and disk space, due to the adaptation and caching node.

Interestingly, however, our investigations also show that sending a media stream with a quality which can satisfy multiple terminals is in certain cases less efficient than sending individual streams. For example, we consider *SVC 4* and assume terminals requesting QCIF-30fps (which requires 175kbps) and 4CIF-15fps (which requires 1879kbps). With this configuration two individual streams require 2054kbps.

**Figure 3. Bandwidth / Disk Space Requirements for Different Delivery Scenarios**

However, a single stream which satisfies both terminals requires 4CIF-30fps quality which induces a bit rate of 2389kbps. This shows that for fewer terminals with particular quality requirements, our approach is less efficient than traditional delivery.

6 Conclusion and Future Work

In this paper we introduced new delivery concepts which exploit the synergies of adaptation and caching of multimedia content in order to minimize the required network resources. We provided an evaluation of these concepts which show their benefits, but also point out that this approach is not always advantageous. Future work will focus on evaluating the additional complexities and advantages which are introduced by such aTVM / Caching Nodes.

References

- [1] A. Vetro, C. Christopoulos, and T. Ebrahimi. Special Issue on Universal Multimedia Access. *IEEE Signal Processing Magazine*, 20(2), March 2003.
- [2] H. Kara, C. Edwards. A Caching Architecture for Content Delivery to Mobile Devices. *Proceedings of the 29th EUROMICRO conference*, IEEE Computer Society Press, pp 241-248, September 2003.
- [3] T. Wiegand, J. Ohm, G. Sullivan, and A. Luthra. Special Issue on Scalable Video Coding - Standardization and Beyond. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(9), September 2007.
- [4] A. Hutter, P. Amon, G. Panis, E. Delfosse, M. Ransburg, and H. Hellwagner. Automatic Adaptation of Streaming Multimedia Content in a Dynamic and Distributed Environment. In *International Conference on Image Processing*, Genova, Italy, September 2005.
- [5] J. Liu, J. Xu. Proxy Caching for Media Streaming over the Internet. *IEEE Communications, Feature Topic on Proxy Support for Streaming on the Internet*, 42(8):88-94, August 2004.