Rate Adaptation Techniques for WebTV

Hugo Santos[1], Rui Santos Cruz[1], and Mário Serafim Nunes[2]

[1] Instituto Superior Técnico, Lisboa, Portugal
hugo.santos@tagus.ist.utl.pt, rui.cruz@ieee.org
[2] IST/INESC-ID/INOV, Lisboa, Portugal
mario.nunes@ieee.org

Abstract. This paper presents a novel hybrid **RTSP**, **SIP** and **HTTP** solution for providing converged WebTV services with automatic Rate Adaptation of multimedia content. The prototypical solution, being developed under the scope of the *My-eDirector 2012* european project, was deployed in a high-bandwidth LAN and some preliminary tests were performed.

Keywords: IPTV, WebTV, DCCP, Multimedia Streaming, Content Adaptation, Rate Adaptation.

1 Introduction

Telecommunications evolved from System centric to Network centric, and nowadays to User centric (allowing users to access services in a customized way and with a single authentication, always in a unique user session) [4].

An increasing large number of people is already having access to all kinds of multimedia content from all kinds of networks using a large variety of terminal devices, requiring different types of content formats with different quality levels.

Some adaptation systems have been created, enabling users to watch multimedia content on their terminal devices with the maximum quality possible, but in a very strict way, largely dependent on the devices/networks of each Service Provider. The biggest challenge is still the choice of the adequate adaptation technique to use for each multimedia content, in each of the Access Networks of the Provider and for each of the terminal devices connected.

While IPTV is a technology with capabilities to provide live video streaming and television services, over managed IP access networks, with mechanisms to ensure interactivity and appropriate Quality of Service (QoS) and of Experience (QoE) [9, 11, 18], WebTV addresses similar objectives but over the global Internet. While IPTV is usually based on Multicast only requiring IGMP, WebTV requires a more complex signaling protocol in order to initiate, modify and terminate sessions, a control protocol for multimedia contents and transport protocols, responsible for transporting the multimedia streams to the end device.

The Session Initiation Protocol (SIP) [15] and the associated Session Description Protocol (SDP) [7] are the protocols in the ETSI's TISPAN [3] architecture for signaling and session control of IP multimedia applications in NGN [12]. Coupled with Real Time Streaming Protocol (RTSP) [16] methods, for multimedia

P. Daras and O. Mayora (Eds.): UCMedia 2009, LNICST 40, pp. 161–168, 2010.

control, they become a very flexible solution for managing multimedia sessions over IP and for E2E QoS signaling, related to the quality being experienced by the user, to achieve continuous QoS guarantees throughout the session.

This paper presents a novel hybrid **RTSP**, **SIP** and **HTTP** solution for providing converged WebTV services with automatic Rate Adaptation of multimedia content, and shares the experience of its development and evaluation, realized within the scope of the european project *My eDirector 2012* [2]. Several preliminary tests were carried out using high-bandwidth LAN connections.

In this paper, Section 2 describes the State of the Art and Section 3 presents the architecture and functionalities of the WebTV solution prototype. Section 4 presents some preliminary performance evaluation of the prototype. Section 5 concludes the paper.

2 State of the Art

There are several adaptation techniques that can be applied to the multimedia contents that flow over the network. But the efficiency of the adaptation techniques, rely on information that needs to be known before the adaptation technique takes in [13]:

- Information about the **characteristics of the client device**, like the size of the display, the colors that the device support and the size of the buffer.
- Information about the **content**, like the size of the buffer, the minimum streaming bitrate and compression formats.
- Information about the **network to which the client device is connected**, like bandwidth, jitter, packet loss, delays and all the variations of the characteristics that happen in the channel.

From the perspective of an Adaptation System there are three components to consider: the adaptation of the Content, the adaptation to Network conditions and the adaptation for an adequate Quality of Service and of Experience for the End user.

2.1 Video Content Adaptation

The adaptation techniques for multimedia contents can be typically included in three classes [17]:

Format Conversion: This type of technique simply transcodes the original content into another format (e.g., MPEG-4 to MPEG-2).

Selection/Reduction: This type of technique aims to reduce the number of frames of a video or to reduce the resolution of a stream. The most common are the **Resolution Reduction**, **Frame Dropping**, **DCT coefficients dropping** or a combination of the last two techniques.

Substitution: This type of technique replaces the original content by certain elements of it, being the most common the **Video-to-Text**, the **Video-to-Audio** and the **Video-to-Image**.

The points where the adaptation techniques can be used may be located directly in a Service Platform (Server) at the Provider Core, in the End user device (client application) or in a Proxy server (between the Service Platform and the End user device).

2.2 Adaptation to Network Conditions

The Datagram Congestion Control Protocol (DCCP) [10] is a standardized protocol that fills the gap between *TCP* and *UDP* protocols. Unlike *TCP*, it does not support reliable data delivery and unlike *UDP*, it provides a *TCP-friendly* congestion control mechanism in order to behave in a fair manner with other *TCP* flows. *DCCP* includes multiple congestion control algorithms that can be selected, depending on user *QoS* requirements. *DCCP* identifies a congestion control algorithm through its *Congestion Control ID* (CCID).

The most appropriate congestion control algorithm for video streaming is the CCID3 [5, 6], allowing a simple and fast computation of the most adequate transmission rate, and can be expressed by the following equation:

$$THR = \frac{s}{RTT.\sqrt{(\frac{p.2}{3})} + RTO.\sqrt{(\frac{p.27}{8})}.p.(1 + 32.p^2)} \tag{1}$$

where: THR is the transmission rate in bytes/second, s is the packet size in bytes, RTT the round trip time in seconds, p is the loss event rate (between 0 and 1.0) and RTO is the TCP retransmission timeout value in seconds.

2.3 Adaptation for End User Quality Perception

End user perception is crucial for the successful deployment of WebTV services. A practical way to measure End user quality is to use a parametric objective opinion model to estimate the subjectively perceived quality of the video, taking as quality factors the coding quality and potential problems due to transport on the network.

ITU-T has standardized a parametric computational model, as Recommendation G.1070 [8], for evaluating QoE of video-telephony. The model estimates the QoE of video-telephony services based on quality design/management parameters in terminals and networks. This model can be extended to live streaming WebTV environments to compute the best video quality V_q (MOS) at any specific moment in time, by the equation:

$$V_q = 1 + I_{coding} \exp\{-\frac{P_{plv}}{D_{P_{plv}}}\} \tag{2}$$

where I_{coding} is the basic video quality affected by the coding distortion, D_{Pplv} is the packet loss robustness factor, expressing the degree of video quality robustness due to packet loss, and $P_{plv}[\%]$ is the packet-loss rate.

3 Outline of the Proposed Solution

The architecture of the streaming server for the WebTV solution (named InStream Server) is being developed in the scope of the *My-eDirector 2012* european project. A final decision is not yet being taken, but a prototype being developed by the authors considers a hybrid **RTSP**, **SIP** and **HTTP** Server with an **Adaptation System** implemented on it. The basic functionality of the architecture is the following:

- The *Encoders* send live streaming contents to the *InStream Server* but not necessarily all of them at the same time. Each live streamed content is encoded in MPEG-4 and/or H.264 and transmitted with several resolutions.
- A large variety of terminal devices, with a *SIP* module installed, can connect to the *InStream Server*, that starts a *SIP* session with the client after receiving the *RTSP* live streaming content request.
- The quality/bitrate of the streamed content is selected by the Adaptation System implemented in the *InStream Server* based on the periodic reports sent by the client during the session.
- If the user wants to switch to a different content channel, a *HTTP* request is sent to the *InStream Server*, requesting the new stream (channel).

3.1 Adaptation System Model

The main scope of the Adaptation System is to know the exact moment when to switch a stream to a client. For the Adaptation System to be cost-effective and lightweight, not requiring much computational resources during operation, and provide good prediction performance within its scope, the best choice falls on a parametric model. The model that was designed for the Adaptation System uses information from the Network conditions by computing THR (equation 1) from Congestion Control ID 3 (CCID3) algorithm of DCCP [10] and the estimate of the subjective MOS for Video Quality, V_q (equation 2), adapted from ITU-T G.1070 standard [8], with parameters related to the base quality for a given codec, the bitrate and packet loss rate.

V_q will be computed at the same time of THR. With the results, the Adaptation System is able to identify the bitrate/video quality that the Client must receive for the transmission channel conditions. The Stream-Switch Decision algorithm of the Adaptation System at the InStream Server is the following:

A: The Server collects network information sent by clients.
B: With that information, THR is computed.
C: The THR value is then compared to the standard bitrate streams on the Server (from 128 kbps to 1024 kbps) to find the closest upper and lower standard bitrates.
D: The two values are then used (one at the time) for V_q computations.
E: With the results, a specific stream bitrate, with the closest higher video quality is then selected.

3.2 The WebTV Streaming Server Prototype

The *InStream Server* prototype has four main modules (Figure 1): the *RTSP* module, the *SIP* module, the *Link* module and *Adaptation System* module.

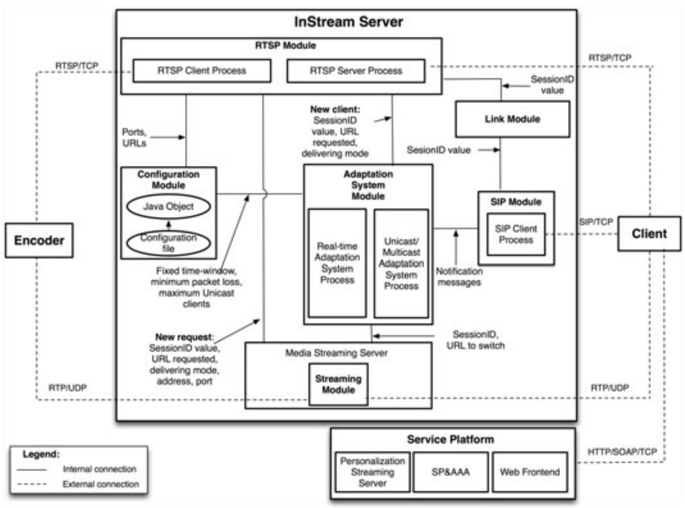


Fig. 1. The InStream Server prototype architecture modules

– **The RTSP Module:** This module is responsible for the content negotiations between the *InStream Server* and the *Encoders*, and also for the negotiation and the streaming of each multimedia content to the clients.
– **The SIP Module:** This module is used to exchange (send/receive) signaling and control messages with clients.
– **The Link Module:** This module provides the liaison between the *RTSP* and the *SIP* modules.
– **The Adaptation System Module:** This module implements the adaptation algorithms for the selection of the most appropriate stream for each client at any moment in time.

The SIP messages: The specific SIP messages used during each session are the **Subscribe Message** and the **Notify Message** [14].

– **Subscribe Message:** This message is created by the *InStream Server* and sent to the client device, to initiate a *SIP* session. The header contains the Unique Identifier created in the *RTSP* session and the body is empty.
– **Notify Message:** This message is sent by the client, periodically (every second), and contains the identifier of the *SIP* session (previously received in the Subscribe message) and, in the body, the information about the transmission channel conditions.

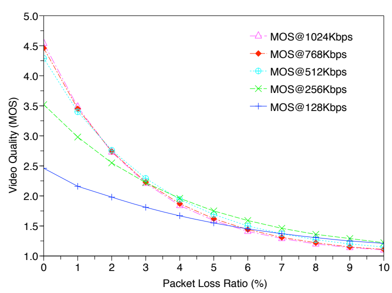
4 Performance Evaluation

The goals while testing the WebTV Streaming Server/Adaptation System prototype were to evaluate the functionalities it provides. A basic Test Architecture was implemented, in a controlled network environment, and a Test Methodology was designed to collect the system response times. The test architecture consists of a Content Streaming Server that stores five multimedia streams of the same content, but with different fixed bitrate/qualities (128 Kbps, 256 Kbps, 512 Kbps, 768 Kbps and 1024 Kbps) and streams them to the InStream Server. The End user client application is a hybrid RTSP+SIP player able to request streams from the InStream Server. The tests were carried out on a private High-Speed 100 Mbps LAN network. To simulate different network speeds and degradation (and introduce some loss) a Bandwidth Controller was implemented at the client system (Bandwidth Controller Standard Edition for Windows [1]). The Test Methodology considered three types of tests:

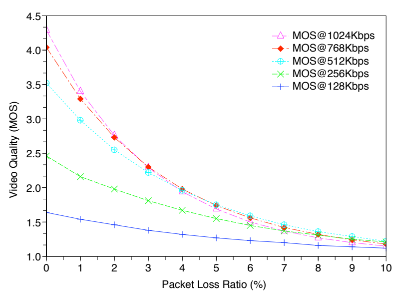
1. Test the effect of Packet-Loss on Video Quality
2. Test the Video Quality vs. bitrate
3. Test the Response times of the InStream Server vs. bitrates

The type 3 Test consisted on a suite of ten different client sessions where the Session Startup, Session Teardown and Stream Switching (10 samples per session) times were collected. A session starts when a stream (request) is ordered to the InStream Server and ends with a Teardown request. For each initial request the InStream Server starts sending the lowest quality stream to the client. The adaptation algorithm, upon receiving feedback information, in a certain timeframe, from the client (player) takes decision about the most appropriate quality class (bitrate/quality) stream to be served to the client.

Test Results: The effect of Packet-Loss for different bitrates can be observed in Figure 2. As expected, the increased packet loss reduces the perceived quality



(a)



(b)

Fig. 2. Video Quality vs. Packet-Loss rate (different bit-rates): (a) H.264; (b) MPEG-4

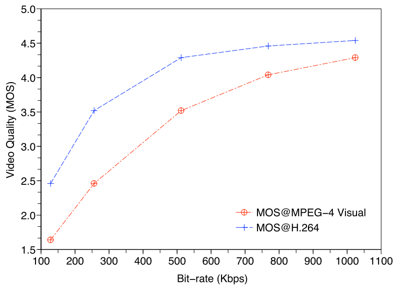


Fig. 3. Video Quality vs. bit-rate

Table 1. Response time of InStream Server

	Stream Switching time (ms) MPEG-4	Stream Switching time (ms) H.264	Session Startup time (ms)	Session Teardown time (ms)
Average	522.2	40.0	1055.4	1.0
Minimum	14.0	11.0	335.0	1.0
Maximum	2519.0	74.0	1390.0	1.0
Standard Deviation	374.7	17.1	357.2	0.0

in a non-linear way. In addition, the quality ranking is in agreement with the bit rate ranking. For the same packet loss rate, the 1024 Kbps sequence exhibits the best quality, whereas the 128 Kbps shows the worst quality. Figure 3 shows the video quality with no losses, where it can be observed that the Video Quality increases with the bitrate, as expected. Table 1 summarizes the results from Response times of the InStream Server. On average, the stream switching time at the InStream Server is around half a second for MPEG-4 and less than 0.1 second for H.264. For the Session Startup time an average of around one second is a very interesting figure.

All the measurements were collected at the InStream Server, reflecting just the time that the application (InStream Server/Adaptation System) takes to react to a certain command or request. No measurements were taken for the network transit time (from client to server, or server to client).

5 Conclusion and Future Work

This project is currently under development. Almost all components are implemented, but not fully tested. One of the key issues in the developed environment is related to stream synchronization. All the streams that come from the *Encoders* should bear the same start time reference (or close, just separated by a few milliseconds). If this doesn't happen, when the stream switching is made, there is a possibility of selecting a stream that is delayed. Future work on this project will include the deployment of components not yet implemented (for scalability of the solution, objective quality estimations at the Client, etc.) the fix of eventual bugs that may exist and an exhaustive test plan (with both Functional and Performance tests), to prove the effectiveness of this solution.

References

1. Bandwidth Controller, `http://bandwidthcontroller.com/index.html`
2. My eDirector (2012), `http://www.myedirector2012.eu`
3. ETSI: TISPAN - Telecoms and Internet converged Services and Protocols for Advanced Networks, `http://www.etsi.org/tispan/`
4. ETSI: Specialist Task Force 360: QoS management at the Network Interfaces to optimize quality of the service delivered to the User (2009), `http://portal.etsi.org/STFs/STF_HomePages/STF360/STF360.asp`
5. Floyd, S., Handley, M., Padhye, J., Widmer, J.: RFC 5348-TCP Friendly Rate Control (TFRC): Protocol Specification. IETF (2008)
6. Floyd, S., Kohler, E., Padhye, J.: RFC 4342-Profile for Datagram Congestion Control Protocol (DCCP) CCID 3: TCP-Friendly Rate Control (TFRC). IETF (2006)
7. Handley, M., Jacobson, V.: RFC 2327-SDP: Session Description Protocol. IETF (1998)
8. ITU-T: Rec. G.1070 Opinion model for video-telephony applications (2007)
9. Kerpez, K., Waring, D., Lapiotis, G., Lyles, J., Vaidyanathan, R.: IPTV service assurance. IEEE Communications Magazine 44(9), 166–172 (2006)
10. Kohler, E., Handley, M., Floyd, S.: RFC 4340-Datagram Congestion Control Protocol (DCCP). IETF (2006)
11. Lee, C.S.: IPTV over Next Generation Networks in ITU-T. Broadband Convergence Networks, 1–18 (2007)
12. Mikoczy, E.: Next generation of multimedia services - NGN based IPTV architecture. In: IWSSIP 2008, pp. 523–526 (2008)
13. Mohan, R., Smith, J.R., sheng Li, C.: Adapting Multimedia Internet Content for Universal Access. IEEE Transactions on Multimedia 1, 104–114 (1999)
14. Roach, A.B.: Session Initiation Protocol (SIP)-Specific Event Notification. RFC 3265, Updated by RFC 5367 (2002), `http://www.ietf.org/rfc/rfc3265.txt`
15. Rosenburg, J., Shulzrinne, M., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., Schooler, E.: RFC 3261-SIP: Session Initiation Protocol. IETF (2002)
16. Schulzrinne, H., Rao, A., Lanphier, R.: RFC 2326-Real Time Streaming Protocol. IETF (1998)
17. Shanableh, T., Ghanbari, M.: Heterogeneous video transcoding to lower spatio-temporal resolutions and different encoding formats. IEEE Trans. Multimedia 2, 101–110 (2000)
18. Volk, M., Guna, J., Kos, A., Bester, J.: IPTV Systems, Standards and Architectures: Part II - Quality-Assured Provisioning of IPTV Services within the NGN Environment. IEEE Communications Magazine 46(5), 118–126 (2008)