Journal of Zhejiang University SCIENCE A
ISSN 1009-3095 (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Delivering audiovisual content with MPEG-21-enabled cross-layer QoS adaptation

AHMED Toufik, DJAMA Ismail
(*CNRS-LaBRI Laboratory, University of Bordeaux-1.351 Cours de la Libération, Talence F-33405, France*)
E-mail: tad@labri.fr; djama@labri.fr
Received Dec. 1, 2005; revision accepted Feb. 17, 2006

Abstract: Future multimedia communication systems have to support the user's needs, the terminal capabilities, the content specification and the underlying networking technologies. The related protocols and applications must be designed from this integration perspective in a cross-layer centric manner. In this paper, we propose an implementation of a streaming service (e.g., Television over IP service) with a unified QoS management concept that enables an IP driven integration of different system components (terminal, user, content, and network). The MPEG-21 framework is used to provide a common support for implementing and managing the end-to-end QoS. The main focus of this paper is on the architecture design, protocols specification and implementation evaluation. Performance evaluations using PSNR and SSIM objective video quality metrics show the benefit of the proposed MPEG-21-enabled cross-layer adaptation.

Key words: Cross-layer adaptation, MPEG-21 multimedia framework, Real-time streaming, QoS
doi:10.1631/jzus.2006.A0784 **Document code:** A **CLC number:** TN919.8

INTRODUCTION

The increasing need for Quality of Service (QoS) assurance mainly for higher bit rates data transfer, minimum packet losses and stringent delay boundary has led to the necessity of squeezing to the utmost the available resources at different layers of multimedia system components (i.e., Network/Terminal/Content). Thus, a new design paradigm recently considered and accepted as beneficial and in the research community is the so-called Cross-layer Adaptation. In fact, this model supports the interlayer communication by allowing one layer to access primitives and data information of another layer even if it is not an adjacent layer. This implies the redefinition of new design strategies, more especially, the interfaces offered in each layer as it breaks the classical open systems interconnection (OSI) model.

The key idea in the cross-layer adaptation is to support the multimedia services at different networking layers. At the application layer, the multimedia services will be adaptive to changing networking conditions and bandwidth availability. At the transport layer, end-to-end congestion control and error resilience mechanisms will be used. At the network layer, resource reservation or differentiation service is used to support end-to-end QoS provisioning. Techniques to provide vertical/horizontal mobility management and connectivity will be used. Feedbacks are used for service adaptation. Routing mechanism needs to be QoS-aware and able to handle mobility. At the data link layer medium access control (MAC) needs to be modified so that reservations are respected and QoS guarantees can be supported. Adaptive power control techniques will be used to manage mobility and maintain active links. Error control techniques will be used to protect against varying error rates. At the physical layer several options are available and will be adapted to the several high level requirements.

In order to support the above mentioned mechanisms, the new MPEG-21 multimedia framework defines several parts to facilitate Digital Item (DI) consumption. MPEG-21 is defined to provide a multimedia framework for multimedia access, delivery and consumption, in order to enable interoperability, content protection and content adaptation in a distributed multimedia system which enables transparent, ubiquitous and pervasive multimedia consumption. In this respect, MPEG-21 aims to create the big picture of a common framework to guarantee interoperability by focusing on how the elements of a multimedia application infrastructure should relate, integrate, and interact. The MPEG-21 defines a data model used to provide the common support for implementing the required functionality and managing of the resources.

In particular, the MPEG-21 part 7 Digital Item Adaptation (DIA) (ISO/IEC, 2004) specifies a set of tools to support interoperable video adaptation, which includes usage environment description (UED), bitstreams syntax description (BSD), and media characteristics description for optimal adaptation to the purpose of the terminal and network QoS. UED includes descriptive information related to user characteristics (e.g., user information and preferences, usage history, presentation preference, audio language preferences, subtitles language preferences), terminal capabilities (e.g., codec capabilities and display capabilities, adaptation capability), network characteristics (e.g., network capabilities and network conditions), and natural environment characteristics (e.g., location and time). UCD provides means for specifying supplementary information which further constrains the usage environment, beyond that possibly defined by UED.

This paper proposes a design and implementation of a cross-layer adaptation for Television over IP services by integrating the management of content, users, networks and terminals based on MPEG-21 standard. The support of universal media access (UMA) with end-to-end QoS provisioning cannot be satisfied without an efficient cross-layer adaptation and optimization. This led us to integrate the related meta-data information for service offering and adaptation. Fig.1 highlights the proposed cross-layer adaptation by integrating content, terminal, user and network profiles. Our solution takes into consideration user/application level QoS and its mapping/ translation to network level QoS.

The rest of this paper is organized as follows: Section 2 presents some related work on cross-layer architecture. Section 3 presents our MPEG-21 enabled cross-layer adaptation. Section 4 describes the implementation issues and some test scenarios. Section 5 presents some performances evaluation and finally Section 6 concludes the paper.

RELATED WORK

In recent years, many papers proposed various solutions addressing the issues of cross-layer optimisation. However, the majority of the papers limits investigations to wireless communication and especially to the MAC and PHY layers interaction. Some of these works do not consider the interaction of user/application-level QoS and its mapping to network-level QoS. We present some parts of these papers in what follows.

In (Zhang *et al*., 2005), a general block diagram of end-to-end QoS supporting video delivery in the network-centric cross-layer solution was proposed. The authors considered QoS mapping across different layers as one of the key issues for their general block diagram. However, application-level QoS was not con-

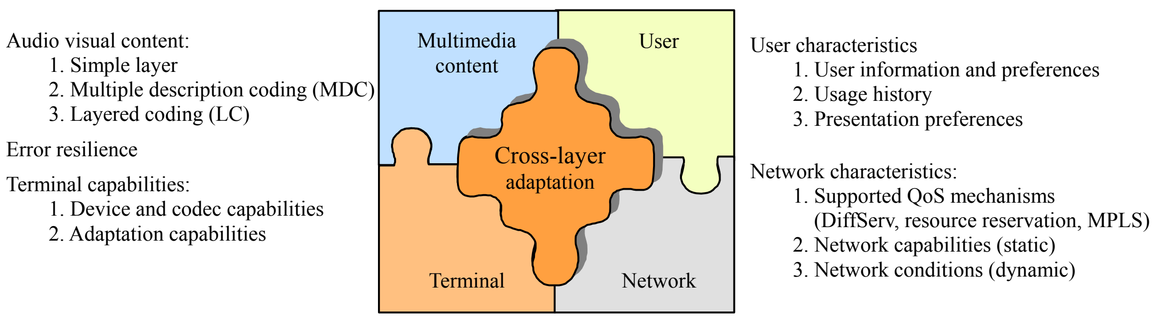

Fig.1 Cross-layer adaptation based MPEG-21 meta-data

sidered. van der Schaar and Shankar (2005) formalized the cross-layer problem, discussed its challenges, and presented a classification of cross-layer solutions. The paper illustrates the relationship between layers and the challenge to find the best configuration that optimizes different metrics at different layers. In the same manner, an overview was provided of the cross-layer paradigm where the physical and MAC layer knowledge of the wireless medium is shared with higher layers (Shakkotai *et al*., 2003; Chen *et al*., 2004; Butala and Tong, 2004; De *et al*., 2004). In (Krishnamachari *et al*., 2003), different error control and adaptation mechanisms available in the different layers are evaluated for robust transmission of video. Based on this evaluation, a new adaptive cross-layer protection strategy for enhancing the robustness and efficiency of scalable video transmission is proposed. Shan and Zakhor (2002) proposed two techniques for adaptive video streaming over network wireless. One of these techniques is based on cross-layer solution where the application layer can access to link layer for retransmission of lost radio link packets. Ahmed (2003) and Ahmed *et al*.(2005) presented an innovative cross-layer content delivery architecture that can receive information from the network and adaptively tune transport parameters, bit rates and QoS mechanisms according to the underlying network conditions. This service-aware IP transport architecture is composed of an automatic content-level audiovisual object classification model; a reliable application level framing protocol with fine-grained TCP-friendly rate control and adaptive unequal error protection; and a service-level QoS mapping/packet tagging algorithm for seamless IP differentiated service delivery. The presented results demonstrate that breaking the OSI protocol layer isolation paradigm and injecting content-level semantic and service-level requirements within the transport and traffic control protocols, lead to intelligent and efficient support of multimedia services over complex network architectures.

There are now some research activities and some R&D projects which deal with the cross-layer interactions for multimedia delivery with QoS support. European IST FP projects are designing and developing next generation systems. The cross-layer paradigm has been also addressed in a number of projects, e.g., BRAIN, MIND, DRiVE, EVERES, PHOENIX, 4MORE and ENTHRONE. The new major topic for European 7th Framework Projects namely "eMobility" will also focus on the cross-layer integration and optimization.

The 4G MC-CDMA Multiple-Antenna System on Chip for Radio Enhancements, (4MORE[1], IST-507039) project and Jointly Optimizing Multimedia Transmission in IP-Based Wireless Networks (PHOENIX[2], IST-001812), addressed cross-layer integration issues. The goals of the IST Network of Excellence Project on Wireless Communications (NEWCOM[3], IST-507325) include identification of the existing gaps in European knowledge of cross-layer, and preparation of an action plan for filling them by capitalizing on project researchers skills. The ENTHRONE[4] (IST-507637) project proposes an integrated management solution covering an entire audiovisual service distribution chain, including content generation and protection, distribution across networks and reception at user terminals by mapping the application level QoS translation to network level QoS.

This paper is aimed to propose a new dimension and extension of the cross-layer interaction and adaptation based on MPEG-21 standard through the support of the user's needs, the terminal capabilities, the content/application specification and the underlying network technologies. Our solution is not limited to the wireless transmission as we take into consideration translating of user/application level QoS to network level QoS through the integration of user, terminal network and content capabilities using cross-layer adaptation and optimization. We define a general framework that considers QoS mapping at different layers of the system. To this extent, an efficient end-to-end QoS signaling mechanism is implemented (for instance based on SDP/SDPng) to deal with capabilities negotiation and QoS translation and mapping. The proposed system is presented in the next section.

CROSS-LAYER ADAPTATION BASED ON MPEG-21 MULTIMEDIA FRAMEWORK

The purpose of cross-layer adaptation is to en-

[1] http://www.ist-4more.org
[2] http://www.ist-phoenix.org
[3] http://newcom.ismb.it/public/index.jsp
[4] http://www.enthrone.org/

sure seamless translation of user/application level QoS to network level QoS which should be maintained across different layers of the system model. User and terminal capabilities along with application and content specifications are crucial for ensuring this requirement. All these capabilities, specifications and profiles represent the meta-data information encoded using the future MPEG-21 standard. The cross-layer adaptation is implemented for a streaming service (live content, VoD, Television service). The service offering is implemented through an IP network for a multitude of heterogeneous terminal (phone, PDA, laptop, PC, etc.) and networks technologies (LAN, WLAN, etc.).

The target architecture is described in Fig.2. It is mainly composed of a DVB (Digital Television Broadcast) terrestrial receiver that allows the reception of MPEG-2 TV programs (up to 2~6 Mbps). The DVB-T receiver is connected to a streaming server that provides audiovisuals content with quality adaptation through a multitude of signaling protocols (e.g., RTSP, HTTP and MMS). These communications protocols use a common platform to express media and session descriptions, namely the session description protocol (SDP) (Handly and Jacobsen, 1998).

User/application level QoS description using MPEG-21 meta-data

SDP is widely used (1) to describe session parameters for announcements and invitations and (2) to describe the capabilities of a system and possibly provide a choice between a number of alternatives. However, the many uses of SDP have led to requests for numerous extensions and have led to recognition of several flaws in the protocol design. For instance, SDP does not provide information related to user characteristics (user information and preferences, usage history, presentation preferences, etc.), and less for terminal capabilities such as adaptation capabilities and network characteristics. For example, QoS-Parameters for different protocols domains such as traffic specification and flow specification or DSCP (DiffServ Code Point) for IP QoS differentiation are not integrated in SDP. These QoS parameters need to be specified in another manner (out of band signaling). Actually the IETF MMUSIC working group is working on new generation of SDP protocol namely SDPng (SDP of next generation) to support various long-term extension.

On the other hand, MPEG-21 can fill the SDP gaps by providing meta-data related to the user, terminal, content and network characteristics. We believe that building SDP on top of MPEG-21 will give intelligent solutions to enable multimedia content access under a wide range of delivery conditions and usage environments. The IETF draft (work in progress) presents a practical approach for harmonizing MPEG-21 with SDPng (Guenkova-Luy *et al*., 2005). Indeed, the MPEG-21 DIA specifies a set of tools to support interoperable video adaptation, including usage environment description, bitstream syntax description (BSD), and media characteristics description (i.e., Adaptation QoS) for optimal adaptation to the purpose of terminal and network QoS. Thus, based on MPEG-21 tools and meta-data information, we propose a cross-layer adaptation system. The adaptation is implemented based on user perceived QoS, application QoS and network QoS. This system

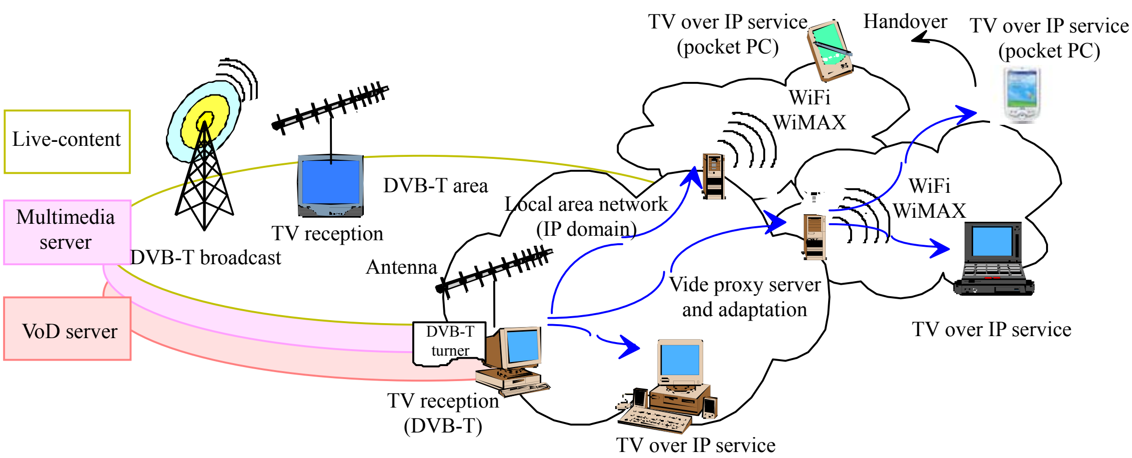

Fig.2 Target architecture: universal media access

allows universal media access anywhere, anytime, using any type of terminal.

Cross-layer adaptation

Fig.3 shows a networked layered architecture model with a subset of protocols and possible cross-layer interactions between different layers. User characteristics, terminal capabilities and content specification are described in MPEG-21 meta-data based on XML data format (see Appendix) and constitute an important layer that should be also considered in the system. For example, without loss of generality, let us take a terminal that supports 8-bit color scheme, which could be expressed by an MPEG-21 XML element (<m21-dia:ColorBitDepth blue="8" green="8" red="8"/>). The supported frame rate of the terminal could be "refreshRate" element and the display capability of its screen is expressed by two elements such as "horizontal" and "vertical" pixels values. These elements are carried by the SDPng description protocol at the connection phase. The streaming server uses the SDPng description to map the UED to the application layer. Another type of mapping is considered at the streaming phase using end-to-end feedbacks (for instance RTCP-based feedbacks) which allow the server to be aware of the changing network conditions. The application layer uses the RTCP feedbacks to adapt the stream to network conditions, not only network conditions that can be dynamically changed. The adaptation is based on the whole changing values in UED (screen format, battery depletion level, user preferences, etc).

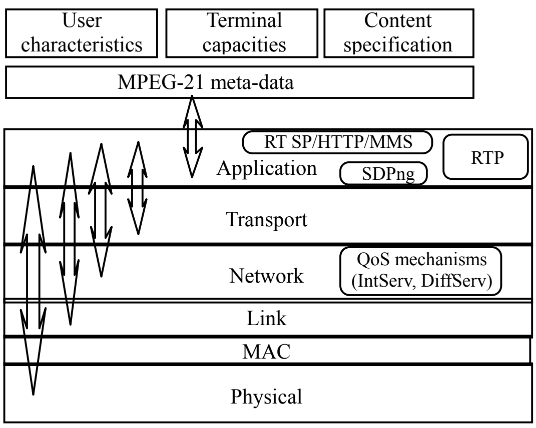

Fig.3 MPEG-21-enabled cross-layer interaction

At the network layer, an advanced interaction with upper layer and lower layer is necessary to assure QoS. An MPEG-21 driven QoS matching and packet tagging algorithm for seamless IP differentiated service delivery mechanisms is used to map the application level QoS to network level QoS. Let us take a video content structured as a single layer stream with I, P, and B frames. Each video frame could be carried over the IP DiffServ network with a particular DSCP code point.

It is admitted that I-frame is more important than P-frame which is more important than B-frame. From this statement, IP packet carrying I-frame data will be marked with low drop precedence compared to IP packet carrying P-frame data and so on. The interface between upper layer and lower layer is subject to the cross-layer interaction that breaks the limit and the border of each layer.

It is clear that allowing universal media access cannot be performed without taking into account dynamic conditions (dynamic UED). Let us take the network characteristics which are considered as static and are used generally for the adaptation at the connection phase. It is described by an MPEG-21 element "NetworkCharacteristic" to provide restrictions like maximum bandwidth "maxCapacity", average throughput, supported QoS classes, etc. These values refer to the Service Level Agreement (SLA) described in the Service Level Specification (SLS). Dynamic features of the network performance is described by the "NetworkConditions" such as network load, IP Packet Delay Variation (IPDV), One-way Packet Loss (OWPL), Round-trip Delay (RTD) or network congestion level such as Random Early Detection (RED), Early Packet Discard (EPD), Early Congestion Notification (ECN), and ICMP source quench message. These elements are also considered for dynamic adaptation. Some of them can be carried by end-to-end protocols such as RTCP. Many types can be adapted by the server. This paper highlights only the benefit of the cross-layer adaptation and does not discuss deeply the adaptation itself.

IMPLEMENTATION ISSUES

The general block diagram of the overall MPEG-21-enabled cross-layer adaptation architecture is depicted in Fig.4 composed of a media server

(Fig.4a) and a client (Fig.4b). The server streams audiovisuals content to the client via an IP DiffServ network using RTP protocol. In case of Television over IP service, the MPEG-2 content can be adapted using many techniques (trans-coding, trans-rating, color down-sampling, etc.). The Transport Stream (TS) Packet-Data Units (PDUs) are encapsulated over UDP. If the underlying network is based on Ethernet, then seven PDU (188-byte length) are multiplexed to form one UDP packet. The client decodes and renders the content based on its adaptation capabilities and its usage environment descriptors.

Let us consider a case of streaming unicast television service over IP for a multitude of users and devices. At the connection phase, the client requests the content through RTSP protocol. The MPEG-21 descriptions of the UED are carried through SDPng protocol (in HTTP or RTSP messages). The server's quality adaptation engine adapts the content based on UED descriptors (color scheme, refresh rate, screen size, codecs capabilities, and network characteristics, etc.). This is a first stage to take into consideration user/application Level QoS and its mapping to lower level QoS.

At the streaming phase monitoring, information can be used to dynamically refresh UED descriptor and for allowing the server to adapt to changing UED conditions. This issue is still open in our solution since we developed only a simple model based on end-to-end RTCP receiver feedbacks and we believe that it needs to be deeply investigated in future work. RTCP "fraction lost", "cumulative number of packets lost" and "inter-arrival jitter" alert the server to be aware of long-term changes in network conditions. This information is used, for instance, to dynamically upgrade/downgrade the received quality by changing the server's sending throughput. This can later be coupled with error-resilience and other adaptation schemes such as Forwarding Error Correction (FEC) and/or real-time transcoding.

In our implementation, both the client and the server are based on open source project VideoLan (VLC) (Available at http://www.videolan.org/vlc/). An extension of its protocols for carrying MPEG-21 meta-data is achieved. VLC is a highly portable multimedia server and player for various audio and video formats (MPEG-1, MPEG-2, MPEG-4, DivX, mp3, ogg, etc.) as well as DVDs, VCDs, DVB-S/T and various streaming protocols. It is used as a server to stream in unicast or multicast in IPv4 or IPv6 network. The server is based on VLC-0.8.4 running Linux 2.6.14 kernel.

The media stream adaptation can be applied at different sides: (1) at the client side, (2) at the content server side, and (3) at any node in the network. Applying the adaptation at the client side is generally useless or impossible because in most cases the client resources are limited (bandwidth and CPU speed). In the same way, the adaptation at the content server side can lead to overloading problem since the server needs

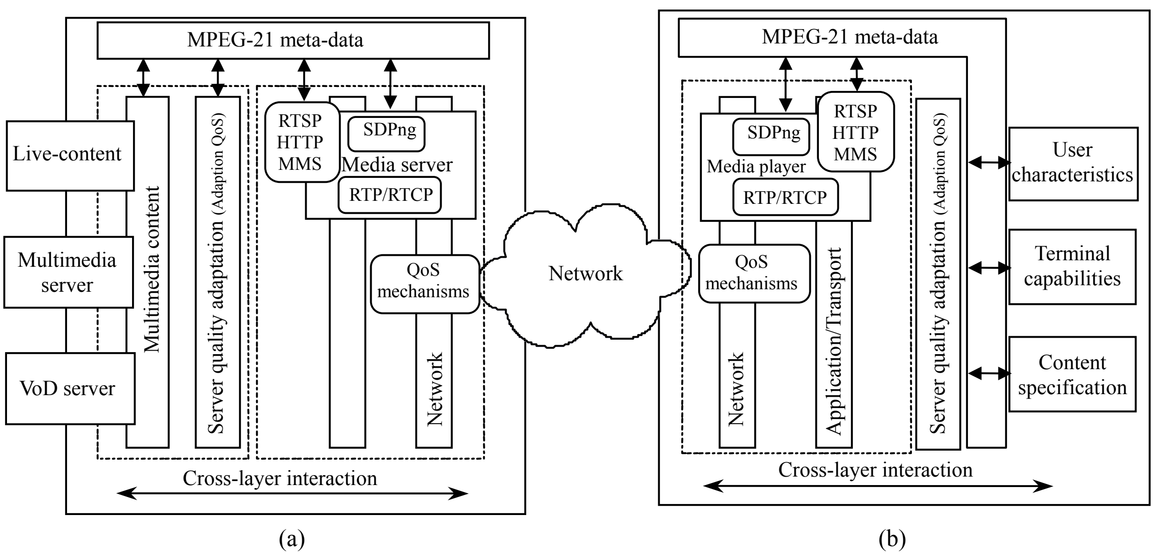

Fig.4 Overall MPEG-21-enabled cross-layer adaptation architecture. (a) Server; (b) Client

to deal with a huge number of clients. Thus, using an independent adaptation server at some strategic node in the network seems to be the best configuration. The best way is to use the adaptation engine between the core and the access network to be as close as possible to potential clients. When the client asks for a particular audiovisual content, the content server sends the stream to the adaptation engine which performs all the necessary adaptation and then sends back the stream to the client. However, our adaptation is implemented at the content server side because we aim to validate our architecture but not to study its performance.

Audiovisual adaptation is carried out in our architecture based on the built-in ffmpeg API of VLC (Available at http://ffmpeg.sourceforge.net/index.php). Different type of adaptation can be done such as real-time transcoding options for video and audio with different bitrate and different scale/aspect ratio adaptation. These options allow heterogeneous terminal with different capabilities to receive real-time multimedia content according to the MPEG-21 UDE descriptors.

PERFORMANCE EVALUATION

This section presents the experimental results of the proposed MPEG-21-enabled cross-layer adaptation. The testbed configuration is shown in Fig.5.

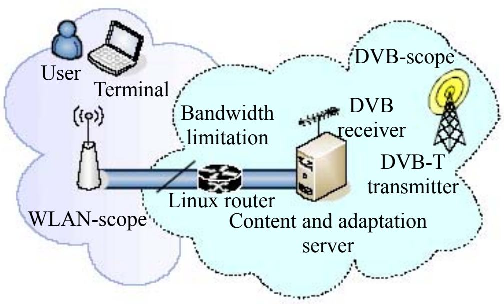

Fig.5 Testbed configuration topology

In our experiment, we use only 60 s length of MPEG-2 audiovisual traffic [© i-TELE (Chaine d'information en continu)] which we believe is sufficient to show clearly the benefit of our architecture. Video quality is achieved by using objective metrics that show the difference between video quality perceived by the client with and without the proposed cross-layer adaptation.

The instantaneous throughput of the MPEG-2 video traffic used in the testbed is shown in Fig.6. It has 720×576 pixels size, an average throughput of 1913 kbps, and a peak rate of 2565 kbps. The audio (MPEG-II Audio) throughput is 192 kbps. This audiovisual traffic will be called "Reference Sequence" (Fig.6).

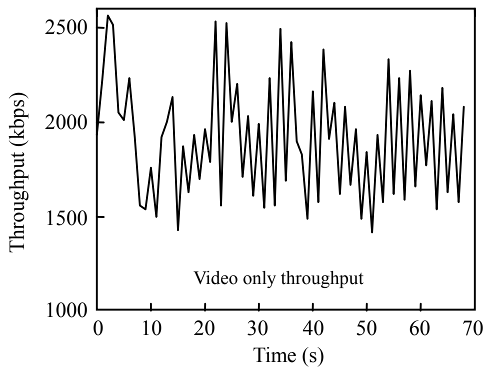

Fig.6 Instantaneous throughput of the reference video

We have used a mobile wireless client to receive audiovisual stream. It is based on 1.5 GHz processor and running a modified version of VLC-0.8.4.

In the following section, we do not consider the adaptation carried by SDPng and performed at the connection phase. We focus only on adaptation performed at the streaming phase. We believe that both of them have the same importance, but we emphasize concretely the benefit of the second adaptation in our test scenarios.

Three test scenarios (Scenarios 1~3) are conducted to check the ability of our system to adapt dynamically to end-to-end congestion feedback (based on RTCP report) using the MPEG-21 cross-layer adaptation. To simulate network congestion, we used TC command line interface provided in Linux iproute2 package to limit the available bandwidth in the network up to 2 Mbps. These scenarios are presented in what follows:

Scenario 1: reference sequence streaming without bandwidth limitation.

Scenario 2: reference sequence streaming with bandwidth limitation (2 Mbps).

Scenario 3: reference sequence streaming with bandwidth limitation (2 Mbps) and using the cross-layer adaptation. We implement two cross-layer adaptations (i.e., Scenario 3.A and Scenario 3.B respectively) policies separately to deeply test the received video quality. Each adaptation is carried out after the reception of the RTCP reception reports from the client. If the "fraction lost" is more than 10 packets loss among 256 packets, we activate one of the following mechanisms to achieve a seamlessly smooth cross-layer adaptation.

In the first cross-layer adaptation (Scenario 3.A), we halve the video throughput of the original sequence. The throughput is halved using "ffmpeg" target bitrate option in the transcoding configuration. In the second cross-layer adaptation (Scenario 3.B), we halve the video throughput of the original sequence and do a video resizing keeping aspect ratio still using "ffmpeg". In this case, to achieve good video quality, we choose the size giving the value of bit/pixel at the level of 0.15 or higher. In our case, 480×384 pixels size is used which gives a level of bit/pixel equal to 0.224.

To measure the objective quality of the received video, we calculate the PSNR (Peak Signal to Noise Ratio) and SSIM (Structural Similarity Index) (Wang *et al*., 2003). The PSNR measures are estimates of the quality of a reconstructed image compared with an original image, while the SSIM measures the similarity between two images, it gives the percentage of similarity for two videos. The SSIM is the newest method and more accurate. To calculate the PSNR and SSIM, we use for each scenario the video sequence received at client and the reference sequence. Fig.7 illustrates the instantaneous received throughputs at the client side for each scenario. The objective quality meter for the received video at the client is given in Fig.8 for PSNR meter and Fig.9 for SSIM meter.

In Fig.7, the throughputs generated by Scenario 3 (A and B) are not stable compared to those in Scenarios 1 and 2. The oscillations are caused by the adaptation since we halve the throughput of video traffic. The server does an adaptation when it receives an RTCP packet with "fraction lost" being more than 10 packets losses among 256 packets and stops this adaptation and return back to normal situation in the contrary case. Even with bandwidth oscillation, the received quality is quite stable over time compared to scenario without adaptation at all. The received video files are available at http://www.labri.fr/perso/tad/download/rv.avi. Fig.8 shows clearly the difference between perceived video qualities by the client in each scenario. The adapted video sequence in Scenario 3 (A and B) provides enhanced quality compared to Scenario 2. The adaptation in Scenario 3.A produces more enhanced quality than the adaptation in Scenario 3.B. This result is confirmed in SSIM chart (Fig.9)

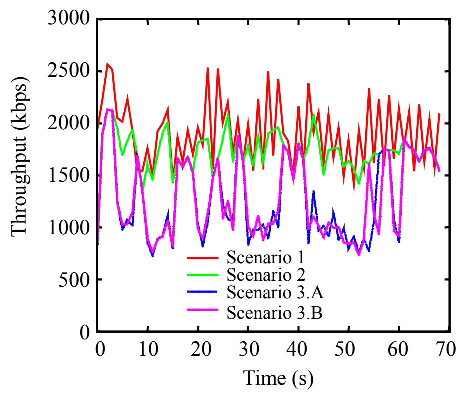

Fig.7 Instantaneous throughputs of the client received video traffic for the different scenarios

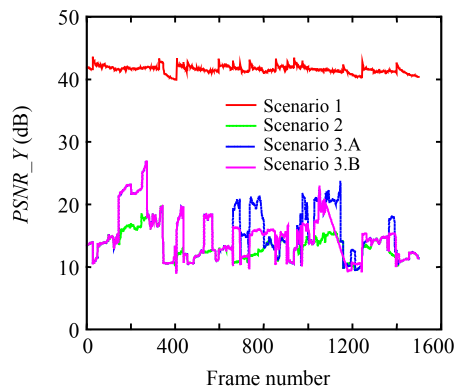

Fig.8 Perceived video quality (*PSNR*) measurements for the different scenarios

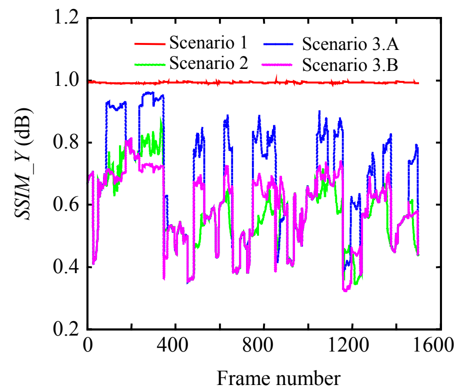

Fig.9 Structural similarity (*SSIM*) Index

with an average similarity index of 66% of the reference video sequence. We can even conclude that the reduction of throughput without video resizing is sufficient and provides superior quality even if we take into consideration the aspect ratio. Some visual results are depicted in Fig.10 for frame number 750. These figures show clearly the advantage of the cross-layer adaptation mechanism compared to system without adaptation. The video quality perceived by the client is enhanced with less bandwidth consumption.

CONCLUSION

In this work, we investigated and implemented an MPEG-21 enabled cross-layer adaptation system for achieving the ultimate goal of Universal Media Access (UMA). Our system focuses strongly on enabling the access to multimedia content anytime, anywhere and for any type of terminals by enabling the integration of user characteristics, terminal characteristics, contents descriptions, and networks capabilities. For this end, our system enables the supporting of multimedia service at different layers of networking and QoS service. The design and the implementation of this system have been presented. Two test scenarios have been conducted using different adaptation policies. The presented results showed that implementing long-term cross-layer adaptation provides enhanced quality compared to traditional system. For instance, the adaptation in the server side is based only on RTCP receiver's feedback for long-term adaptation. Further studies will be done to integrate end-to-end monitoring and network feedback to allow short-term adaptation. Indeed, our previous simulations for network QoS provisioning (Unequal Error Protection, DiffServ Video Marking, Object-based MPEG-4 Video Adaptation, TCP-friendly rate control for video traffic) can be easily integrated in this proposal in future extension of the system to complete the whole cross-layer chain (Ahmed *et al*., 2001; 2005).

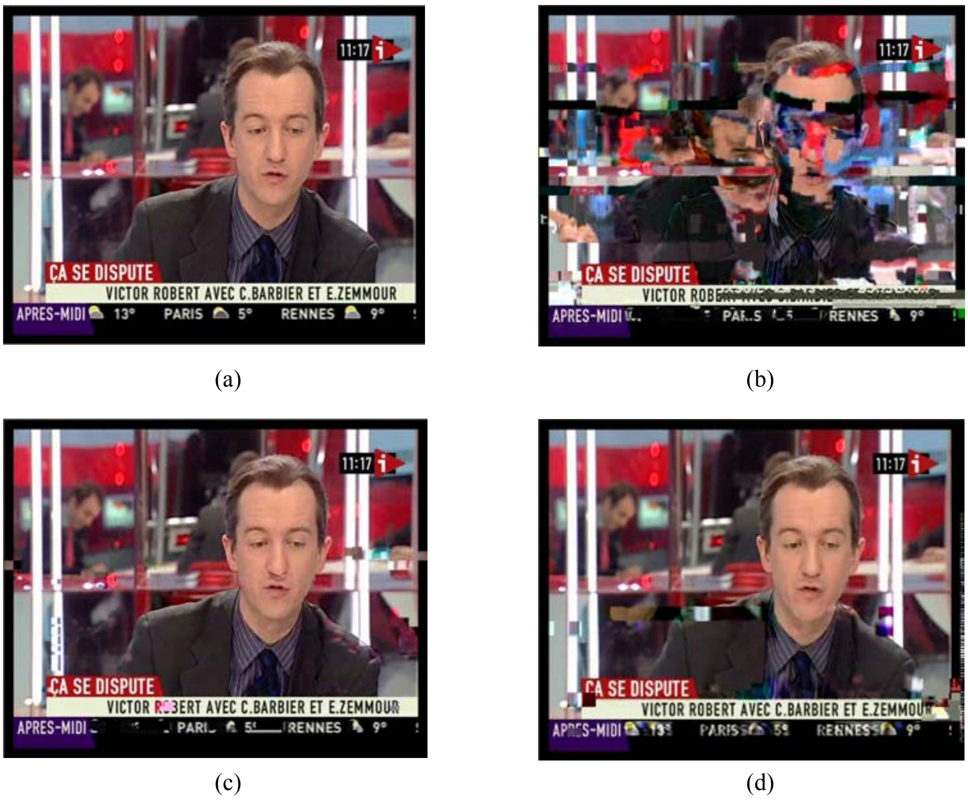

Fig.10 Received video presentation at Frame #750 for different scenarios. (a) Scenario 1; (b) Scenario 2; (c) Scenario 3.A; (d) Scenario 3.B

References

Ahmed, T., 2003. Diffusion Adaptative des Paquets Vidéo sur IP: Une Approche Cognitive. Ph.D Thesis.

Ahmed, T., Buridant, G., Mehaoua, A., 2001. Encapsulation and Marking of MPEG-4 Video over IP Differentiated Services. 6th IEEE ISCC. Hammamet, Tunisia, p.346-352.

Ahmed, T., Mehaoua, A., Boutaba, R., Iraqi, Y., 2005. Adaptive packet video streaming over IP networks: a cross-layer approach. *IEEE Journal on Selected Areas in Communications*, **23**(2):385-401. [doi:10.1109/JSAC.2004.839425]

Butala, A., Tong, L., 2004. Cross-layer design for medium access control in CDMA Ad hoc networks. *EURASIP Journal on Applied Signal Processing*, (2):129-143.

Chen, J., Lv, T., Zheng, H., 2004. Cross-Layer Design for QoS Wireless Communications. Proceeding of International Symposium on Circuits and Systems. ISCAS, p.176-220.

De, S., Qiao, C., Pados, D.A., Chatteriee, M., Philip, S.J., 2004. An integrated cross-layer study of wireless CDMA sensor networks. *IEEE Journal on Selected Areas in Communications*, **22**(7):1271-1285. [doi:10.1109/JSAC.2004.829364]

Guenkova-Luy, T., Schorr, A., Hauck, F., Kassler, A.J., Wolf, I., Feiten, B., Timmerer, C., Romijn, W., 2005. Harmonization of Session and Capability Descriptions between SDPng and MPEG-21 Digital Item Adaptation. IETF MMUSIC WG Internet-Draft.

Handley, M., Jacobsen, V., 1998. SDP: Session Description Protocol. RFC 2327.

ISO/IEC, 2004. Information Technology—Multi-media Framework (MPEG-21), Part 7: Digital Item Adaptation. 1st Ed.

Krishnamachari, S., Schaar, M.V.D., Choi, S., Xu, X., 2003. Video Streaming over Wireless LANs: A Cross-Layer Approach. The 13th International Packetvideo Workshop. Nantes, France.

Shakkotai, S., Rappaport, T., Karlsson, P., 2003. Cross-layer design for wireless networks. *IEEE Communications Magazine*, **41**(10):74-80. [doi:10.1109/MCOM.2003.1235598]

Shan, Y.F., Zakhor, A., 2002. Cross layer techniques for adaptive video streaming over wireless networks. *IEEE ICME*, **1**:277-280. [doi:10.1109/ICME.2002.1035772]

van der Schaar, M., Shankar, S., 2005. Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms. *IEEE Wireless Communications*, **12**(4):50-58. [doi:10.1109/MWC.2005.1497858]

Wang, Z., Simoncelli, E.P., Bovik, A.C., 2003. Multi-Scale Structural Similarity for Image Quality Assessment. Proceeding of IEEE Asilomar Conference on Signals, Systems, and Computers. Asilomar.

Zhang, Q., Zhu, W., Zhang, Y.Q., 2005. End-to-end QoS for video delivery over wireless Internet. *Proceedings of the IEEE*, **93**(1):123-134. [doi:10.1109/JPROC.2004.839603]

APPENDIX

The following section is an example of an MPEG-21 DIA describing a terminal capability:

```
<DIA>
 <Description xsi:type="UsageEnvironmentType">
  <UsageEnvironmentProperty xsi:type="TerminalsType">
   <Terminal>
    <TerminalCapability xsi:type="DisplaysType">
     <Display id="primary_display">
      <DisplayCapability xsi:type="DisplayCapabilityType">
       <Mode>
        <Resolution horizontal="720" vertical="480"/>
       </Mode>
      </DisplayCapability>
     </Display>
    </TerminalCapability>
   </Terminal>
  </UsageEnvironmentProperty>
 </Description>
</DIA>
```

This following example describes the structure of SDPng document:

```
<?xml version="1.0" encoding="UTF-8"?>
<sdpng xmlns = "http://www.iana.org/sdpng"
      xmlns:sdpng="http://www.iana.org/sdpng"
      xmlns:rtp="http://www.iana.org/sdpng/rtp">
      <cap> ... </cap>
      <def> ... </def>
      <cfg> ... </cfg>
      <constraints> ... </constraints>
      <info> ... </info>
</sdpng>
```