Future Issues and Challenges in Distributed Media Synchronization

Fernando Boronat, Mario Montagud,
Universitat Politecnica de Valencia
C/ Paraninf 1, 46730, Grao de Gandia, SPAIN
fboronat@dcom.upv.es, mamontor@posgrado.upv.es

Hans Stokking
TNO
Brassersplein 2, Delft, the Netherlands
hans.stokking@tno.nl

Pablo Cesar, Rufael Mekuria
CWI: Centrum Wiskunde & Informatica
Science Park123, Amsterdam 1098 XG, Netherlands
{p.s.cesar, r.n.mekuria}@cwi.nl

***Abstract*—Current technological advances are resulting in a complex media ecosystem, in which streaming media can reach a variety of end-devices using a heterogeneous set of delivery methods. Such development allows ubiquitous media consumption, but comes at a cost: interoperability. While users can now consume content whenever they want and wherever they are, their alternatives for interacting with others, while watching, are limited. Due to incompatibility issues between different delivery methods and end-devices, shared media experiences across locations are not yet a commonplace. Looking ahead to the future, this paper identifies key issues for enabling truly remote shared experiences (cross-location watching of media, 3D tele-immersion) in a consistent/synchronous way. In particular, it details future challenges for achieving distributed media synchronization across disjoint users using heterogeneous networking technologies and end-devices.**

***Index Terms*— IDMS, Inter-destination Media Synchronization, IDES, Inter-device Synchronization, multimedia synchronization**

I. Introduction: The problem of heterogeneity

In the last years we are witnessing a proliferation of media delivery methods. Media consumption is becoming ubiquitous as xDSL rollouts progress, new generation wireless networks are developed, and deployment of fiber networks into the ground is becoming a reality. Individual consumption of media is becoming easier than ever before, and soon broadcast and broadband services will converge. Nevertheless, the interactivity capabilities of the users have not evolved from the traditional remote control model (start, play, pause, stop). Acknowledging substantial improvements on personal media consumption, it is disappointing to see the small advances on (remote) group media consumption.

This article focuses on emerging (remote) group consumption of media content, with particular interest on one key technological enabler: distributed media synchronization. It identifies a number of challenges ahead, including, diversity of networks and end-devices, variety of media sources and media types, and current possibilities for cloud computing (e.g., for transcoding). The final intention of this article is to pave the way for the future, so novel synchronized media services can emerge.

Lately, media delivery protocols and formats have evolved, making ubiquitous video consumption a reality. Some approaches aim at digital broadcasting (e.g., DVB or DAB), while others are intended for improving media streaming in the Internet (e.g., progressive download). In the second cluster some current solutions include plain Real Time Messaging Protocol (RTMP) based on TCP and HTTP streaming; RTMP and HTTP Dynamic Streaming; and live streaming using RTSP/RTP (Real-Time Streaming Protocol/ Real-time Transport Protocol), RTMP, HTTP, and RTMFP (Real-Time Media Flow Protocol) based on UDP. In addition, there are a number of proprietary delivery methods such as Apple's HTTP Live Streaming, Microsoft's Smooth Streaming, etc. Existing platforms do not fully support the synchronization of media streams that are delivered over different distribution networks due to interoperability issues (e.g., codecs, transport protocol).

Apart from the diversity on delivery networks, the media viewing experience is becoming richer than before due to the variety of media sources and media types. Currently, there are an increasing number of end-devices that support all kind of media content (audio, video, and 3D models). Moreover, users have the opportunity to choose among different media sources that can stream such content. These developments bring new challenges regarding media synchronization in terms of inter-stream (different streams or media types coming from different sources) and inter-destination (as different end-devices might process media in different manners). At the same time, as end-devices in the same physical location cooperate, inter-device synchronization will also be needed. For example, multi-screen experiences at home require the synchronization of media content between the television set and the mobile device (e.g., extra views of a given television program in the mobile device) [1].

Cloud computing can play an important role on media processing. Mature server and networking virtualization technologies can help in offloading media processing from mobile end-devices to the cloud. This allows for better

ISBN - 978-90-5986-410-8

multimedia support on low processing and computational power end-devices (e.g., smartphones and tablets). Some examples include in-video subtitling, content scaling, video transcoding and picture-in-picture [2]. While the cloud opens new opportunities, the effect on synchronization has still to be further evaluated.

This paper discusses in detail the three different upcoming challenges (networking, sources and end-devices, and cloud computing) for enabling (remote) shared experiences. In particular, the article focuses on distributed media synchronization (i.e. inter-stream, inter-destination [3] and inter-device synchronization), as an essential factor for a good quality of experience (QoE). Purposely, other issues such as universal session handling, scalability, presence awareness, and privacy concerns [4] are considered out of scope of this work.

II. Challenges in Multimedia Synchronization

Due to the increased relevancy that IDMS is acquiring for supporting shared experiences, several works in the past have addressed this topic. Some illustrative articles discuss a number of use cases [3], propose qualitative comparisons between synchronization solutions [5], and report on working prototype implementations ([6], [7], [8], [9]). Moreover, standardization of IDMS has been pursued within ETSI TISPAN [10] and it is currently a milestone for the IETF AVTCORE WG [11] (but only for those cases in which RTP is used as the transport protocol). Recognizing the importance of these previous works on the design, definition and classification on IDMS, here some future challenges are highlighted based on three key issues: delivery protocols, end-devices, and the alternative of offloading processing power to the cloud.

A. *Delivery Protocol*

Recent experiments have demonstrated that different delivery technologies cause different end-to-end delays. Measurements show synchronization differences of up to 6 seconds for Internet-based video watching (including P2P) and of up to 5 seconds for TV-based reception [12]. Such delays might become annoying when watching sports at home (assuming thin walls or open windows), but they are unacceptable for (remote) group media watching scenarios. First subjective experiments indicate that a playout difference of over 2 seconds will not be acceptable for distributed consumption of media content [13].

One key challenge for the future is the development of mechanisms for accurate synchronization of media content at the receiver side, independently of the delivery protocol in use. This will become particularly important in convergent broadcast/broadband applications. HbbTV [14] can be used here as an example of such convergence and its associated synchronization challenges, as it is a new standard that aims to provide convergent broadcast/broadband services. However, synchronization in HbbTV is challenging and inaccurate, as it is based on DVB stream events. There are two main problems with this solution. First, stream events can incur on delays of one or two seconds, so frame accuracy is not possible. Second, set-top boxes are programmed to react to the first stream event, ignoring the following ones, so re-synchronization becomes difficult to implement. In general, this solution cannot be considered to be precise enough for advanced synchronization requirements like lip-sync, sub-titles or frame-accurate video synchronization.

In particular, there are a number of challenges for achieving synchronization in convergent broadcast/broadband environments (as identified by the HbbTV-Next[1] European project) [1]:

- Different protocols have different feedback mechanisms and timing models.
- Media streams from different sources and/or over different transport channels have different transmission delays and jitter. Also varying de-packaging processing leads to different processing delays.
- The content delivery nature may be different. Whereas one stream is delivered continuously as a DVB broadcast or IP multicast, another may be transmitted on demand.

In general, convergent scenarios impose a number of challenges in terms of inter-stream synchronization (as services are composed at the receiver side) and of inter-destination synchronization (in remote shared experiences). Due to the lack of interoperability between protocols (e.g., HTTP DASH uses a different paradigm than DVB) and the different delays that can be accumulated when using each one of them, (streaming) protocol independent synchronization solutions are required. Such future solutions can affect media rendering at the receiving side, independently of the protocol used, even more if the requirement is to synchronize streams originated from multiple remote sources.

B. *Inter-Device Synchronization*

Multiple screen multitasking is a reality. On the one hand, a tablet study by the Online Publishers Association[2] reveals that second-screen multitasking (85% of the time) — and even third-screen multitasking (66% of the time) — is a natural phenomenon with the tablet. Likewise, studies on the US market[3] reveal that up to 70% of the people are using auxiliary IP-ready end-devices while watching regular TV. While most multitasking is not directly related to the TV content, around 30% of the users actively search for related information (e.g., via Wikipedia). These facts underline the future importance of well-integrated multiple screen services, which will have to be synchronized with the content being shown in the main screen.

Inter-Device Synchronization (IDES) focuses on the synchronization of different streams on multiple end-devices at a single user side. For example, when a single source of multimedia content is distributed to multiple playout devices, the audio and video content require synchronous playout for multi-channel stereo sound and lip-sync, and video content

[1] http://www.hbb-next.eu

[2] http://onlinepubs.ehclients.com/images/pdf/MMF-OPA_--_Portrait_of_Tablet_User-Wave_2_--_Jun12_%28Public%29.pdf

[3] Synchronized Second-Screen technologies panorama, November 15. 2011, http://blog.eltrovemo.com/529/

ISBN - 978-90-5986-410-8

can also be transferred between end-devices' screens. Other examples are community gaming around TV content on a second screen, rating systems for talent shows and live interactive quiz shows.

IDES may also be seen as a superset of IDMS, as it generally implies synchronization efforts on different end-devices (forming a group), eventually of different types and classes. Unlike IDMS, in which the challenge focuses on compensating long, global delays, without having a local timeline available to synchronize content, the main challenge in IDES is on satisfying consumers in a common space during a joint media event with different end-devices involved. In such situation, users may visually compare performance, and hence synchronization requirements become stricter. In practice, recent research studies regarding IDES show that the delay between two synchronized video streams should not exceed 15 milliseconds [15].

Google TV, Yahoo Connected TV and Microsoft SmartGlass provide platforms for application developers to create second-screen experiences. They all recognize that one of the biggest problems of second-screen experiences is synchronization. Some IDES techniques include Automatic Content Recognition or ACR (e.g., fingerprinting and watermarking techniques), Wi-Fi communications, manual-sync (user actions), adding timelines as a separate content to both MPEG-2 and IP-based content delivery paths ([16]) or the use of BD-Live-sync [17] (it enables end-devices sharing time-sync codes with one another over the LAN). At present, there exist some commercial applications such as Nielsen Media Sync Platform[4] using watermarking; Civolution Videosync[5], Myspace[6] , ConnectTV[7] and the Yahoo Audio-Sync Engine IntoNow[8] using fingerprinting; and Disney second screen app including audio-sync (fingerprinting), manual-sync (user actions) and/or BD-Live-sync.

The upcoming challenge is to provide non-proprietary solutions that can work across delivery networks, end-devices, and types of media. Synchronization of user end-devices in a local environment should not depend on a centralized, ubiquitous server, as it is often applied in IDMS. Involved end-devices should be able to synchronize media between them on their own, by employing a distributed scheme. The synchronization effort should provide a smooth transition between the end-devices, trigger the TV display by user events on the mobile end-device without too much delay, or allow the exchange of services and user's events between end-devices with minimum, and thus unnoticeable, delays. Some other issues that are still open are the communication protocol among end-devices in the same domain, and also among end-devices in different domains, session mobility and viewers' annotations [18].

[4] http://blog.nielsen.com/nielsenwire/online_mobile/nielsen-and-abcs-innovative-ipad-app-connects-new-generation-of-viewers/

[5] http://www.civolution.com/syncnow/

[6] http://www.myspace.com

[7] http://www.connectv.com

[8] http://www.intonow.com

C. *Cloud offloading*

Cloud offloading is becoming a popular technology for media processing, particularly for content intended to mobile devices. While such alternative has the benefit of providing support to a heterogeneous set of end-devices (as content can be prepared for them), it comes with a number of challenges. In terms of synchronization, the main challenges include the extra media processing delay and scalability.

Cloud offloading should provide a central control point for the service provider to control the QoE perceived by the end users and provide homogeneous services across heterogeneous clients and end-devices, content, networks and protocols. Unfortunately, media processing in the cloud adds extra delay that can negatively affect distributed media synchronization (inter-stream synchronization, IDMS and IDES). This problem will be emphasized when media content follow different paths: some streams are live (broadcast), others are downloaded, and others need of cloud processes. There are some approaches that can be adopted to address this problem (processing delays). For example, intelligent scheduling of tasks in the cloud can be applied and network conditions measurements can be used for optimizing such processes.

Scalability is another important issue when using cloud offloading. The cloud offloading mechanisms must scale with the number of users in the system without significantly increasing the cost to the service provider. This is of particular concern since almost any type of video processing can only be accomplished by working with raw audio and video streams. This means that video and audio streams have to be decoded, processed, and then re-encoded. One possible way to keep processing costs in check is to detect duplicate processing tasks requested on multiple clients and then replicate the output of one processing pipeline to all the multiple requesting clients. This poses challenges of duplicate detection and appropriate duplication of outputs across multiple clients without losing the one-to-one interactive capabilities of the applications. Another approach to this problem would be to accurately ascertain the end-device capabilities and share the processing burden between the cloud and the end-device.

III. Conclusion

This paper details upcoming challenges for achieving distributed media synchronization in novel group media sharing scenarios. In particular, it focuses on three main challenges due to the various network protocols used (incompatibility between them and different delays between them), the end-device (and potential synchronization between different end-devices at the same or remote locations), and the possibility of offloading media processes to the cloud.

Currently, existing distribution technologies do not handle the distributed media synchronization problem in an optimal way. As delay variability exits, it becomes a problematic barrier for offering interaction between users in different locations synchronously watching the same piece of content. Further research on devising network-independent, adaptive, flexible, interactive, and scalable solutions is needed.

ISBN - 978-90-5986-410-8

Moreover, there is a need to overcome other challenges regarding heterogeneity of consumer end-devices, cloud scalability (when the cloud is in the path of the media content), and real-time processing bottlenecks. Nevertheless, technical conditions for Inter-Device Synchronization or IDES might be easier to handle, since hybrid media streams must only be synchronized at the same end-device, which is mostly a home media environment.

In any case, before novel solutions for distributed media synchronization are proposed and developed, there is a need to better understand the problem at hand. Currently, the exact ranges of asynchrony levels tolerable by users have not been sufficiently determined. Depending on the task at hand (e.g., chatting while watching content), it is assumed that different levels of de-synchronization will be acceptable [3]. While some initial experimentation has taken place ([19] and [13]), exhaustive subjective tests are needed in order to find out the acceptable bounds for each use case.

Therefore, it can be concluded that further and intensive research on distributed media synchronization is needed in order to address a growing demand for (remote) media sharing experiences.

ACKNOWLEDGMENT

This work has been financed, partially, by Universitat Politècnica de València (UPV), under its R&D Support Program in PAID-11-02-331 Project and in PAID-01-10. CWI's work has been partially funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. ICT-2011-7-287723 (Reverie project). TNO's work has been partially funded by European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. ICT-2011-7-287848 (HBB-NEXT project).

REFERENCES

[1] P. Cesar; D.C.A. Bulterman; and A.J. Jansen, "Usages of the Secondary Screen in an Interactive Television Environment: control, enrich, share, and transfer television content," Proceedings of the European Interactive TV Conference (EuroITV2008), Salzburg, Austria, July 3-4, 2008, pp. 168-177.

[2] HBBNext Project, Deliverable 4.1, ANALYSIS: Cloud-Based Services and Service/Content Synchronisation, March 2012.

[3] M. Montagud; F. Boronat; H. Stokking and R. van Brandenburg, "Inter-destination multimedia synchronization: schemes, use cases and standardization", Multimedia Systems Journal (to appear in 2012) 2012, DOI: 10.1007/s00530-012-0278-9.

[4] I. Vaishnavi; P. Cesar; D. Bulterman; O. Friedrich; S. Gunkel; D. Geerts, "From IPTV to synchronous shared experiences challenges in design: Distributed media synchronization", Signal Processing: Image Communication, vol. 26, no. 7, Aug. 2011, pp. 370-377.

[5] F. Boronat; J. Lloret; M. García, "Multimedia group and interstream synchronization techniques: A comparative study", Information Systems, vol. 34, no. Mar. 2009, pp. 108-131.

[6] F. Boronat; J.C. Guerri; J. Lloret, "An RTP/RTCP based approach for multimedia group and inter-stream synchronization", Multimedia Tools Appl., vol. 40, no. 2, Nov. 2008, pp. 285-319.

[7] T. Hashimoto; Y. Ishibashi, "Group Synchronization Control over Haptic Media in a Networked Real-Time Game with Collaborative Work", Netgames'06, Singapore, October 2006

[8] Y. shibashi; S. Tasaka, "A comparative survey of synchronization algorithms for continuous media in network environments, Local Computer Networks, 2000. LCN 2000. Proceedings. 25th Annual IEEE Conference on , vol., no., pp.337-348, 2000, doi: 10.1109/LCN.2000.891066,

[9] I. Vaishnavi, "Coherence in Synchronous Shared Experiences," PhD Thesis, 2011. http://dare.ubvu.vu.nl/handle/1871/19782

[10] ETSI TS 183 063 V3.5.2 (2011-03), "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); IMS-based IPTV stage 3 specification".

[11] R. van Brandenburg; H.M. Stokking; M.O. van Deventer; F. Boronat; M. Montagud; K. Gross, "RTCP for inter-destination media synchronization", draft-ietf-avtcore-idms-06.txt, IETF working draft, July 2012.

[12] R. Mekuria; P. Cesar; D.C.A. Bulterman, " Digital TV: the effect of delay when watching football," in Proceedings of the European Interactive TV Conference (EuroITV2012), Berlin, Germany, July 4-6, 2012

[13] D. Geerts; I. Vaishnavi; R.N. Mekuria; M.O. Deventer and P. Cesar, "Are we in Sync? Synchronization Requirements for Watching Online Video Together," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, (ACM CHI 2011), 2011, pp. 311-314

[14] Hybrid Broadcast Broadband TV (HbbTV), ETSI TS 102796 v1.1.1, EBU-UER, Technical Specification, 2010.

[15] Y. Bang; J. Han; K. Lee; J. Yoon; J. Joung; S. Yang and J-K. Kevin Rhee, "Wireless Network Synchronization for Multichannel Multimedia Services", Feb. 15-18, 2009 IEEE ICACT 2009

[16] C. Howson; E. Gautier; P. Gilberton; A. Laurent; Y. Legallais; "Second screen TV synchronization," Consumer Electronics - Berlin (ICCE-Berlin), 2011 IEEE International Conference on , vol., no., pp.361-365, 6-8 Sept. 2011

[17] Blu-ray Disc Association 2010 - White Paper: Blu-ray Disc™ Format (General). 2nd Edition, October, 2010

[18] M.J. Montpetit; N. Klym and T. Mirlacher, "The Future of IPTV: Adding Social Networking and Mobility", ConTEL, 2009.

[19] M.O. van Deventer; H.M. Stokking; O.A. Niamut; F.A. Walraven; V.B. Klos, "Advanced Interactive Television Service Require Synchronization", IWSSIP 2008, Bratislava, June 2008.

ISBN - 978-90-5986-410-8