Design, Development and Evaluation of an Adaptive and Standardized RTP/RTCP-based IDMS Solution

Mario Montagud
Universitat Politècnica de València (UPV), SPAIN
mamontor@posgrado.upv.es

ABSTRACT
Inter-Destination Media Synchronization (IDMS) is essential for enabling pleasant shared media experiences. The goal of my PhD thesis is to design, develop and evaluate an advanced RTP/RTCP-based IDMS solution fitting the requirements of the emerging distributed media consumption paradigm. In particular, standard compliant extensions to RTCP are being specified to allow for an accurate, adaptive and dynamic IDMS control when using RTP for streaming media. Moreover, the feasibility and suitability of several architectural schemes for exchanging the IDMS information, algorithms for allowing a dynamic IDMS monitoring and control, as well as adjustment techniques are being investigated. Objective and subjective testing are being conducted to validate the satisfactory performance of our IDMS solution and to provide insights about the users' tolerance on asynchrony levels in different IDMS scenarios.

Categories and Subject Descriptors
C.2.4 [**Computer-Communication Networks**]: Distributed Systems; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems.

General Terms
Design, Human Factors, Performance, Standardization.

Keywords
IDMS, Media Synchronization, RTP/RTCP, Simulation.

1. INTRODUCTION & MOTIVATION
Traditional research on media synchronization (sync, hereafter) has been primarily focused on orchestrating the play out of incoming media streams locally at single devices. On one hand, *intra-stream sync* deals with the maintenance of the temporal relationships between the Media Units (MUs), e.g. video frames or audio samples, within a specific media stream. On the other hand, *inter-stream sync* refers to the preservation of the temporal dependences between MUs of independent, but correlated, media streams, being *lip-sync* the most representative example [1].

With the development of distributed multi-party media services, such as video-conferencing or Computer Supported Cooperative Work (CSCW), the need for a third type of media sync emerged: *Inter-Destination Media Sync* (IDMS) [2, 3], which involves the simultaneous synchronization of the media playout of a specific media stream across separated devices. The relevancy of IDMS is increasing in the emerging media consumption landscape. Nowadays, innovative forms of shared experiences are gaining momentum [4], allowing geographically distributed users to socially interact (e.g., using text, audio or video chat) within the context of simultaneous content consumption. Some examples are Social TV, networked games, 3D Tele-Immersion, synchronous e-learning, etc. However, realizing those shared interactive services faces a lot of challenges [5]. In particular, end-to-end (e2e) delay differences when streaming the same media content to a group of geographically distributed clients must be compensated for. Previous works have revealed that that e2e delay differences can range from a few milliseconds up to 8 s, depending on the technology in use [4, 6-8]. More specifically, the measurements in [7] showed that the e2e delays for the shortest and longest paths of an IPTV scenario can differ up to 6 s. These delay differences are significantly larger than acceptable limits in most of the IDMS use cases [4]. As an example, a controlled user study analyzed the effects of de-synchronization on the QoE in shared video watching scenario, in which users communicated each other (via text and voice chat) [9]. It was concluded that delay differences up to 1 s might not be perceptible by users, but differences over 2 s really become annoying for most of them. These results provide initial empirical evidence that actual e2e delay differences lead to high rates of user dissatisfaction. Moreover, Social TV is not, by far, the most restrictive IDMS use case, and other scenarios require stringent sync levels [4]. Consequently, this motivates the design, development and (objective and subjective) evaluation of an inter-operable, adaptive and accurate IDMS solution to compensate such e2e delay variability, thus enabling seamless and coherent shared media experiences. These are the *final goals* of my PhD thesis.

The design and specification of an IDMS solution comprises several aspects. A key issue is to choose or to specify suitable protocols to stream the media and to exchange timing information between the involved sync entities. Besides, different architectural schemes between the sync entities can be adopted. In addition, several monitoring and control algorithms, as well as adjustment techniques, need to be designed to achieve IDMS. Accordingly, *associated sub-goals* of my PhD thesis deal with the analysis of the feasibility and suitability of different alternatives for all the above components of the IDMS solution under development.

This paper provides an outline of the methodology, contributions and impact on the research community of my PhD thesis.

2. METHODOLOGY
In order to achieve the objectives of my PhD thesis I intend to use the methodology sketched in Fig. 1, which is briefly described in this Section.

MM'13, Oct 21-25 2013, Barcelona, Spain
ACM 978-1-4503-2404-5/13/10.
http://dx.doi.org/10.1145/2502081.2502219

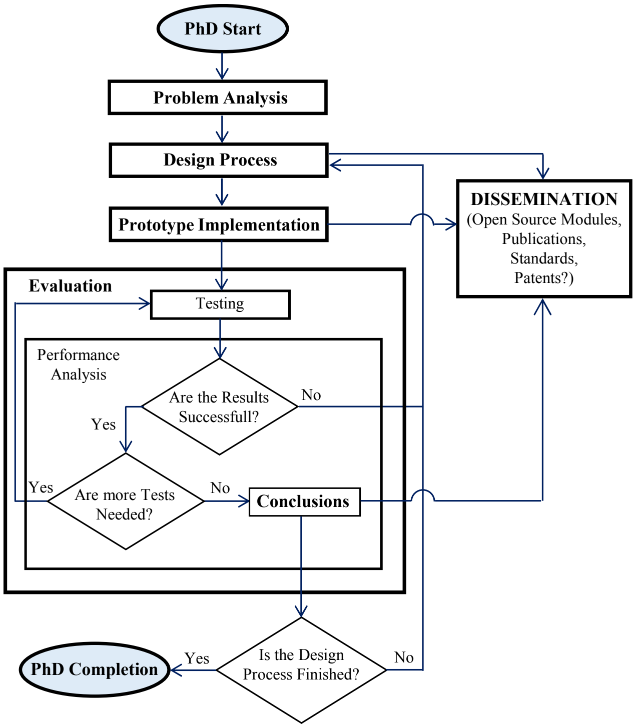

Figure 1. Workflow of the PhD thesis.

2.1 Problem Analysis

I started my research with a thorough review of the start-of-the-art and a compilation of use cases requiring IDMS [4]. This allowed me: i) to better understand the research problems and the associated challenges; ii) to identify the strengths and limitations of existing IDMS solutions; and iii) to derive the technical requirements that must be provided.

Several works have addressed this topic up to date (e.g., [2, 5, 10]). In [3], our research group presented a survey of existing IDMS solutions, which will be extended and updated in my PhD thesis. But apart from numerous papers, two recent PhD dissertations have addressed the IDMS problem. The PhD thesis by I. Vaishnavi [11] was focused on three relevant areas to provide coherence in shared media experiences in cross-domain scenarios: QoS, distributed media synchronization, and user mobility. In this research work, the *local-lag* algorithm [10], which was originally devised for networked games, was adapted to be used for shared video watching. The PhD thesis by Z. Huang [12] proposed a dissemination framework to meet the sync requirements (including IDMS) for multi-site tele-inmmersive systems involving time-correlated multi-sensory media streams.

2.1.1 Requirements for IDMS

Most of the existing IDMS solutions (surveyed in [3]) define new proprietary protocols that may increase the network load and make compatibility between (third-party) implementations more difficult. In addition, some of these IDMS solutions are only based on compensating the network delay variability, but do not take into account variable delays that are originated at the client side [4, 6, 7]. Some other solutions uniquely rely on syncing specific events (e.g., stream content or position updates), but do not support continuous monitoring and control processes.

My thesis acknowledges all the previous advances on IDMS, but mainly aims to design, develop and evaluate, both objectively and subjectively, a wider-applicable, inter-operable, adaptive and highly accurate IDMS solution, being capable of overcoming the above limitations while additionally fitting key requirements for IDMS, such as the following ones:

1. The IDMS solution must rely as much as possible on existing technologies (standard, if possible). This will assure interoperability when being deployed in large-scale scenarios, involving third-party infrastructure and communication devices.
2. Provisioning of timelines (and additional metadata, such as sequence numbers, payload type and source identifiers) in the media delivery units. This would enable a simpler reconstruction of the original media timing and allow for aligning timelines from different streams at the client side.
3. The IDMS solution must rely on wall-clock timing information. Additionally, the existence of mechanisms for negotiating the use of common global wall-clock sources between the involved sync entities is desired. This would allow a coherent framework for stamping, interpreting an aligning timelines in the media delivery chain.
4. Compensation of the e2e delay (i.e., from capture/retrieval at the server side to presentation at the client side) variability between distributed sync entities.
5. Need for a feedback control channel. This would enable a continuous monitoring and control of the IDMS process. This signaling channel, or an additional one, will also be very useful to negotiate key aspects for IDMS (e.g., group membership or clock sources).
6. Inherent rate adaption techniques (at the server-side, at the client-side, or both) to adjust the IDMS timing, allowing to dynamically maintain and restore the original synchronicity.
7. Support of the IDMS solution by both network and end-system entities. This would allow a future deployment of a highly scalable IDMS solution (e.g., for IPTV services) [8].

2.2 Design Process

The next step of my PhD thesis consists in the design of all components of the IDMS solution. This includes the protocol specification (fitting the above requirements), architectural schemes, monitoring and control algorithms, as well as adjustment techniques to achieve IDMS. More details about each of these sub-parts of the IDMS solution under design are in Section 3.

2.3 Prototype Implementation & Evaluation

According to the schedule of the design process, the prototype implementation and evaluation processes will not be performed at a single stage, but will be repeated for each individual component of the IDMS under design, as well as for the global IDMS solution at a later stage (see Fig. 1). In order to validate the performance of the IDMS solution, two main evaluation methodologies are being employed: simulation and user perception tests. Therefore, the first step consisted in selecting the most appropriate frameworks for deploying appropriate prototypes. Network Simulator 2 (NS-2) and GStreamer were chosen, respectively. The implementation in a simulation framework enables better flexibility for double checking assumptions and for identifying critical issues during the design and evaluation processes. The repeatability of the experimentation is more feasible than in real-world assessments. Moreover, it facilitates the assessment of the performance of our IDMS solution in heterogeneous network environments (under different and controllable conditions). Through simulation tests, the responsiveness, effectiveness and consistent behavior of all deployed control schemes, algorithms and adjustment techniques for IDMS can be assessed, by providing measurements about networking-related aspects, objective QoS statistics, as well as delay differences (asynchrony) between Sync Clients. In [13, 14], readers can find the obtained simulation results and conclusions up to date. However, even though considering the convenience of simulation tests, they cannot definitively validate the performance of our IDMS solution. This is because the user satisfaction (QoE) on the perceived shared media experience is the key aspect that will determine the benefits provided by our IDMS solution, and this cannot be obtained through simulation tests. Hence, the implementation of a prototype of a real media sharing application is also required, which is a key stage of my PhD thesis (work in progress). This will also allow me to carry out objective testing, but, most importantly, a subjective evaluation in real scenarios.

As a summary, the evaluation methodology I am employing in my PhD thesis, as well as the associated metrics under assessment, are sketched in Fig. 2. Based on the findings and conclusions from the complete evaluation process I will hopefully have gathered the necessary outputs to conclude my PhD thesis.

3. RESEARCH CONTRIBUTIONS

This section outlines the ongoing and expected research contributions of my PhD thesis.

3.1 Current Progress

3.1.1 Protocol Specification: RTP/RTCP for IDMS

Several standard protocols were analyzed for assessing their suitability for being extended for IDMS purposes [14]. But, definitely, RTP/RTCP were selected because of their compliance with most of the above-described requirements. Moreover, these protocols are extensively used in managed streaming services, they provide both intra- and inter-stream sync mechanisms, and their specification (RFC 3550) allows further extensions to them. After selecting RTP/RTCP for IDMS, the next step was to choose the most suited extension points to accomplish our requirements. RFC 5968 provides common guidelines for extending those protocols, while allowing for interoperability and protocol maintenance. Accordingly, a newly-defined RTCP XR block (RFC 3611) for IDMS, called IDMS report, was specified to enable Sync Clients to provide feedback about reception and/or presentation times for specific RTP packets. But apart from using RTCP for monitoring purposes, control of the timing of the Sync Clients is also needed. Therefore, an additional RTCP packet type for IDMS, called IDMS Settings packet, was specified to provide guidance on when to playout the media. Additionally, an SDP attribute was defined to signalize the use of these RTCP messages. The format of the RTCP messages and their exchange process to achieve IDMS can be found in [4, 14, 15].

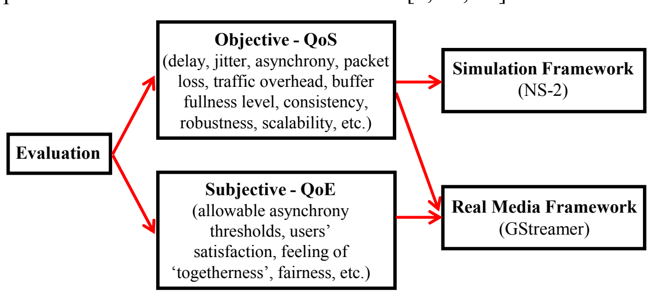

Figure 2. Evaluation Methodology.

I am actively involved in the standardization process of the protocol specification for IDMS within the IETF [15]. In my opinion, standardization will help the uptake of implementations, ensuring a more widespread use of IDMS in practice.

3.1.2 Group-based IDMS Control

One of the targeted goals of my thesis was to enable the co-existence of various logical groups of users in a shared media session, while allowing an independent control of their playout processes. The proposed RTCP messages and the SDP attribute for IDMS have taken into account this requirement [15]. Also, further control algorithms have been designed to achieve this capability, as proved in [13, 14].

3.1.3 Architectural Schemes for IDMS

Three main architectural schemes for IDMS can be distinguished [4]: two centralized schemes (Master/Slave or M/S Scheme and Synchronization Maestro Scheme or SMS) and a distributed one (Distributed Control Scheme or DCS). The involved sync entities and their communication process to achieve IDMS in each one of them is illustrated in Fig. 3. A qualitative comparison among these schemes, taking into account several key factors for IDMS, is provided in my PhD thesis [4] (summarized in Table 1). Each scheme has its own strengths and weaknesses, and is best suited for specific IDMS use cases. Subsequently, I also decided on their development in an extended version of our IDMS solution [14], which was initially based only on the use of SMS [13]. The consistent behavior and effectiveness of these control schemes for IDMS is validated in my PhD thesis [14]. Moreover, several techniques are being designed to enhance their responsiveness regarding specific features, such as coherence for DCS [14], as well as scalability and interactivity for SMS (see Section 3.2).

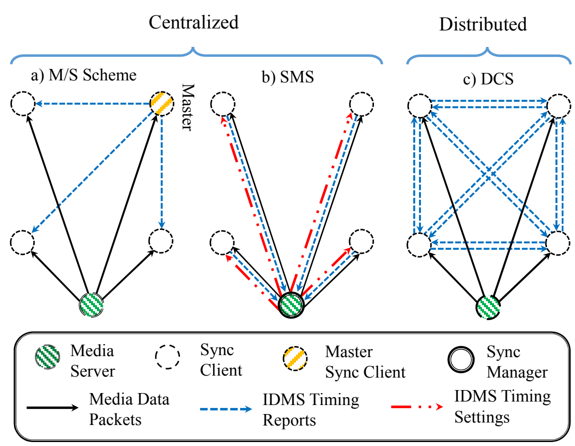

Figure 3. Architectural Schemes for IDMS.

Table 1. Qualitative Comparison Among IDMS Schemes

		Factors									
		Robustness	**Scalability**	**Traffic Overhead**	**Interactivity**	**Consistency**	**Causality**	**Coherence**	**Flexibility**	**Fairness**	**Security**
Scheme	**DCS**	1	2	3	2	3	3	2	2	1	3
	SMS	2	3	2	3	1	2	1	1	2	1
	M/S	3	1	1	1	2	1	3	3	3	2

1. Best Scheme, 2. Good Scheme, 3. Worst Scheme

3.1.4 Master Reference Selection Policies

Once the IDMS reports from all Sync Clients in a specific group have been gathered, a reference IDMS timing to synchronize with has to be selected. The feasibility and suitability (in terms of interactivity, efficiency, fairness, buffer fullness level, etc.) of different dynamic master reference selection policies, for each architectural scheme in use, are being explored (see e.g., [13, 14]).

3.1.5 Adjustment Techniques

After selecting the IDMS reference, the next step is to perform the required adjustments to synchronize. A compilation of server- and client-based techniques that can be used to prevent, maintain or restore sync between separated clients will be provided in my PhD thesis. From these, two reactive client-based techniques have been designed. The first one is based on aggressive playout adjustments (*skips & pauses*) whilst the second one makes use of smooth playout adjustments (Adaptive Media Playout or AMP). The performance of these techniques, for each one of the deployed IDMS schemes, is being assessed in my thesis (see e.g., [13, 14]).

3.2 Next Steps

This sub-section discusses the next steps to complete my PhD thesis, which are scheduled for the next 20 months.

3.2.1 Event-Driven RTCP Feedback Mechanisms

Novel event-driven RTCP reporting mechanisms will be designed for enhancing the performance of our IDMS solution in terms of interactivity, flexibility and accuracy when using SMS. In particular, these mechanisms will be targeted to allow the transmission of Early RTCP packets (RFC 4585) as a response to observed events of interest, while still adhering to the RTCP traffic bounds (RFC 3550). Three main advantages are expected to be provided: i) earlier correction of out-of-sync situations; ii) concurrent presentation of dynamic media-related events in a fine-grained synchronized way with the piece of content they refer to; and iii) delay reduction when joining IDMS-enabled sessions.

3.2.2 Optimized AMP Technique

Further research on AMP for IDMS will be targeted to meet two additional QoE-related goals. The first one is to find the best curve for minimizing abrupt changes in the playout rate. The second one is to meet a trade-off between a proper playout buffer occupancy in all Sync Clients and an overall IDMS control.

3.2.3 QoE Perception Tests

A key future step of my PhD thesis is to analyze the benefits on the QoE provided by our IDMS solution. To accomplish this, rigorous subjective testing in different media sharing scenarios will be carried out, complementing previous studies (e.g., [9, 11, 12]. These user perception studies will be targeted to answer the following research questions (one of the main contributions of my PhD thesis): i) which asynchrony levels are noticeable or annoying to users in different cases (as done by Prof. Steinmetz for inter-stream sync in [1]); ii) which architectural schemes and adjustment techniques are best suited for IDMS; and iii) the sync accuracy that can be achieved with our IDMS solution in real setups. A set of questionnaires will be employed and multiple interviews will be conducted to investigate the influence of several aspects on the perceived QoE. Statistical analysis will also be performed to confirm the validity of the experimentation.

3.2.4 Standardization

We plan to continue with our efforts for standardizing the proposed RTP/RTCP extensions for IDMS.

4. ACKNOWLEDGMENTS

This work was supported by UPV, under its R&D Support Program in PAID-01-10 and PAID-00-12 Projects. Moreover, I would like to thank my supervisor, Fernando Boronat (UPV), and Pablo Cesar (CWI) for their continuous support and very helpful advices.

5. REFERENCES

[1] Steinmetz, R., Human perception of jitter and media synchronization, *IEEE J. Sel. Areas Commun*, 14(1), 1996.

[2] Akyildiz F., Yen W., Multimedia group synchronization protocols for integrated services networks, IEEE J.Sel. A. Commun. 14, 1, 162-173, January 1996.

[3] Boronat, F., Lloret, J., García, M., Multimedia group and inter-stream synchronization techniques: A comparative study, Inf. Syst, 34(1), 108-131, March 2009.

[4] Montagud M., Boronat F., Stokking H., Brandenburg R. van, Inter-destination multimedia synchronization: schemes, use cases and standardization, *Multimedia Systems Journal*, 18(6), 459-482, November 2012.

[5] Vaishnavi I., Cesar P., Bulterman D., Friedrich O., Gunkel S., Geerts D., "From IPTV to synchronous shared experiences challenges in design: Distributed media synchronization", Signal Processing: Image Communication, Vol. 26, Issue 7, pp. 370-377, August 2011.

[6] ITU-T: G.1050 network model for evaluating multimedia transmission performance over internet protocol, 2007.

[7] van Deventer M.O., et al., Advanced Interactive Television Service Require Synchronization, *IWSSIP'08*. Bratislava.

[8] Stokking H., et al., IPTV inter-destination synchronization: A network-based approach, *ICIN'10*, Berlin, October 2010.

[9] Geerts, D., et alter., Are we in sync?: synchronization requirements for watching online video together, CHI 2011, New York (USA), May 2011.

[10] Mauve, M., et alter., Local-Lag and Timewarp: Providing Consistency for Replicated Continuous Applications, *IEEE Transactions on Multimedia*, 6(1), February 2004.

[11] Vaishnavi I., Coherence in Synchronous Shared Experiences, PhD Dissertation, Advisor: Dick Bulterman, Vrije Universiteit (Amsterdam), 2011.

[12] Huang Z., Synchronized distribution framework for high-quality multi-modal interactive teleimmersion, PhD Dissertation, Advisor: Klara Nahrstedt, University of Illinois at Urbana-Champaign, 2012.

[13] Montagud M. and Boronat F., Enhanced adaptive RTCP-based Inter-Destination Multimedia Synchronization approach for distributed applications, *Computer Networks Journal*, Volume 56, Issue 12, pp. 2912-2933, August 2012.

[14] Montagud M., Boronat F., Stokking H., Cesar P., Design, Development and Assessment of Control Schemes for IDMS in an Evolved and Standardized RTCP-based Solution, *Computer Networks*, Article in Review.

[15] Brandenburg R. van, Stokking H., van Deventer M.O., Boronat F., Montagud M., Gross K., RTCP for inter-destination media synchronization, draft-ietf-avtcore-idms-12, IETF AVTCORE Group, Internet Draft, July 2013.