



HAL
open science

World First Retransmission of Medical Images in an Operating Room via DICOM-RTV: Technical Details

Saad El Jaouhari, Patrick Hardy, Guillaume Pasquier, Pierrick Gutter, Eric Poiseau, Cédric Moubri-Tournes, Stéphane Leduc, Sebastien Vincendeau, Laurent Bourgeois, Bernard Gibaud, et al.

► To cite this version:

Saad El Jaouhari, Patrick Hardy, Guillaume Pasquier, Pierrick Gutter, Eric Poiseau, et al.. World First Retransmission of Medical Images in an Operating Room via DICOM-RTV: Technical Details. COMPSAC 2019 - 43rd IEEE Annual Computer Software and Applications Conference, Jul 2019, Milwaukee, United States. pp.294-299, 10.1109/COMPSAC.2019.10222 . hal-02430883

HAL Id: hal-02430883

<https://hal.science/hal-02430883>

Submitted on 7 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

World first retransmission of medical images in an operating room via DICOM-RTV: Technical details

Saad EL JAOUHARI¹, Patrick HARDY², Guillaume PASQUIER¹, Pierrick GUITTER¹,
Eric POISEAU³, Cédric MOUBRI-TOURNES¹, Stéphane LEDUC¹, Sébastien VINCENDEAU⁴,
Laurent BOURGEOIS⁴, Bernard GIBAUD⁵ and Emmanuel CORDONNIER¹

¹b-com, Cesson Sévigné, France, ²Harmonic, Cesson Sévigné, France

³INRIA Rennes, Equipe SED, Rennes, France, ⁴Hospital Pontchaillou, Department of Urology, Rennes, France,

⁵Univ Rennes, Inserm, LTSI - UMR 1099, F-35000 Rennes, France

Emails: ¹{FirstName.LastName}@b-com.com, ²patrick.hardy@harmonicinc.com,

³eric.poiseau@inria.fr, ⁴{FirstName.LastName}@chu-rennes.fr, ⁵bernard.gibaud@univ-rennes1.fr

Abstract—This paper provides technical details regarding the worldwide premiere implementation of the DICOM Real-Time Video (DICOM-RTV), in which an experimental setup was deployed at the Rennes University Hospital in France during five Urological surgeries. The solution allows the real-time retransmission of medical images from different equipment in the operating room in a perfectly synchronized way. In this paper, we show how we have been able to transport and synchronize the video signal from a camera recording the hand gestures of the surgeon, and from an endoscope used by the same surgeon. This paper also demonstrates the interest of using the new DICOM-RTV standard in order to improve the surgical gestures precision and to provide precise, real-time and synchronized instructions to the different operators during a surgery, toward the next generation operation rooms.

Index Terms—DICOM, DICOM-RTV, Real-Time, Video, Medical Metadata, Operating Room, Healthcare, Synchronization.

I. INTRODUCTION

Digital Imaging and Communications in Medicine (DICOM) is considered as the backbone of the image display in modern medicine, and currently the driving force behind the imaging workflow in hospitals around the world. Its main objective is to provide standards to manage all the chain from the medical image acquisition, transfer, processing and interpretation, to the storage, in order to facilitate the interoperability between the different medical imaging equipment. Moreover, all the PACS (Picture Archiving and Communication System), which comprise the acquisition modalities, the storage and the viewing workstations, are implementing the DICOM standard. Additionally, DICOM is recognized as a standard by European Committee for Standardization (CEN), and International Organization for Standardization (ISO).

However, the standard does not handle the streaming of real-time video, audio signal, and the associated metadata yet. Generally, proprietary baseband video or video over IP solutions are being used to handle the transport of the real-time video, yet, they provide limited interoperability since they neither transfer the corresponding medical metadata nor enable the synchronization of multiple videos. Hence,

a new DICOM extension called DICOM Real-Time Video (DICOM-RTV), described in the Supplement 202 [1], was introduced in order to solve these issues. The Supplement defines a new IP-based DICOM Service for transmitting real-time video, together with the corresponding metadata in a synchronized way, to subscribers (e.g. display monitor), with a quality of service comparable to the one obtained via video cables. It also promotes the interoperability of the different medical media equipment, by standardizing the way they communicate (i.e. via DICOM-RTV). Moreover, it upgrades the use of the current systems by the IP-based one, where the only Ethernet interface is present.

The DICOM-RTV extension relies on the SMPTE (ST 2110 family of standards), a rapidly emerging set of standards used in the professional video (e.g., TV studios) equipment, and complements it with data-content specifications consistent with other parts of the DICOM standard. It uses mainly the system architecture specified in SMPTE ST 2110-10 [2]. As a first step, the current version of DICOM-RTV is restricted to the real-time streaming of uncompressed video, since the underlying standards (SMPTE ST 2110 family) do not include any transport of compressed video, yet. Nevertheless, the proposed mechanism for conveying the medical metadata along with the video (and audio) is fully compatible with the transport of compressed video and it is anticipated that when underlying standards embrace compressed video, DICOM-RTV will be extended to support them.

Furthermore, the proposed extension allows the transmission of several flows of data (e.g. video flows, video/audio flows, and data flows) in the same medium, by taking advantage of the multiplexing nature of the IP, with a special focus on the medical metadata which are required to ensure the patient safety (e.g., operating the right patient) and enabling advanced applications using video and data (e.g., image guided surgery).

In order to demonstrate DICOM-RTV, a world first implementation of the standard was realized in collaboration with

the Rennes University Hospital (CHU). The demonstration took place during five urological surgeries, involving the transmission of synchronized real time video of a camera and an endoscope. The demonstration validated the retransmission of medical images from different equipment in the Operating Room (OR) in a perfectly synchronized way. In this demonstration DICOM-RTV allowed the display of two synchronized videos, one from a camera recording the hand gestures of the surgeon, and another one from an endoscope. Fig. 1 is a screenshot of two synchronized frames of the real-time video, where the light in the right top video corresponds to the instant when the surgeon uses the laser. The main purpose of this demonstration was to point out the surgical gestures precision and to provide precise, real-time and synchronized instructions to the different operators during a surgery, towards the next generation ORs.

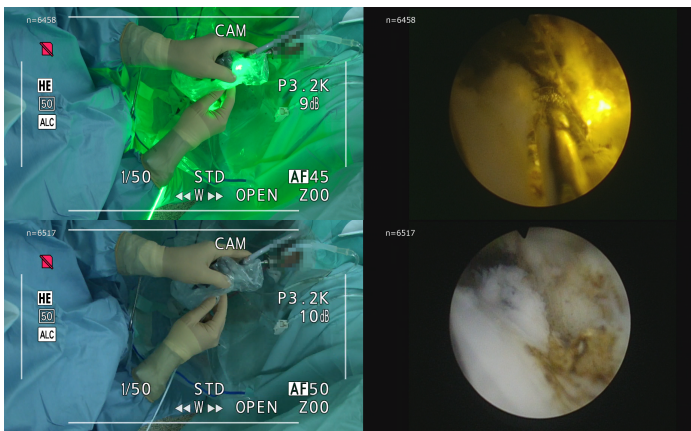


Fig. 1. Synchronized real-time videos from the operating room during a surgery

Thus, this paper underlines technical details regarding this demonstration and it is structured as follows: in section II the related work is presented, which discusses the literature works that deal with the real-time medical videos, mainly inside a medical infrastructure such as a hospital. Then, the global architecture of the DICOM-RTV system is proposed in section III. Next, in section IV, technical details are provided in order to clarify each component of the architecture and the different standards used for this purpose. This section is followed by the result and discussion section V which discusses the obtained latency results and the future challenges, subject to future works. Finally, the paper is concluded in section VI.

II. RELATED WORK

In this section, a literature review regarding the solutions that manage real-time medical videos in the OR is presented. It shows that in the last decade only a few works have been done in order to solve this challenge. In [3] the authors propose a solution to store real-time x-ray images of a patient's organ structure from a fluoroscope, as a part of the patient medical record. The video is captured, digitalized, converted and compressed in an MPEG2 format and integrated into the

PACS in a DICOM file format. The proposed solution does not deal with the synchronization and the real-time display of the video. Another work presented in [4] takes advantage of the image assisted surgical treatment and developed a multimedia electronic patient record system that manages and distributes the real-time multimodality imaging and informatics in order to assist the surgeon during the different clinical phases of the operation (spinal surgery). They used images, clinical forms, waveforms, textual data for planning the surgery, two real-time imaging techniques (digital fluoroscopy) and endoscope video images, and more than half a dozen live vital signs of the patient in order to assist and monitor the surgery. All these data are acquired, displayed and archived in real-time. The acquired data are synchronized using a master clock and displayed live into a customized interface using an LCD screen in the OR. This paper suggests that from the technical point of view, what DICOM-RTV enables has been accessible for a while. However, because OR is a multiple providers infrastructure, no such system have been deployed on the clinical side since. Only a urbanist approach, relying on commonly adopted interoperability standards can make it possible to enable OR to go to this level of performance. In [5] the authors present a development of a radio channel model for an angiography room in a hospital. They use wireless communications in the 60 GHz band to send video data from an X-ray machine to a movable LCD screen in the OR. This paper is anticipating the next step of OR technology, integrating wireless communications for flexibility (e.g., introducing a non planned equipment during the procedure). DICOM-RTV will enable such evolution when compressed video is supported on the one hand, and when wireless network manages prioritization among the flows (slicing) and ultra-low latency, both supported by the 5G.

III. GLOBAL ARCHITECTURE

Inside the OR of the University Hospital of Rennes, the architecture, shown in Fig. 2, was deployed during five urological surgeries. The hospital provided access to the modality **Worklist**, allowing the retrieval of the metadata concerning the patients undergoing the surgeries. It contains information such as the name of the patient, age, sex, the name of the surgeon, and so forth. These metadata were then received by the **Configurator**, which sent them later to a transmitter after treatment. The OR allowed managing also several essences of medias that were used in order to assist the surgeon during the operation. In this demo, a special interest was given to the **Endoscope** and to a **Camera** hanged on top of the surgeon hand in order to monitor his gestures. However, video sources from different medical equipment can be used (e.g., a demonstration of a perfect synchronization between an ultrasound video and an endoscope video was successfully done).

Inside the OR, three private networks were deployed, a **Hospital Network**, a **Configuration Network**, and a **Data**

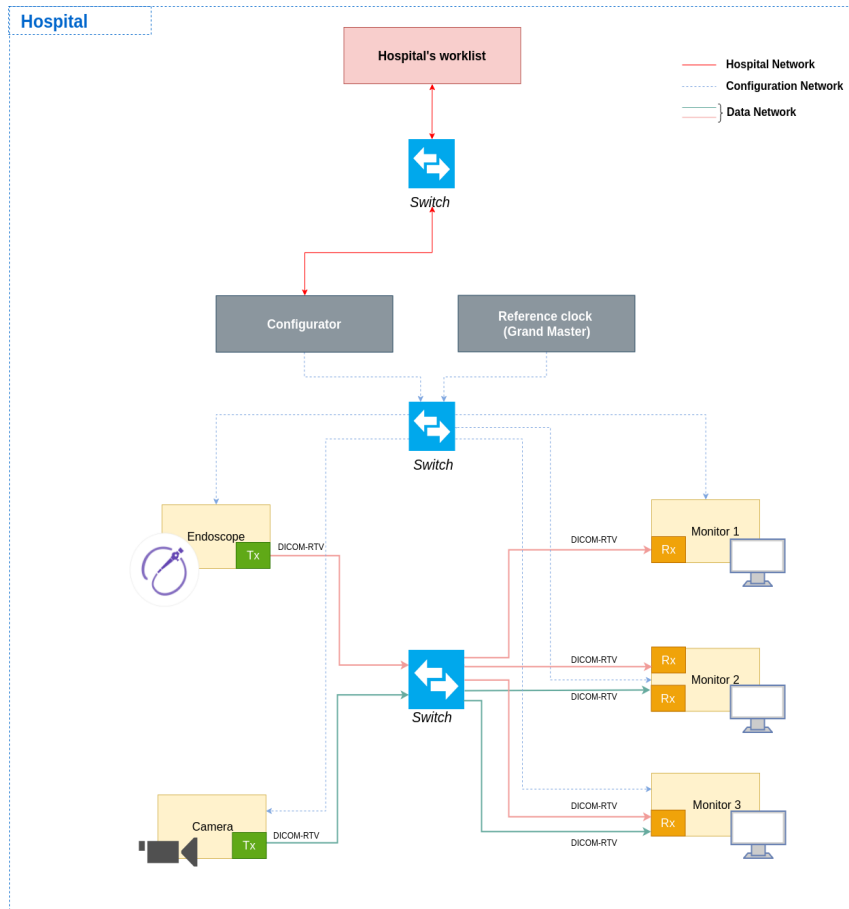


Fig. 2. Architecture of the OR

Network. The Configurator used the hospital network in order to reach, and query, the hospital's workflow. Secondly, the configuration network used an Ethernet switch (1 Gigabit) in order to control and command the different equipment of the OR. The switch was connected, on the one side, to the Configurator and to a reference clock (Grand Master) in order to configure all the equipment and also to lock their clocks for synchronization purpose, and on the other side to the OR's equipment. Finally, the data network, which used a fiber optic network switch with high bandwidth (up to 25 Gigabit) was deployed in order to handle the DICOM-RTV flows. Such high bandwidth was needed since the network had to transport heavy video streams (i.e. uncompressed 1980x1080, 10-bit depth, and 50 frames per second videos).

In this architecture, we assume that the different equipment support DICOM-RTV. It means that the source of the media integrate a DICOM-RTV Transmitter (Tx), and the consumers of the media stream (i.e. the monitors) integrates a DICOM-RTV Receiver (Rx), as explained later in the paper.

The proposed solution offers three main possibilities to display these flows:

- Case 1: to directly display the received flow on a monitor

(i.e., Monitor 1).

- Case 2: to display two flows, received by two corresponding Rxs, on a single Monitor (i.e. a Mosaic Monitor 2). This case is used when synchronization between the two videos is requested.
- Case 3: is to receive two DICOM-RTV flows and to process them to obtain a single video and to display the result in a single monitor (i.e. Monitor 3). In this case, augmented reality algorithms using synchronization may be applied.

IV. IMPLEMENTATION

This section describes the technical details of the different components of the architecture and the implementation choices.

A. Architecture implementation

The only difference with the global architecture is the location of the Tx's/Rx's. Since the equipment currently do not support DICOM-RTV, the Tx's and Rx's have been implemented and deployed as external components as shown in Fig. 3. We argue that with the advancement of the integration of DICOM-RTV extension to the DICOM standard, the Tx's/Rx's can be integrated directly in the modality, which will allow it to communicate in DICOM-RTV.

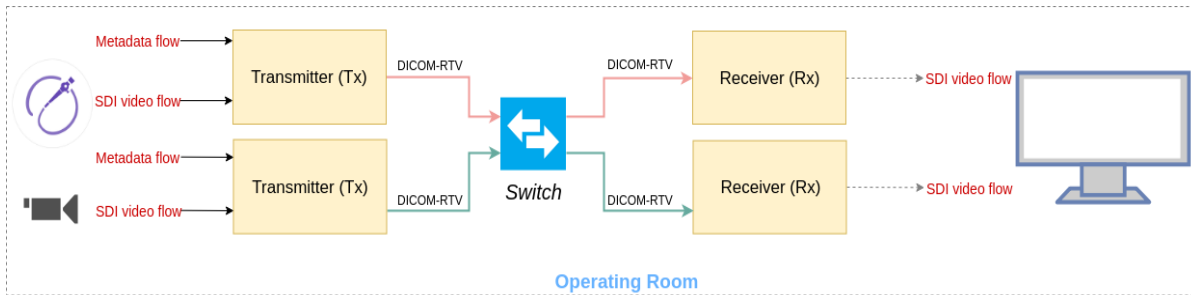


Fig. 3. External Tx/Rxs

B. DICOM-RTV format

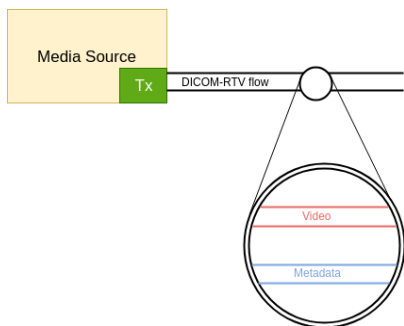


Fig. 4. DICOM-RTV format

As shown in Fig. 4, a single DICOM-RTV flow encompasses two flows, corresponding to the video and to the metadata ones. Firstly, the DICOM-RTV video flow is formatted in compliance with SMPTE ST 2110-20 [6].

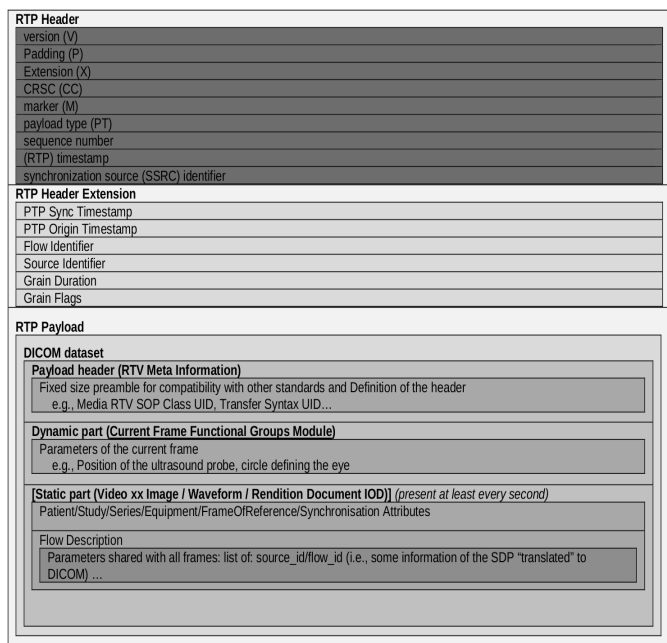


Fig. 5. DICOM dataset encapsulation within RTP

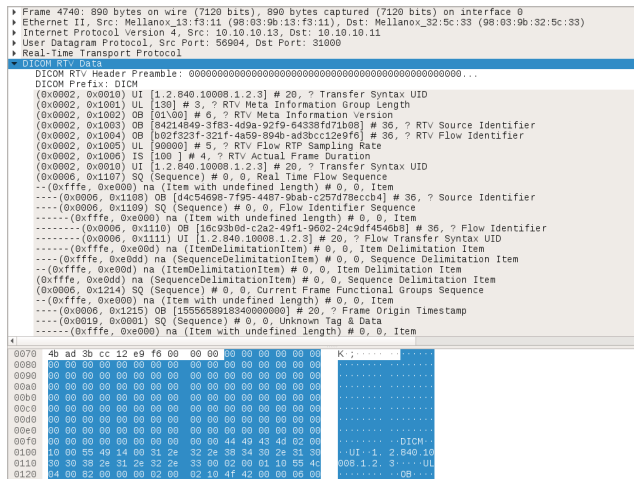


Fig. 6. Wireshark capture of a DICOM metadata packet

Secondly, as synthesized in the Fig. 5, the DICOM-RTV metadata format (explained in detail in the DICOM Supplement 202 [1]), encapsulates the metadata inside a RTP packet, as defined in the RFC3550 [7]. The Wireshark capture, shown in Fig. 6, dissects the RTP packet containing DICOM-RTV metadata sent from a Tx, and it shows that the captured DICOM-RTV packet follows the DICOM tag value format.

C. Transmitter (Tx)

The transmitter ensures mainly the conversion from the input video (i.e. SDI) to the DICOM-RTV video format and the transmission of the resulting flow over IP either unicasted or multicasted, as shown in Fig. 7. It receives also as input the metadata flow that needs to be transported with the video. The two inputs are then transferred inside their respective buffers in the named **Frame Synchronizer**. The reading of those two data flows (video and metadata) is done simultaneously at a specific instant (Alignment Point [8]) provided by the **Time Management** module of the Tx. This time is then hooked to the video and metadata essences as a timestamp information, used later for synchronization purpose, via the **Frame Synchronizer**. The video and metadata flows are then encapsulated inside a RTP packet, paced (in order to avoid traffic congestion in the network) and transmitted synchronously via UDP in the data network.

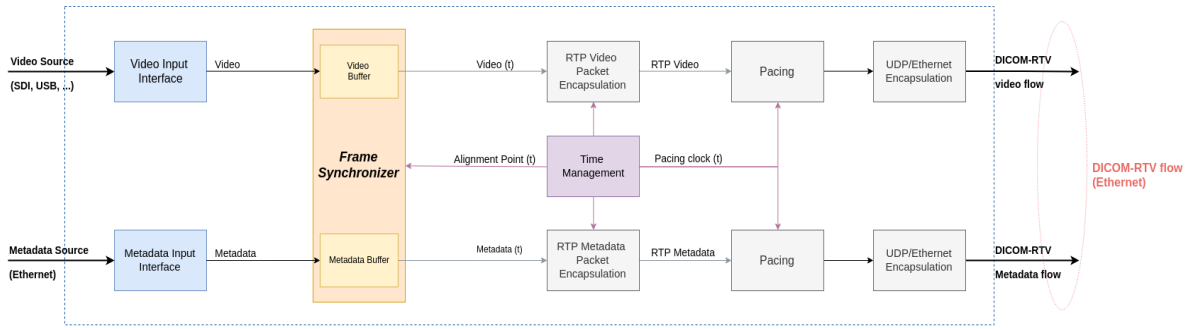


Fig. 7. Transmitter (Tx)

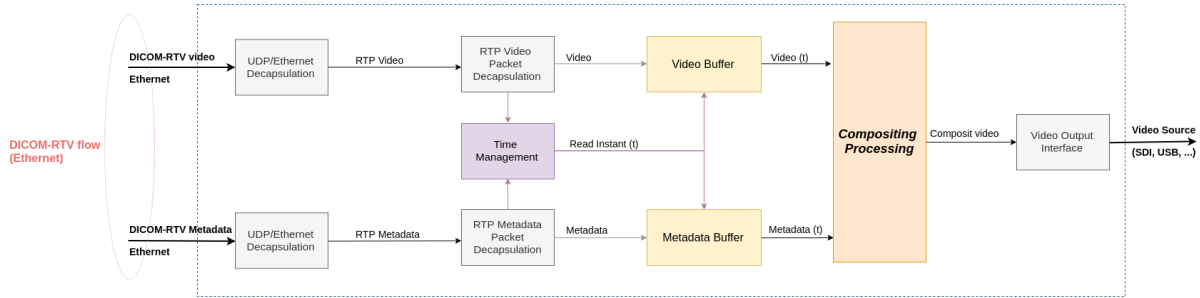


Fig. 8. Receiver (Rx)

D. Receiver (Rx)

The receiver subscribes to the DICOM-RTV flows (in unicast or multicast mode), decapsulates the RTP video and RTP metadata packets and then transfers their payload content in their respective buffers, as shown in Fig. 8. The timestamp transported in each RTP packet is processed by the **Time Management** module which defines the appropriate instant to read the video and metadata buffer. The appropriate instant is calculated from the current time, recovered from the reference clock. The **Compositing Processing** takes then as inputs the two flows, hooked with their respective timestamps, and sends them via a video output interface to the monitor (SDI, USB, etc.). The compositing processing utilizes the timestamp of each input, to synchronize accurately the incrustation of the metadata information in the video output. In the Compositing Processing, the produced image displayed on the monitor is splitted up into 9 zones (a 3x3 matrix), and the **Configurator** explicitly specifies in which zone the metadata will be positioned.

E. Configurator

The Configurator (i.e. a DICOM-RTV Configurator) allows on the one side the configuration, the control and the command of the different TxS and RxS equipment, via the configuration network, in order to enable them to exchange DICOM-RTV flows over IP. On the other side, it is interfacing with the hospital network in order to securely access its information system so as to collect the metadata related to the patients being operated in the OR (the Worklist).

Hence, the Configurator configures the TxS/RxS to transport DICOM-RTV flows. Thus, for the Tx, it specifies:

- The DICOM-RTV metadata: it provides the Tx with the **static metadata** only once at the configuration step and **dynamic metadata** continuously, that will be sent later by the Tx at the video rate. The static metadata are retrieved from the hospital's worklist and contains information related to the patient being operated. The dynamic metadata are retrieved from external sources (e.g. position of the ultrasound probe) by the configurator and relayed to the Tx.
- The video format: contains the parameters and values of the awaited SDI video flow. The Tx will raise an alert if the received flow is not identical to the configured video format.
- The network parameters: it configures the data network parameters of the Tx by specifying the IP address (e.g. the multicast address if multicast), the source and destination UDP ports, and configures the routes if not configured.

The Configurator configures the Rx as follows:

- Sends the SDP packet which specifies the video and metadata information of the awaited DICOM-RTV flow.
- Configures the data network parameters of the Rx, by specifying the reception IP address, the destination ports and configures the routes if not configured.
- Configures the composition component. It mainly configures the position of the overlaid metadata information in the image.

F. Grand Master

The Grand Master is a reference clock, implementing the Precision Time Protocol (PTP), which main role is to lock the

clock of the DICOM-RTV equipment. The recovered clock will be later used by the synchronization mechanisms, in compliance with the SMPTE ST 2059-1 standard [8], in order to synchronize a video with the corresponding metadata, or to synchronize multiple media essences.

V. RESULTS AND DISCUSSION

The latency obtained from the tests, for 66 hours of real-time transmission of the video and metadata flows from the source (e.g. camera) to the destination (e.g. a monitor) reached a maximum delay of 124 ms, and a minimum delay of 84 ms as shown in Fig. 9. The curve with the highest recurrence is nominal and is due to the frame synchronization mechanism implemented on the Tx side. The output SDI interface produces the curve with the lowest recurrence, because the SDI output frequency is not synchronized with the PTP clock. The implementation can still be further optimized. A better, but non-optimal yet, implementation, should provide a latency between 84 ms to 104 ms. The DICOM-RTV processing and the transmission of this flow takes around 30 ms between the Tx and the Rx. In our system, the remaining latency is due to the acquisition SDI cards and the SDI output cards. We argue that this remaining latency can be improved by implementing an optimized solution of the Tx and Rx.

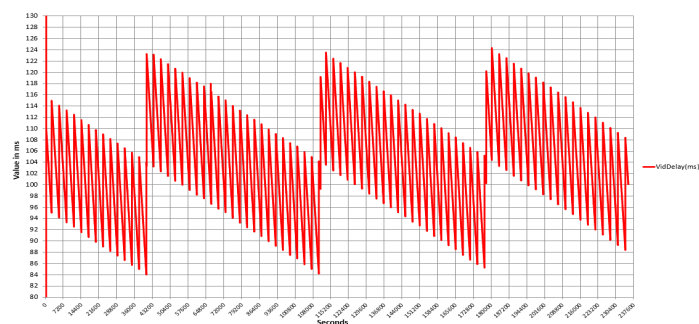


Fig. 9. Latency

Furthermore, it is worth mentioning that the proposed solution is a non-interventional trial, meaning that the worldwide premiere was not meant to change the clinical procedure. It consisted in adding an additional monitor displaying two synchronized video flows, either from an endoscope and an ultrasound, which showed the ability to help the surgeon in the navigation in the prostate thanks to the ultrasound view, or from an endoscope and an external view of surgeons hands via a camera. In the latter case, the interest is to anticipate future telementoring use cases. Dr Vincendeau is one of the two experts who have to train their colleagues on the recent prostate laser vaporization procedure. Such procedure requires a perfect coordination between two complex movements: the position of the endoscope with the left hand, typically, and the orientation and activation of the laser with the right hand. A bad synchronization or worst, activation of laser without a swiping movement, can result in a perforation of the prostate. Only a perfect synchronization of the endoscope view with

the external camera one, which is the purpose of DICOM-RTV, can highlight problem in the coordination of movements.

Moreover, security issues must be addressed, namely: 1) isolating the network to prevent any intrusion and prevent very high speed flows from disrupting the hospital's network, and 2) opening the network to allow equipment sharing between ORs and the flows transmission outside the OR, for training and remote expertise purposes. As a first step, it was decided to completely separate the networks into three private networks, one for transmitting video and metadata flows, one for configuring the equipment, and one for the linking with the hospital's information system. We argue that, generally, as long as the data remains within the boundary of the medical infrastructure, the data are kept safe.

VI. CONCLUSION

In this paper, technical details regarding the worldwide premiere implementation of the DICOM-RTV, in which an experimental setup was deployed at the Rennes University Hospital in France during five urological surgeries. The worldwide premiere demonstrated that the approach proposed by the DICOM-RTV standard is compatible with real use, and that it improves what already exists. Yet, important work still need to be done on the industrialization side, in particular to provide Tx and Rx compatible equipment in the form of compact boxes sold as converters, preparing, thus, for the integration of functions within the equipment itself (i.e., sources and monitors).

ACKNOWLEDGMENT

Special thanks to the healthcare professionals of the CHU of Rennes for their precious contributions.

REFERENCES

- [1] "Digital imaging and communications in medicine (DICOM) Supplement 202: Real-Time Video," March 2019.
- [2] "ST 2110-10:2017 - SMPTE Standard - Professional Media Over Managed IP Networks: System Timing and Definitions," *ST 2110-10:2017*, pp. 1–17, Nov 2017.
- [3] C. Mei-Yen, L. Chien-Tsai, S. Ying-Chou, L. Rheun-Chuan, C. Chian-Fa, M. T. Mu-Huo, and C. Cheng-Yen, "Design and Evaluation of a DICOM compliant Video-fluoroscopy Imaging System," in *2007 9th International Conference on e-Health Networking, Application and Services*, June 2007, pp. 248–251.
- [4] J. Documet, A. Le, B. Liu, J. Chiu, and H. K. Huang, "A multimedia electronic patient record (ePR) system for image-assisted minimally invasive spinal surgery," *International Journal of Computer Assisted Radiology and Surgery*, vol. 5, no. 3, pp. 195–209, May 2010. [Online]. Available: <https://doi.org/10.1007/s11548-009-0387-x>
- [5] M. Kyr, J. Simola, K. Haneda, K. Takizawa, H. Hagiwara, and P. Vainikainen, "Development of a channel model for 60 ghz radio systems in an angiography room," in *2011 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems*, Jan 2011, pp. 39–42.
- [6] "ST 2110-20:2017 - SMPTE Standard - Professional Media Over Managed IP Networks: Uncompressed Active Video," *ST 2110-20:2017*, pp. 1–22, Nov 2017.
- [7] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," IETF, RFC 3550, Jul. 2003. [Online]. Available: <http://tools.ietf.org/rfc/rfc3550.txt>
- [8] "ST 2059-1:2015 - SMPTE Standard - Generation and Alignment of Interface Signals to the SMPTE Epoch," *ST 2059-1:2015*, pp. 1–31, April 2015.