IBC2023

5G JAZZ JAM SESSION – AN EXPERIMENT OF COMBINED ARTS

G. Stante[1], D. Airola Gnota[1], F. Debandi[1], M. Serafini[1], L. Vignaroli[1],
F. Graglia[2], G. Scozza[2], E. Mastromartino[3], M. Fiammengo[4],
L. Guida[4]

[1]RAI – Radiotelevisione Italiana, Italy, [2]Rai Way, Italy
[3]Nokia, Italy, [4]Opnet, Italy

ABSTRACT

The present paper describes the architecture, the design, and the implementation of the "5G Jazz Jam Session" use case as part of the “5G Audiovisivo 2022” project promoted by the Ministry of Enterprises and Made in Italy. The proposed use case enables the integration and collaboration of a heterogeneous group of partners, including artists, content producers, and technology providers, with the purpose of creating a new type of artistic performance. This paper highlights the main outcomes of the experiment, the main challenges and solutions for the 5G network architecture and the TV production environment, as well as the analysis of the network performances in terms of latency, jitter, throughput, and hence the end user experience.

INTRODUCTION

In the framework of the “5G Audiovisivo 2022” project promoted by the Ministry of Enterprises and Made in Italy, the main target of the “5G TV Production – 5G Jazz Jam Session” use case was to leverage 5G network capabilities in a real remote TV production environment and analyse how 5G networks could support various scenarios where high-quality video contents are generated and distributed. In addition to this, the use case wll also analyse content distribution via 5G Broadcast technology [1].

In the field of remote TV production [2], all video contributions are produced in separate remote locations and sent via IP to the main production facility. In this respect, the introduction of content delivery over 5G networks could potentially revolutionize the typical production workflow of media companies. However, applications like the one proposed in this paper are extremely challenging as they require very low and stable delays, ultra-reliable transmissions, and very high bandwidth, in order to achieve a satisfactory result.

The goal of the considered use case is to develop a new experience of artistic performances and a fresh live TV production workflow enriched by the use of innovative audiovisual contents. This contributes to the enhancement of cultural events, through the combination of different artistic disciplines with elements of technology, such as immersive experiences of "extended reality", enabled by 5G technology. As result of the experiment, the end user has a new experience in live events and, at the same time, the media companies have the opportunity to investigate how to exploit new technological features in order to simplify content production and improve the entire television workflow.

USE CASE DESCRIPTION

The “5G Jazz Jam Session” use case aimed at the creation of a new kind of live event where the artistic performance is executed in several remote locations simultaneously but, nevertheless, it is perceived as a unique seamless show by the spectators gathered in the main venue. In particular, this specific artistic experiment involved the musical performance of a Jazz Big Band (the Big Band GP Petrini), whose musicians were displaced in three different spatially separated areas of the venue. The musicians were asked to play together by means of a dedicated 5G Private Network.

In collaboration with the dancer Oksana Romaniuk (Egri Dancing Foundation) [3] and with the participation of the sculptor Osvaldo Moi [4] that installed his sculptures in one of the event halls, the musical performance was enriched with choreographic actions and an amazing scenography. The experience was also sweetened by the story of the encounter between a dancer and a saxophonist. Music and videos (displayed on stage to the spectators) accompanied this "chase" between them.

Some innovative technology features were added to the performance. In particular, the Tilt brush [5] painter Mattia Righi (shown in Figure 1) was able to improvise a portrait of the City of Turin inspired by the rhythm of Jazz. Moreover, the event was also made available in the virtual dimension of the Metaverse to allow users to experience an "extended" event that goes beyond the boundaries of the physical space. The performance was made accessible via VR headsets in the Metaverse within a specially developed virtual space [6]. Users could meet themselves and interact as digital avatars inside a common virtual location where they were able to assist the live musical performance displayed on a screen.

Figure 1 - Tilt Brush painter

On the delivery and distribution side, the experience was also broadcasted live both outside the location, via 5G Broadcast, and inside, via 5G Broadband, to a set of mobile devices. Finally, a new 5G Broadcast receiver feature was also implemented and demonstrated during the event [1].

Figure 2 shows the three locations involved in the event, inside the OGR (Officine Grandi Riparazioni), in Turin. In particular:

- "Fucine" hall, hosting the Main Stage and 13 elements of the Jazz Big Band.
- "Duomo" hall, hosting a soloist of the band and the dancer.
- "Binario 3" hall, hosting two other soloists of the band and the tilt brush painter.

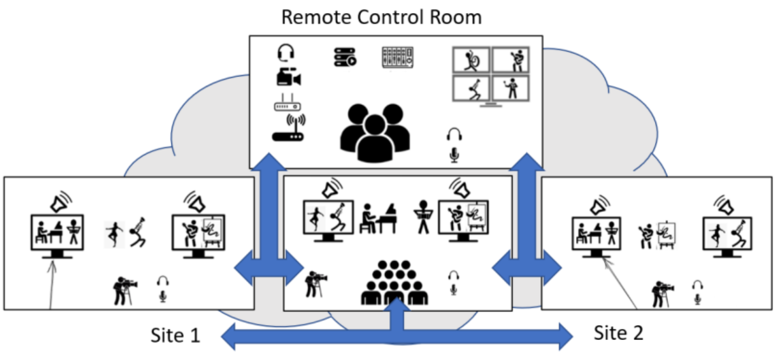

Figure 2 – UC description

In addition, the TV control room was remote and located in the R&D laboratories of the RAI CRITS, in Turin. The two facilities were connected through a 1 Gbps radio link.

Within the 5G network, different type of a/v streams had to travel concurrently. Table 1 summarizes the main characteristics:

- Video contribution streams were generated by Haivision, PRO460 backpack model [7]. These streams travelled via IP to the remote-control room (at RAI Research Centre), where they were decoded by Haivision Streamhub [8] and used for TV production.
- Video Monitoring streams were generated by Vitec encoders/decoders [9][10]. These flows were distributed from each location to the other remote ones and were used by performers to facilitate visual synchronization among musicians.
- Audio monitoring streams were of paramount importance, as the outcome of the entire event relied on them. These streams were used by each remote musician to listen to the performance from his location and to keep in sync with the rest of the band.
- Audio contribution streams were used for sound emission in the main hall and for TV production purpose. These were basically identical to the previous audio streams. The only difference was that they had a better audio quality, at the expense of a slightly higher delay.
- PGM Video was the output of the video switcher in the remote-control room and the end result of the live TV production. It was sent back to Mainstreaming local cache placed in OGR, to enable 5G Broadband trasmission via 5G towards the mobile devices inside Fucine hall equipped with Opnet SIMs.

Stream type	Number of streams	Bit Rate (Mbps)	End-to-End Latency (ms)	Codec & Protocols	Other characteristics
Video Contribution (for TV Production)	4	12	300	H265 SST	ARQ and FEC used.
Video Monitoring	6	16	50-70	H265 UDP TS	-
Audio Monitoring	4	~ 3.5	~ 21	Ravenna / AES67	SMPTE 2022-7 for redundancy
Audio Contribution (for TV Production)	4	~ 3.5	~ 31	Ravenna / AES67	SMPTE 2022-7 for redundancy
PGM Video Distribution (For Distribution)	1	~ 4	-	HLS	Broadcast of the produced content and reception on devices

Table 1 – A/V streams sent over 5G network

ARCHITECTURE

In this chapter the overall architecture is presented, including both the network architecture and the remote TV production setup.

5G network architecture

The goal in this use case was to evaluate the 5G Private SA (Stand Alone) solution for real-time transport of high-quality audio and video content, ensuring a high degree of network reliability, uplink capacity with minimal and constant latency.

The network implementation was a temporary installation, inside the venue, of two macro-blocks: a 1 Gbps radio link for data transmission in “Point-to-Point” mode, and a 5G Private Network (3400_3600 MHz, 60 Mhz band width, licensed to Opnet). Further description of the radio link is reported in the following paragraph.

The employed 5G Private Network (NPN) was based on a Nokia solution, for both Radio and Core segment, including:

- One gNB/BTS with three 5G cells enabled by three dedicated remote radio units. The cells provided full 5G RF coverage in three different halls of the event location: Fucine, Duomo and Binario 3. Figure 3 reports the venue floorplan. Each of these remote radio unit was connected via optical fiber to the Core equipment installed in a rack in the technical room, behind the stage of the Fucine room.
- One dedicated NPN 5G Core Network providing 5G services in Stand Alone mode (so without support of 4G anchoring for signalling). The NPN 5G Core was also located at the technical room and connected to a second building via radio link, hosting the TV production equipment.

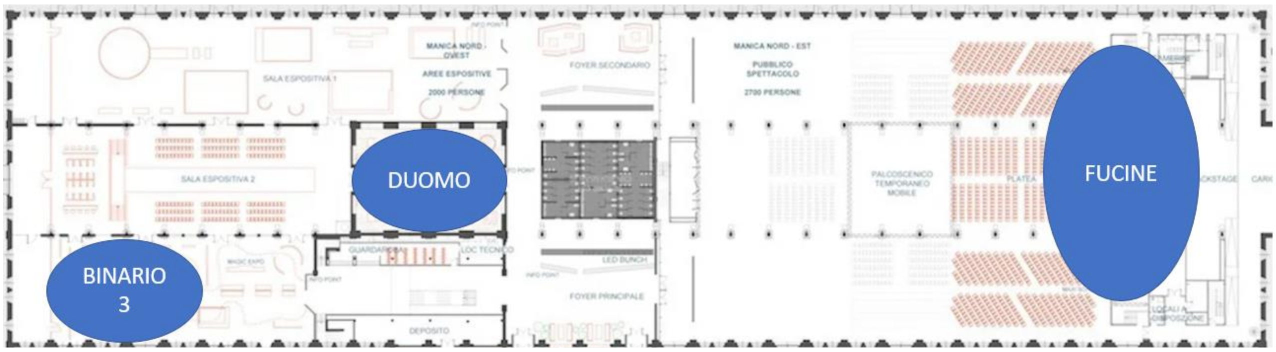

Figure 3 – OGR layout

On the user side, to interface and gather the application traffic, the following setup was used:

- Two pairs of 5G modem (one pair for each remote hall), coupled with IP/MPLS routers, for the transport of the audio/video signals.
- A set of 5G cellular devices, for the reception of the PGM video produced by the TV Control Room and sent back to the mobile devices by means of the Mainstreaming local cache.

The audio/video monitoring traffic was exchanged (in both UL and DL direction) by means of a dedicated 5G network between the main concert hall in OGR and the two remote halls. In each location the traffic was aggregated with the assistance of IP/MPLS routers that were used to create an overlay infrastructure to support the audio/video production with L2 and L3 services over 5G. The same IP/MPLS infrastructure was also used to complement the 5G network, supporting additional L2 & L3 services required by TV production, such as intercom and very-high quality fixed-camera video applications (working at very-high bitrate).

The complexity of the use case lied in the coexistence of many simultaneous live a/v streams. For each of them, ad-hoc network configurations and optimizations had to be adopted, as shown in Table 2, to achieve the expected outcome.

The 5G network was configured to support four main groups of applications: low latency audio signals (for both audio monitoring and contribution), low latency video signals (for video monitoring), video signals for TV production (video contribution coming from mobile cameras), PGM video signal (for PGM distribution to the User Equipment)"

5G network QoS was used and on-purpose tuned to support the above different applications and guarantee the best respect of the applications needs. Specific parameters of 5G ultra reliable and low-latency connection (URLLC) have been used to support the low latency audio signals.

Stream/Signal	QoS Priority	Application requirement	Network coutermeasures exploited	Performance achieved
Audio Monitoring	1 Highest	Minimum & constant delay	ST2022-7 network redundancy Highest QoS priority configuration and scheduler optimization	<20 ms for almost all packets sent
Audio Contribution (for TV Production)	1 Highest	Zero packet loss	ST2022-7 network redundancy through different 5G modem, to minimize the statistical incidence of a packet loss.	<40 ms for all packets sent
Video Monitoring	2 High	High uplink throughput, very low and costant delay, enough video quality for tv production	Radio scheme set to 6:4 and QoS tuning based on needs	Almost no video artefacts using throughput up to 25 Mbps in both DL and UL, with E2E video latency around 75 m
Video Contribution (for TV production)	3 Medium	High uplink throughput, low and costant delay, high video quality for tv production	Radio scheme set to 6:4 and QoS tuning based on needs	Almost no video artefacts using throughput up to 25 Mbps in UL.
PGM Video distribution (via BroadBand)	4 Low	No packet retransmission	Network coverage analysis and optimization	50 Mbps in Downlink within all operation areas and adjacent ones
Operation & Maintenance	4 Low	Remote control of a/v equipment	Creation of indipendent and separated virtual networks	Reliability

Table 2 - Signal coexistence within the 5G network

Network Backhaul

The OGR venue was connected through a radio link to the RAI R&D (CRITS) premise, where a remote-control room was set up.

The IP services delivered through the radio link were distributed over several physical ports, to which different priorities have been assigned, as shown in Table 3.

IP Service	Priority
mobile-camera video signals & Management	2 - high
Production Audio - Stream A	2 - high
Production Audio - Stream B	2 - high
fixed-camera video signals	3 - medium
Intercom audio	1 - Highest

Table 3 - IP service separation on the radio link

The production audio traffic, considered the most critical stream, was purposedly assigned to two different ports. This allowed the exploitation of both ST2022-7 redundancy and non-correlated channel time slots on the radio link.

At the RAI Research premises, the traffic coming from the radio link was carried to the remote-control room by a chain of fanless IP switches, so as not to disturb the audio and video control activities.

TV Production Setup & Audio/Video remote control room

Inside RAI Research facility, a dedicated technical room was setup and used to move all the noisy A/V equipment away from the remote-control room, in order not to disturb TV direction activites.

Inside the technical room, all video streams were gathered by a Sony XVS6000 video switcher, equipped with traditional SDI cards and BNC connectors. As production video format, 1080p50 was chosen for the event, in order to exploit at most the ultra-low-latency feature of Vitec encoder/decoder chains [9][10], used for visual synchronization among musicians. To work properly at 1080p50, this switcher model needs Level-A type SDI video, so all streams from the various sources were certified or converted before injection. In addition, a sync signal to all video equipment was provided and distributed by a sync generator (Tektronix TG8000).

During the event, the video mixer output has been simultaneously recorded via a Blackmagic Hyperdeck 12G player and sent to an Haivision encoder, which was used to send the PGM signal back in the event location.

Table 4 reports the list of all video sources used for TV production.

Video Source	Source Number	Position	Equipment	Codec and Bitrate
Fixed cameras	4	- 2x in Fucine - 1x in Duomo - 1x in BInario3	- 2x Sony PMW F55 camera (cameramen-controlled) - 1x Sony Venice camera (not controlled) - 1x ARRI ALEXA Mini camera (not controlled)	JPEG2K 100 Mbps
Mobile Cameras (backpack cameras)	4	- 2x in Fucine - 1x in Duomo - 1x in BInario3	- 2x Sony PXW-FX9 camera (shoulder camera) - 2x Sony PMW-F55 camera (steadycam with focus puller)	H265 12 Mbps
Control Room Camera	1	- 1 in Remote Control Room	- Blackmagic Micro Cinema Camera	JPEG2K 50 Mbps
Metaverse feed	1	- 1 stream taken from a PC in the Remote-Control Room connected to the Metaverse via the cloud	- Blackmagic Hyperdeck Studio Mini Video Output Card	-
Tilt brush feed	1	- 1 stream taken from a PC in the Binario3 hall	- Blackmagic Hyperdeck Studio Mini Video Output Card	JPEG2K 50 Mbps
RVM clips	4	- 1x in Fucine	- Black Magic Video Player	-

Table 4 – Video sources used during the event.

Regarding the fixed cameras, four video signals were received and decoded in the control room by two Lawo V_Remote 4 [11]. The same devices were used as signal loop-through to manage the conversion from Level-B to Level-A and for black burst syncing.

The four signals of the mobile cameras were encoded and transmitted on 5G network via 5G backpacks, using Haivision PRO460 model [7]. At the remote-control room, these signals were received and decoded through Haivision StreamHub [8]. The stream configuration was set with 300 ms of E2E video delay.

A Blackmagic Micro Cinema Camera was also placed in the control room, to generate a stream to be delivered to Fucine Hall led walls, making possible for the people in the control room to appear for some welcome greetings to the audience before the show.

In addition, as video source, it was also possible to feed the switcher with an Engage instance opened in a virtual world (metaverse) created ah-hoc by Impersive, through the use of a workstation connected to the cloud.

A third Lawo V_Remote 4 device received the video of Tilt Brush [5], a virtual painting in the metaverse.

Finally, a Blackmagic Videohub Clean Switch 12x12 controlled the display through some presets sent via TCP protocol. Using both the same control communication protocol and the development of an ad hoc application, it was possible to send the same command to both RVM players syncing their playout.

AUDIO APPLICATION

Since the whole use case was built over the 5G network “temporal” performance (i.e. latency and jitter), the signal distribution between locations required an AoIP implementation capable of dealing with the latency figures typical of the audio monitoring and, in general, capable of supporting the requirements of a musical live show. Among all possible AoIP solutions, an effective candidate for this role was the Ravenna/AES67 implementation offered by DirectOut® Technologies [12][13], which was finally selected for the event.

A big difference from a wired network implementation is that, as of today, 5G network lacks support for both multicast flows and Precision Time Protocol. While the former is a useful optimization to avoid replicated unicast connections, the latter is an unavoidable requirement to keep the audio location synchronized (Ravenna/AES67 requires PTPv2). The adopted solution to overcome the multicast limitation was developed by Vivivaldy®, by creating multiple Virtual Private Networks (VPNs) between each location. Broadly speaking, the Vivivaldy® solution was used to tunnel the audio streams from each audio processor, placed in each remote location, towards a Vivivaldy® Server installed in the main site, which was physically connected to the 5G core network, hence enabling the exchange of multicast Ravenna audio streams.

In the following, more details about the deployed architecture are provided. Each location was equipped with a DirectOut® Prodigy.MP processor with a Ravenna card on board for the exchange of AoIP streams, while the audio signals exchanged with the resident audio mixing consoles running the live event was made by using legacy audio formats (i.e. analog, AES3 and MADI). On the synchronisation side, each location was provided with a PTPv2 reference signal, through a complementary wired IP network (the solution used was the Meinberg Lantime 3000 connected to GPS). Therefore, every location was time-locked to a common time reference. To improve the quality of the “low latency” audio monitoring streams (the ones used as reference signal for the remote locations), a redundant network topology was adopted, in order to take advantage of the implementation of SMPTE 2022-7 standard available on the DirectOut® Ravenna cards.

During some preliminary test sessions performed with the same musicians at Nokia labs, the implemented solution via 5G network showed an one-way average latency of 16-18 ms with peaks exceeding 25ms (this value encompasses also “routing and switching delay” of the overall network infrastructure used). For this reason, an audio buffer of 1024 samples (1024samples@48kHz equals to 21.3 ms) was used for the audio monitoring streams delivered to the performers' in-ear monitor. It must be noted that in-ear monitoring was adopted for both reducing any unnecessary latency and, most importantly, avoiding any potential "echo" that could affect the musician's performance.

Following this approach, audio dropouts are anticipated. In this case, in fact, a dropout is not caused by a packet loss but by a packet, or part of it, which does not arrive in time to fill the receiving sliding buffer in the correct order. However, these signal disruptions were reckoned short enough and not overly discontinuous to allow musicians to play together.

The preliminary test session allowed to gather several feedback from musicians regarding the distributed performance.

The first feedback was that some musicians, at least in the early stages, were not familiar with playing remotely and through IEM (In-ear monitors). However, after a warm-up, they got used to the different setup and finally managed to achieve a first simultaneous musical performance.

Regarding the low-latency streams used as audio monitoring, it was observed that musicians were quite able to bear dropouts rather than latency increases. In fact, their general preference leaned towards receiving the audio with as little delay as possible (even if this would imply the presence of audio dropouts) as long as the tempo was clearly audible and recognizable. Moreover, according to their opinions, one-way delays greater than 30 ms were considered excessive and resulted in musicians being unable to play together. This threshold value depended on the specific type of song played. Songs with faster tempo required even lower latency figures to allow the band to play together.

Beside the "low latency" streams used for the performers' in-ear monitors, additional audio streams from the remote location were received and used at the main site for the specific purpose of being directly played on the P.A. system. Since the audio streams were multicast, no further bandwidth was necessary. For these "production" audio streams, a receiving buffer of 1500 samples (thus 31.25 ms) was used. Even if a bigger buffer implied a slightly extra amount of delay, it assured, nonetheless, a greater resilience against audio dropouts, which could negatively impact the audience experience. Only expert listeners reported to perceive such greater time-difference between the sound-mix coming from the main stage, and the performance coming from the remote locations.

Figure 4 shows the layout of audio streams employed during the event, overlaid with stream latencies. As described in the previous paragraphs, two different types of audio streams were configured: a low-latency audio monitoring for musicians (in red), and a high-quality audio contribution (in blue), for P.A. sound distribution and TV production purposes. Black arrows represent, instead, local audio stream carried out via cable or optical fiber.

It must be noted that:

- The remote musicians (in Duomo and Binario 3 halls) did not listen to each other.
- The remote musicians were provided with the IEM submix of the main stage, with a delay of 21 ms.

- Musicians on the main stage (Fucine) listened to the submix of all instruments, including the remote ones. The IEM round-trip delay of the remote instruments from the main stage was around 42 ms, obtained by the sum of the forward audio stream delay (21 ms) plus the return audio stream delay (21 ms).
- The audience in the auditorium listened to the high-quality audio contribution streams. Since the delay of these streams from remote locations were slightly greater (31 ms) than the audio monitoring, the delay of the remote musicians perceived by the audience was around 52 ms.
- The same audio mix played by the P.A. system was delivered to the remote-control room, where it was temporally aligned to the video signals.

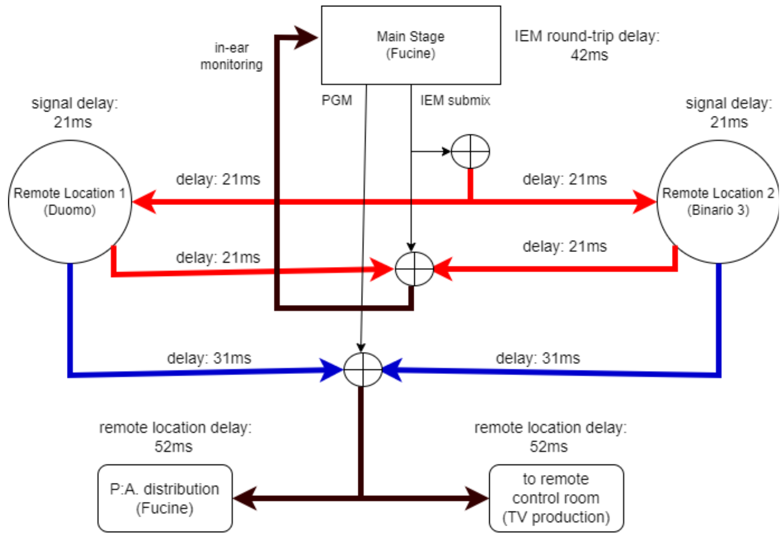

Figure 4 - audio streams with latencies

The audio streams exchanged during the live event were recorded and successively analysed. Analyses show that the 1024 samples buffer size assured minimal audio samples loss (also due to the full ST2022-7 redundancy). This is pointed out in Figure 5 which represents the audio stream from the main site (Fucine) to one remote location (Binario3).

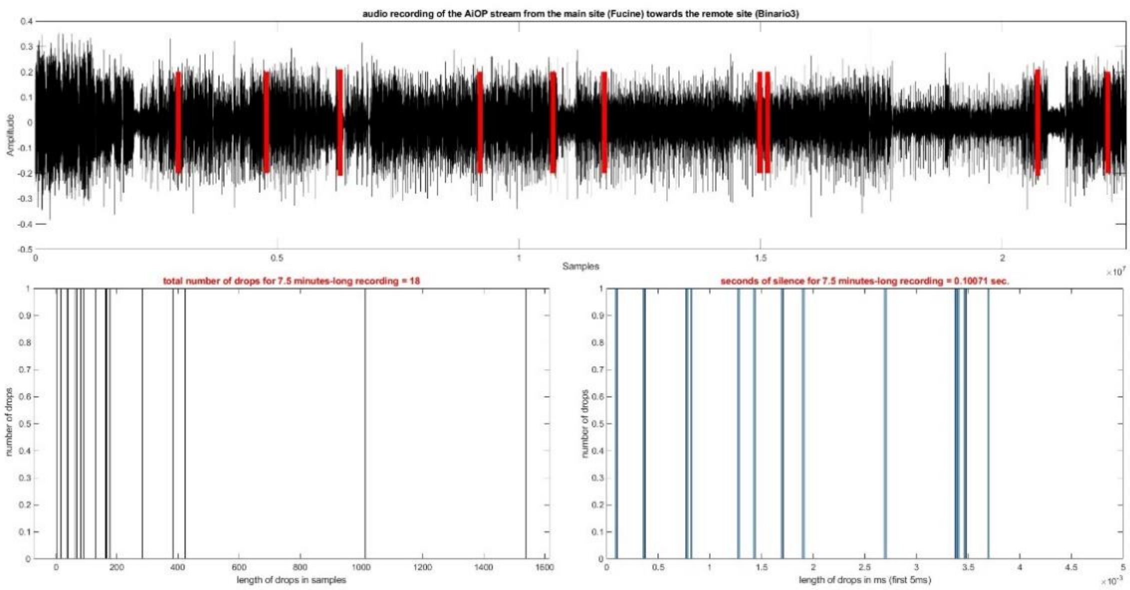

Figure 5 - positions, lengths, and numbers of audio drops within 7.5-minute audio trace.

The figure upper panel shows the exact time when audio drops (vertical bold red lines) have occurred during a piece of 7.5 minutes long song. The bottom-left panel shows a histogram representing the length, in samples, of the audio drops. It reports there were a total of 18 drops. The bottom-right frame of Figure 5 shows the same histograms using a time scale (ms). It reports that the total audio drop resulted in 0.1 second of silence for a 7.5-minute audio trace.

VIDEO APPLICATION

Diagram of video monitoring and contribution signals

As described in the previous paragraphs, there were several video signals traveling over the 5G network. Each of these was characterized by a different delay, depending on their usage.

Figure 6 shows the layout of the main video streams, overlaid with corresponding latencies. Red arrows represent the low-latency video monitoring streams (70 ms of delay) which were used as an additional visual aid for the synchronization among musicians, despite the slightly longer delay compared to the audio monitoring signal. Low-latency video signals were sent from each location to all the others, so that it was possible from each area to see the remote performance. Blue arrows, instead, represent the video contribution streams. They were characterized by a greater robustness (and therefore higher delay, up to 300 ms). The longer delay just posed a minor problem for the TV production team, which, thanks also to the preliminary analysis of the scores played by the different musicians, managed to mix the cameras correctly. These video streams were sent to the remote-control room through 5G network. Black arrows represent other contribution video streams sent to the TV production room via fixed network, without the involvement of the 5G network.

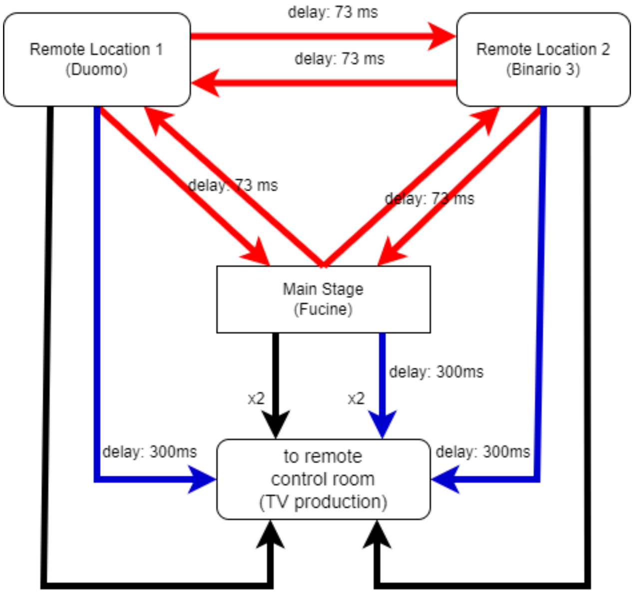

Figure 6 – video streams with latencies

Delay measurement and video quality assessement

Some video latency and quality measurements were carried out on the Vitec and Haivision encoder-decoder chains. Figure 7 shows the laboratory setup used for measuring the delay introduced by the Vitec and Haivision chains. The latency is measured by means of an oscilloscope and a photodiode. [14]

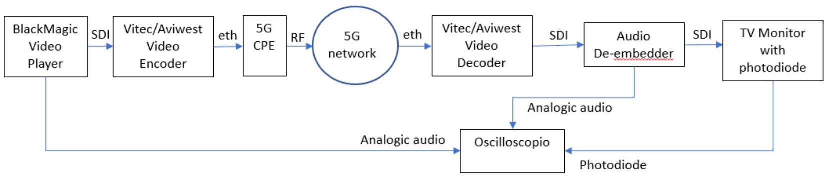

Figure 7 – Video delay measurement setup

To this purpose, an ad-hoc video was edited and employed. Figure 8 shows a frame of this video. It consisted of a short clip composed of a sequence of heterogeneous scenes. Within this clip, a second smaller video has been superimposed. The smaller video consisted in a sequence of black frames with silent audio interspersed by periodic few white light frames, at which a beep sound is emitted. Since the beep sound was synchronous with the light frames by construction, the overall latency could be measured on the oscilloscope as time difference between the photodiode response and the source analog audio coming out from the video player.

Figure 8 - A frame of the video used for latency measurement and the photodiode

Table 5 reports the results obtained for the musicians' low latency video monitoring signal, generated by the Vitec equipment. The values shown in the table also include the processing delay of the monitor used during tests, which we evaluated as 15 ms. However, TV monitor with a slightly lower processing delay (13 ms) were used during the event.

Decoder Buffer (ms)	E2E Delay (ms)	RF Link
10	52	Uplink
30	73	Uplink
50	95	Uplink
10	43	Downlink
30	63	Downlink
50	84	Downlink

Table 5 - E2E latency for musicians' video monitoring signal

End-to-end delay and video quality of the Vitec chain varies depending on the network buffer value set. A larger network buffer better accommodates network jitter, reducing packet loss. Conversely, a smaller buffer reduces overall latency resulting in higher values of packet loss, inducing possible video artefacts.

To find the best trade-off, subjective quality evaluations with duration of about 5 min were performed. They showed that a buffer value of 10 ms resulted in sporadic video artefacts. Instead, the buffer value of 30 ms showed no visible video artefacts during the observed interval. Therefore, this value was chosen for the event. Figure 9 reports the corresponding latency value measured.

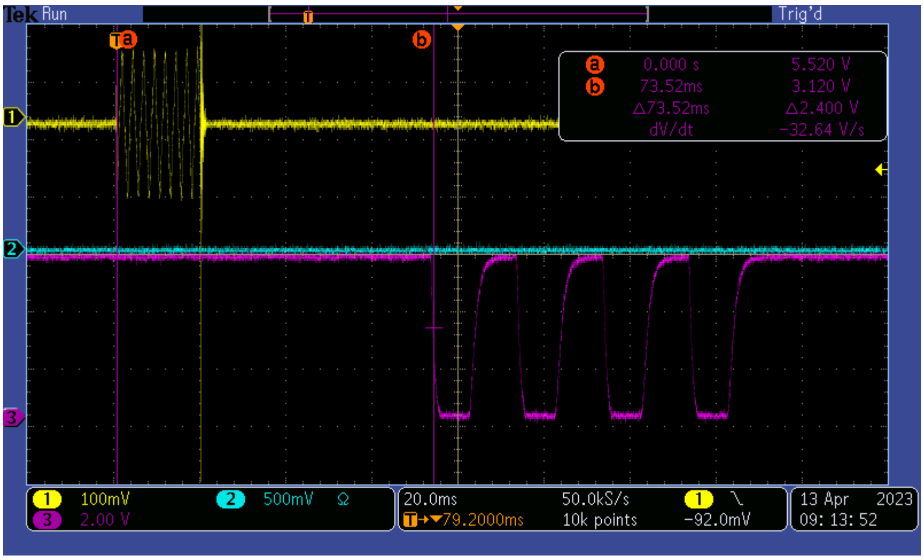

Figure 9 – E2E latency of video monitoring signals (with 30ms of decoder buffer in UL)

Regarding video contribution signals generated by Haivision equipment, similar tests were carried out in lab, to find the best operational parameters. A bitrate value of 12 Mbps was selected as a fair compromise between video quality and uplink network bandwidth. In addition, among possible transmitting protocols, we selected SST, that seemed to be more suitable in terms of latency and reliability. Haivision PRO460 allowed to set a parameter called “E2E latency”, which is an estimation of the overall expected latency. Lower E2E latency values, in our use case, implied a lower robustness of the decoded video signal. On the other hand, high values resulted in high latencies that were considered inadequate by the director’s office for the tv production. For the live event, an “E2E latency” value of 300 ms was chosen as operational value.

CONCLUSIONS

In this experiment, we managed to create a jazz musical performance between a big band placed in a main stage and musicians positioned in two remote locations.

Thanks to ultra-low latency audio/video encoders/decoders and the joint use of a private 5G network configured ad-hoc for the specific purpose, it was possible to keep the latency toward the remote musicians down to a minimum of 21 ms, while, for the latency of the remote performers perceived by the audience, it was achieved around 51 ms of delay.

The latency values obtained were considered low enough to enable the implementation of the considered use case. The ultra-low latency capabilities of the 5G network and of the audio/video equipment employed were validated by the musicians, who managed to play in a satisfactory way, without any major synchronization issues or loss of tempo. Moreover, the user experience of the audience was also good, although a trained ear might have noticed a slight delay of the remote instruments. Some participating spectators were later interviewed and all of them confirmed the perception of the performance as a unique show.

In addition to that, it was possible to create a real live TV production workflow, whose final product was also transmitted live using 5G Broadcast technology in the cities of Palermo and Turin.

Lastly, the live experience of the event was enriched through new immersive technologies, for greater user involvement, such as virtual world (metaverse) and VR painting.

In conclusion, the combined use of different typologies of arts with elements of technological novelty has resulted in a new type of artistic product, which was more comprehensive, interactive, engaging, and accessible to end users.

REFERENCES

1. Lucco Castello A., Milanesio D., De Vita A., Graglia F., Scozza G., Reggiani L., Valente P., Meroni G. 2023. Efficient delivery of audiovisual content to mobile devices combining 5G Broadcast and CDN technologes, IBC 2023 (accepted for presentation).
2. L.Vignaroli, D. Desirello, A. Trogolo, G. Sacco, R. Rahav and B. Altman. 2021. The itinerant orchestra, a 5g, multi-camera, remote and distributed video production experiment. September 2021. Technical paper IBC 2021.
3. EGRI Dancing Foundation website, available at: https://www.egridanza.com/
4. Osvaldo Moi website, available at: https://www.osvaldomoi.it/
5. Tilt Brush website, available at: https://www.tiltbrush.com/
6. Engage website, available at: https://engagevr.io/
7. Haivision webpage; data sheet: Haivision Pro460; available at: https://www.haivision.com/wp-content/uploads/haivision-pro-datasheet.pdf
8. Haivision webpage; data sheet: Haivision Streamhub; available at: https://www.haivision.com/wp-content/uploads/haivision-streamhub-datasheet.pdf
9. Vitec webpage; MGW Ace Encoder data sheet; available at: https://www.vitec.com/fileadmin/downloads/Collateral/VITEC_ACE_Encoder__Datasheet__22-H.pdf
10. Vitec webpage, MGW Decoder data sheet, available at: https://www.vitec.com/fileadmin/downloads/Collateral/VITEC_Ace_Decoder__Datasheet__22-X.pdf
11. AVC-Group webpage; data sheet: Lawo V_Remote4, available at: https://www.avc-group.com/assets/products/Lawo/pdf/lawo-br-v_line.pdf
12. DirectOut GmbH, 2020. White Paper: AoIP over WAN with DirectOut Equipment. Lockdown Rock – The show must go WAN. DirectOut Academy, June 2020.
13. Giaroli L. Webinar: Lockdown rock – The remote sessions, RAVENNA 2020 Webinar Series, DirectOut Academy, December 2020.
14. EBU Group, 2019. TR 047-Testing HDR Picture Monitors, Technical report, January 2019.

ACKNOWLEDGEMENTS

The project described in this paper was supported by the Ministry of Enterprises and Made in Italy. It is led by Rai Way and the aggregation is made up of: Rai, with its Technological Innovation and Experimentation Research Center Department (CRITS) and the Turin TV Production Center (CPTO), the Municipality of Turin, Rohde&Schwarz, OpNet, Politecnico di Milano, MainStreaming, Impersive, Kinecar, RETESETTE, La Sicilia Multimedia, in addition to the collaboration with Nokia and the Big Band GP Petrini.

A special thanks for the success of the event goes to the project coordinator Mauro D'Onofrio (RaiWay), the TV director Maria Baratta (Rai) the scultor Osvaldo Moi, the tilt brush painter Mattia Righi, the dancer Oksana Romaniuk (Egri Dancing Foundation) and the band director Gianpaolo Petrini (Big Band GP Petrini).

The authors of the present paper wish to thank all of the other project colleagues for their contribution to the work done: Mattia Santilli and Alfredo Stefano (Rai Way), Angelo Bruccoleri, Alberto Ciprian, Giancarlo De Biase, Assunta De Vita, Vittoria Mignone, Federico Maria Pandolfi, Davide Zappia (Rai CRITS), Andrea Baroni, Francesco Buonomo, Francesco Gurian, Sandro Piras, Flavio Ribelli (Rai CPTO), Stefano Luise, Pierluigi Novelli, Daniele Moro, Giorgio Ronchi (Nokia), Paolo Bossi, Cosimo Buccella (Opnet), Giacomo Meroni, Luca Scrivano (Mainstreaming), Guido Geminiani, Filippo Infantino (Impersive), Massimo Silla (Sirti).