

Latency and Reliability Analysis of a 5G-Enabled Internet of Musical Things System

Luca Turchet¹, Senior Member, IEEE, and Paolo Casari¹, Senior Member, IEEE

Abstract—The availability of high-performance embedded audio systems, along with high-bandwidth and low-latency connectivity options provided by 5G networks, is enabling the Internet of Musical Things (IoMusT) paradigm. A central component of this paradigm is represented by networked music performances (NMPs), where geographically displaced musicians play together over the network in real time. However, to date, IoMusT deployments over 5G networks remain scarce, and very limited statistical results are available on the actual latency and reliability of 5G networks for IoMusT and NMP scenarios. In this article, we present a private 5G IoMusT deployment and analyze its performance when supporting NMPs. Our IoMusT system is composed of up to four nodes and includes different background traffic conditions. We focused on the assessment of the sole wireless link, as the measurements can be easily transferred to a realistic NMP architecture involving a wide area network (WAN) by compounding them with those of the WAN. Our results show that latency increases with the number of nodes and with the presence of background traffic, whereas the reliability did not vary with the complexity of the conditions. For all tested scenarios, the average measured latency was below 24 ms (including a jitter buffer of 10.66 ms), whereas packet losses occurred with a probability of less than 0.01. However, irregular spikes were found for all latency and reliability metrics, which can significantly reduce the quality of service perceived by the users of NMP applications. Finally, packet loss and latency resulted to be uncorrelated, which suggests that they have different root causes.

Index Terms—5G networks, Internet of Musical Things (IoMusT), low-latency wireless communications, networked music performance (NMP) systems.

I. INTRODUCTION

THE FIFTH generation (5G) is the latest generation of mobile cellular networks standardized by the 3rd Generation Partnership Project (3GPP). 5G was conceived to overcome a number of shortcomings of 4G networks, while providing significantly better key performance indicators (KPIs) [1]. These include lower radio access network (RAN) latency, higher-bandwidth data communications, faster transmission scheduling through higher numerologies, as well as a more flexible core network (CN), including virtualized network functions and edge-side computation. Thanks to these

features, it is expected that the cellular connectivity provided by 5G and its KPIs might support novel Quality-of-Service (QoS)-driven applications [2].

One emerging field of application for 5G networks is networked music performances (NMPs), where geographically displaced musicians play together over the network [3], [4]. QoS is crucial to enable realistic collaborative interactions between musicians over distant locations, as the end-to-end transfer of audio information through the network must incur low latency and be very reliable.¹ These tight requirements represent a major challenge for current 4G networks, and call instead for ultrareliable and low-latency communications. 3GPP showed interest in 5G-enabled audio streaming distribution during live performances in their technical report TR22.827 [5, Sec. 5.2], which collected preliminary requirements for such a use case. While this does not necessarily mean that 5G technology is mature for such interactions, the interest in making cellular networks an enabler of live performances is likely to increase steadily.

A number of hardware- and/or software-based solutions have been developed to support NMPs, either at the commercial or at the experimental level. During the recent COVID-19 pandemic, such systems have received increasing attention and demand from professional and amateur musicians for a variety of situations including online rehearsals, performances and lessons [6]. Although the majority of them were originally conceived as software programs executable on general purpose machines, recent advancements leverage dedicated hardware platforms specifically designed to minimize audio acquisition, processing and buffering delays. Relevant examples in this space are JackTrip [7], Elk LIVE [8], LOLA [9], and fast-music [10].

The availability of high-performance embedded digital boards for audio sampling and processing, along with reliable low-latency connectivity options, is enabling the application of the Internet of Things (IoT) concept to the musical domain. This has yielded a vision for the emerging paradigm of the Internet of Musical Things (IoMusT) [11]. The IoMusT vision relates to the network of “Musical Things,” i.e., computing devices embedded in physical objects dedicated to the production and/or reception of musical content. According to this vision, future musical instruments and interfaces will embed intelligence and communications capabilities. All devices that

Manuscript received 16 February 2023; revised 17 March 2023 and 14 May 2023; accepted 19 June 2023. Date of publication 26 June 2023; date of current version 25 December 2023. This work support by the European Space Agency under Grant 4000132621/20/NL/AF. (Corresponding author: Luca Turchet.)

The authors are with the Department of Information Engineering and Computer Science, University of Trento, 38122 Trento, Italy (e-mail: luca.turchet@unitn.it).

Digital Object Identifier 10.1109/JIOT.2023.3288818

¹We remark that minimizing latency usually takes priority over reducing the bandwidth consumption of an NMP system. Hence, the use of audio compression algorithms is typically not preferred, because such algorithms introduce latency.

support NMPs are a fundamental component of the emerging IoMusT paradigm, and 5G is expected to be an enabler for it [12], [13].

While telecommunications operators roll out the first private and public deployments of 5G cellular networks worldwide, only a few (often special purpose) 5G architectures have been investigated to date for the case of NMPs [12]. Preliminary tests with early 5G hardware often target feasibility rather than an in-depth statistical analysis of the 5G network's actual latency and reliability performance for NMP applications [14], [15]. Only one study, to the best of our knowledge, has very recently focused on the long-term collection of latency and packet error traces for audio transport over 5G infrastructure [16], albeit only two musical endpoints are considered in it. Such preliminary experiments confirm that not every feature specified in 5G standards is available in state-of-the-art 5G networks: foreseeably, only the features with the most promising market viability will be implemented. Therefore, the potential of 5G cellular systems in this context remains largely unexpressed, and a systematic evaluation of the 5G network performance in realistic IoMusT and NMP scenarios remains an open research avenue. In particular, to the best of authors' knowledge, NMPs over 5G have been studied only involving two endpoints and without considering concurrent background traffic [17].

In this article, we make a further step forward in the analysis of 5G-supported NMPs by presenting a private 5G communication architecture that connects an IoMusT system of up to four nodes. Each node represents a musician; scenarios with multiple nodes reproduce well the network and traffic conditions occurring when musicians play together in a band or classic quartet. From the point of view of 5G connectivity, having all musical things closely co-located makes their transmissions more subject to interference among users and to scheduling conflicts. Thus, to the best of our knowledge, it represents a previously unseen configuration in the literature. Moreover, we apply competing background traffic that would saturate the available 5G radio resources in the absence of IoMusT communications, using both the user datagram protocol (UDP) and the TCP transport protocols. This represents a challenging if not worst-case scenario from the point of view of radio access management, and ensures that the results of our experiment cater to public network deployments, which are designed to avoid bandwidth saturation as much as possible.

In this setup, we collect latency and reliability performance metrics that help assess the feasibility of each architecture for NMPs, and perform a statistical analysis on our data.

Our work is driven by the following research questions.

- 1) Is the performance of 5G networks sufficient to support the requirements of IoMusT deployments?
- 2) Can we quantify the performance of an NMP application supported by a 5G network in terms of the packets' latency (which relates to the feasibility of the NMP itself) and reliability (which relates to the quality of the sound perceived by the musicians and audience)?
- 3) How does the performance of a 5G network supporting an NMP vary as a function of the number of IoMusT nodes and of different background traffic levels?

Our main purpose in this article is to answer the above questions using state-of-the-art technologies for NMP systems and 5G networks. To achieve this, we will rely on a state-of-the-art private standalone (SA) 5G network composed of: up to four NMP devices [8], connected into a peer-to-peer NMP system via one ZTE MC801A1 5G SA customer premise equipment (CPE) per device; an RAN composed of a ZTE V9200 baseband unit (BBU) and of a ZTE QCell R8149 antenna-based device, configured to work in the n78 3GPP band (from 3.3 to 3.4 GHz) using time-division duplexing (TDD) and a sub-carrier spacing of 30 kHz; a ZTE ZXRRAN U9003 multiaccess edge computing (MEC) server; a CN located close to the BBU and running 5G CN functions (e.g., authentication, mobility, session and user management, etc.). The antenna device, BBU, MEC server, and CN were connected through fiber-optic cables. Moreover, two extra ZTE Axon 10 Pro 5G smartphones act as sources of intervening radio traffic.

We remark that the above equipment is standard, and not modified to optimize its performance in our specific scenarios. Both the RAN and the CN hardware and software used in our experiments are the same versions available in the market. Moreover, radio access parameters are standard (e.g., up to three retransmissions, proportional-fair scheduler, and 30-kHz subcarrier spacing). The details of our considered deployments are provided in Section III.

We aim to assess whether 5G has the potential to be a fundamental enabler of the IoMusT paradigm, that will overcome the packet latency and reliability limitations of current 4G cellular networks [18]. While we are aware that the most relevant case for NMPs over 5G would be to include a wide area network (WAN) connecting the nodes, in this study we focus on the wireless access component in isolation from the performance of the WAN. Decoupling the networking performance of IoMusT devices, the 5G RAN and the CN from the performance of WANs and long-range backhauling yields more general results. In fact, we do not tie ourselves to a specific operator network topology (like in [16]) or to custom network configurations (like in [17]). Rather, we can assess the delay sources for IoMusT deployments in detail, *while measuring how much transport delay can other network components afford*. This information enables future 5G network design to account for these measurements, and make more informed choices about, e.g., how far IoMusT devices can be located to operate correctly, or which MEC server should host network functions involved in IoMusT service provisioning.

In these terms, the closest study related to our work [16] is akin to our setup, as their metropolitan link introduces an estimated delay of <1 ms. In any event, both our results and those of [16] can be easily transferred to a realistic NMP scenario with a nonnegligible WAN transport component because the measurements on latency and reliability can compound with the WAN delay contribution (and the statistics thereof) in mixed architectures using 5G and WAN.

As 3GPP continues the characterization of the QoS of multiple applications that may be supported by 5G networks [19], it is of paramount importance to substantiate whether current or future 5G architectures already support these QoS levels, or whether they need to be technically

improved and revised, or rather if support for such QoS is unlikely under the current 5G specifications [20] and a consideration for future-generation cellular architectures such as 6G. In this context, we note that the IoMusT has some characteristics in common and several differences with respect to typical applications of interest for 5G. For example, applications relying on 5G for (massive) machine-type communications (MMTCs) [21], [22], [23], [24], [25], [26], [27] typically focus on dense deployments of machine-type devices, with intermittent or erratic communication patterns. While some IoMusT deployments can be dense, especially if they involve interactions between the performers and their audience [11], IoMusT communication patterns are predictable and periodic for the whole duration of a performance.

While ultrareliable low-latency communications (UR-LLC) were designed to support fast exchanges between 5G devices with vanishing errors [28], [29], [30], [31], [32] and would thus be ideal for NMPs on paper, research on UR-LLC is still progressing. For instance, it remains unclear whether target UR-LLC error and latency figures will constitute minimum or average values, to which user density would UR-LLC apply, whether UR-LLC will best apply to episodic communications or to periodic and possibly prolonged data exchanges, and whether it would support a potentially large set of users as may appear in a typical IoMusT scenario.² Moreover, it is still under discussion how to let UR-LLC co-exist with other traffic types, such as massive IoT [29], [33] or eMBB [34], [35] and UR-LLC is still not fully supported (e.g., the European Parliament's Research Service plans UR-LLC-capable deployments not earlier than 2025 [36]), except at the level of exploratory demos [37]. In several cases, massive IoT deployments are even considered a feature of future sixth-generation (6G) networks [38]. Even the density of users foreseen for the high-performance MMTC network slice formalized in [19] may be insufficient for several IoMusT deployments involving several performers acting simultaneously and interacting with an audience.

The above discussion should clarify that the IoMusT represents a new paradigm, where a stable flow of real-time packet exchanges needs to be supported with very limited delay budgets, very high reliability, and possibly involving a large number of devices. These elements demand that the IoT community investigates the performance of real IoMusT deployments in-depth [39], [40] and that it is relevant to do so using current state-of-the-art 5G technology, starting from smaller scenarios and progressively scaling up to denser and more demanding ones. Such rigorous investigations will be instrumental to emphasize the actual achievements of current technology, as well as the technical improvements required to fully support a given service on the field. Notably, such considerations are in line with the push of industrial groups to categorize different IoT embodiments with (possibly extremely) different requirements [41].

Our study should serve as a first stepping stone to foster additional investigation on optimized 5G architectures to support NMP through 5G networks, as well as a first definition of benchmark scenarios of interest for NMP tests.

In the remainder of this article, we discuss the main service requirements for NMP systems, focusing on latency and reliability (Section II), and discuss the materials and methods of our performance evaluation (Section III). We then proceed to describe our result and findings (Section IV) before drawing some final remarks in Section VI.

II. LATENCY AND RELIABILITY REQUIREMENTS FOR NMP SYSTEMS

NMP systems aim to render the same conditions as acoustic-instrumental on-site performances. An effective remote and distributed music performance entails extremely strict QoS requirements, such as very low communication latency, low and constant jitter (i.e., the variation of latency), and high audio quality (i.e., low packet losses that generate unperceivable dropouts in the signal) [42], [43]. Therefore, audio transfer through a wireless channel must be reliable, fast, and should experience no outages. Connectivity interruptions may happen, so long as their frequency of occurrence is low enough for low-complexity error correction schemes to compensate. Such techniques include packet loss concealment methods [44], [45], [46], [47]. Satisfying these KPIs is necessary to maintain a stable tempo and to ensure a satisfactory auditory perception, thus enabling synchronicity among performers and, more generally, a high-quality interaction experience [3, Ch. 3].

In more detail, several studies have determined that the end-to-end latency that guarantees performative conditions to be as close as possible to traditional in-presence musical interactions amounts to 20–30 ms [48], [49], [50], [51], [52], [53]. Such a delay corresponds to the propagation delay of a sound wave covering a distance of 8–10 m in air. This distance is typically assumed to be the maximum displacement that different performers can still tolerate, while ensuring a stable interplay in the absence of further synchronization cues (e.g., a metronome, or the gestures of an orchestra conductor).

Reliability, in the context of NMPs, refers to the capability to guarantee successful message transmissions within a defined latency bound. There is currently no consensus on a minimum threshold value for this metric. Notably, scarce research has been conducted thus far to determine exact KPIs for reliability in NMP systems. On the one hand, this might be a consequence of conducting academic assessments of NMP systems through networks with inherently high reliability, so as to focus on the effects of latency [4], [54]. On the other hand, the definition of the term reliability in NMP contexts is still unclear [16], hindering the coherence between different experiments. In fact, the relationship between packet loss, the distribution of packet loss over time, and perceived audio quality has not been univocally determined, yet. Only a few studies have preliminarily investigated such a complex matter [55], [56]. In any event, there is consensus that consecutive packet losses cause the most harmful impact on the perceived

²IoMusT scenarios include not only large performances with several actors, but also complex interactions between performers and (not necessarily co-located) audiences.

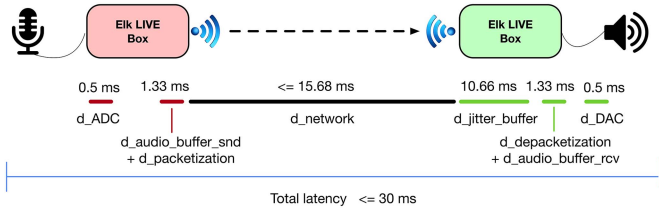


Fig. 1. Schematic representation of the components contributing to the overall latency, with the indication of the configurations utilized in the deployed 5G architecture.

audio quality, and need to be avoided as much as possible. In fact, depending on the length of the error burst, packet loss concealment methods may fail to successfully reconstruct the missing audio data.

One of the requirements of NMPs is to have a constant jitter, i.e., the latency should not fluctuate significantly, otherwise, this would negatively affect the synchronization among the musicians, and introduce artifacts, such as audio glitches in the audio stream [4]. Usually, jitter buffers at the receiver side are utilized to compensate for the varying transmission latency of individual packets. For a given network, it is possible to realize different latency budgets by selecting different sizes for the jitter buffer.

Once the jitter buffer is set in place, latency becomes constant. In more detail, the overall audio latency path from a musician acting as a sender to a musician acting as a receiver is composed as follows (see Fig. 1):

$$\begin{aligned} \mathcal{L} = & d_{\text{ADC}} + d_{\text{audio_buffer_snd}} \\ & + d_{\text{packetization}} + d_{\text{network}} \\ & + d_{\text{jitter_buffer}} + d_{\text{depaketization}} \\ & + d_{\text{audio_buffer_rcv}} + d_{\text{DAC}} \end{aligned} \quad (1)$$

where all variables represent instantaneous, time-varying values, and in particular as follows.

- 1) d_{ADC} represents the delay due to the analog to digital converter.
- 2) $d_{\text{audio_buffer_snd}}$ is the delay due to the acquisition of the signal to be sent, which is stored in an audio buffer having a size configured according to the audio host utilized.
- 3) $d_{\text{packetization}}$ represents the delay due to the packetization of the digital signal.
- 4) d_{network} is the delay determined by the transport network latency.
- 5) $d_{\text{jitter_buffer}}$ represents the delay caused by the jitter buffer used to compensate the network jitter for a sufficient number of packets, which relates to the buffer size.
- 6) $d_{\text{depaketization}}$ is the delay due to the depacketization of the signal received from the jitter buffer.
- 7) $d_{\text{audio_buffer_rcv}}$ is the delay due to the acquisition of the received signal in packets (which is stored in an audio buffer having a size configured according to the audio host utilized), as well as the mixing of such a signal with that generated by the musician using the local device.
- 8) d_{DAC} represents the delay due to the digital to analog converter.

The jitter buffer size not only affects the overall latency, but it may also affect reliability and, as a result, the perceived audio quality. Lost packets are the result of actual packet losses in the network plus late packet arrivals that the jitter buffer cannot compensate for. For instance, a jitter buffer lasting 5 ms will handle packets which are at most 5 ms later than the fastest packet, and all packets received afterwards will be lost even if they carry noncorrupted audio data. As pointed out in [16], the choice of the jitter buffer size is not trivial and needs to be carefully considered, since it trades off latency for audio quality.

An additional issue is the occurrence of bursty errors and the average length of error bursts. This is a well-known problem in wireless networks [57], and requires better statistical models than a uniform error distribution to understand the impact of error bursts on other network protocols as well as applications. For NMPs, error bursts are strongly related to the reliability, which depends not only on packet losses and on how they distribute over time but also on the time duration of packetized audio samples (hence on how many samples are included in a single packet, and thus on the sampling frequency). Realistic packet error ratios may range from 10^{-6} up to 10^{-4} , although sufficiently powerful error concealment techniques may compensate for higher ratios. However, large bursts are more likely to impair concealment algorithms [16].

A complicating factor is that, to the best of the authors' knowledge, there currently exists no widely accepted method to objectively evaluate the impact of network-based packet losses in NMP settings. The state-of-the-art in this area is the perceptual evaluation of audio quality (PEAQ), an international telecommunication union (ITU) standard conceived to measure perceived audio quality by taking psychoacoustic effects into account. However, it has been argued that PEAQ might not be appropriate for the evaluation of the impact of packet loss on perceived audio quality, as it was not designed to reflect the specific properties of networked systems [55]. Therefore, the definition of a clear reliability threshold for NMPs still represents an open research challenge.

III. MATERIALS AND METHODS

A. Apparatus

The end-to-end network implemented in our experiments was a private 5G SA network composed of three elements [58] as follows.

- 1) *User Equipment (UE)*: Any device directly employed by an end user to communicate.
- 2) *RAN*: The infrastructure that includes radio base stations (the gNBs) and bridges the connection between the UEs and the CN.
- 3) *CN*: The central part of a network that implements key connectivity services (including, e.g., authentication, security, access management, traffic shaping, slicing, and mobility management) for users connected through the RAN; moreover, the CN enables the transmission of IP packets to external networks such as the Internet.

Fig. 2 provides a schematic of the network architecture and the data flow, where we depict all components involved in

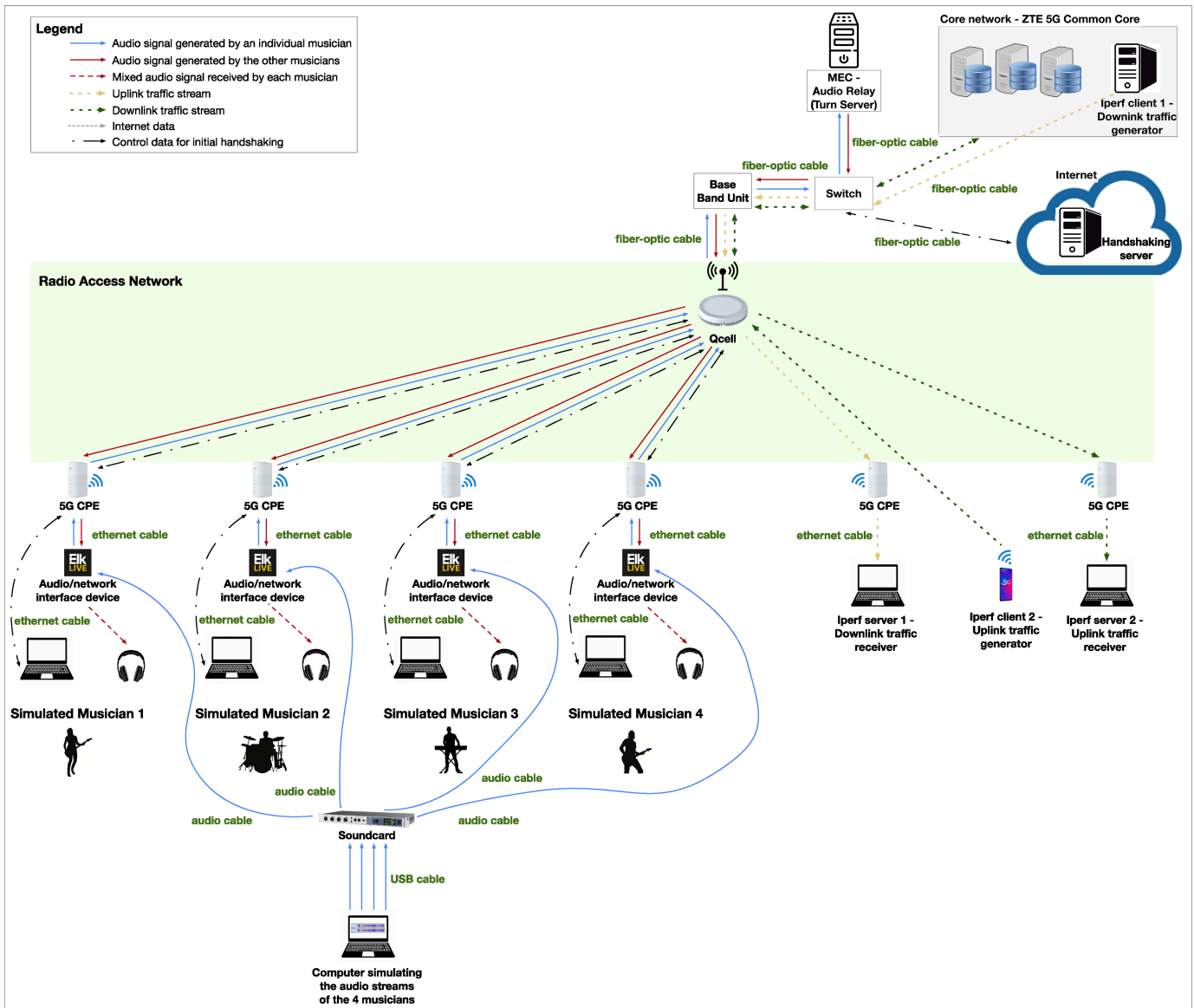


Fig. 2. Schematic representation of the deployed 5G SA architectures and the corresponding data flow.

the architecture and setup for the sake of completeness. The network was deployed in an indoor space of the ZTE Italia Innovation & Research Center (ZIRC) located in the city of L'Aquila (Italy). The base station was placed on the ceiling, about 3 m away from six UEs placed on a table (see Fig. 3). Four of the six UEs acted at the same time as the sender and receiver of audio signals. The remaining two UEs were used for the generation and reception of background traffic. The average available bandwidth was measured via ZTE's proprietary data rate metering software, yielding 1000 Mbit/s in downlink and 270 Mbit/s in uplink.

1) *User Equipment*: Each of the four UEs used for audio transfer consisted of a CPE (i.e., a 5G/WiFi/Ethernet router, and specifically a ZTE model MC801A1) connected via Ethernet to an audio/network interface device (an Elk LIVE box [8]) providing a peer-to-peer NMP system. We did not involve human subjects to perform live music. Rather, to fully

automate the measurement sessions, we simulated the audio signals they would have produced. We achieved this via an ad-hoc software coded in the pure data real-time audio programming language. The four signals corresponded to the audio recordings of four musicians playing together (electric bass, drums, keyboard, and electric guitar players) but recorded separately. The files were played back at the same time and the resulting signals were routed from a laptop to an RME Fireface UFX II soundcard. Such a laptop was not connected to any CPE, and only served the purpose of generating the audio signals.

The four audio signals travelled along audio cables from the soundcard to the input of each NMP device. Each box mixed the sound produced by one simulated performer with the sound received from the other boxes (one, two, or three depending on the experimental conditions). The resulting mix could then be heard from headphones connected to each box.



Fig. 3. Picture of the setup of the 5G SA architecture, showing the base station, the six CPEs, the four NMP devices, the four headphones, the sound card, the six laptops, and the smartphone.

The connection between the NMP devices requires a preliminary handshaking procedure, which was controlled by laptops, one for each box. This preliminary handshake was mediated by an external sever connected to the Internet, which is handled by the NMP service provider. After this initial phase, the boards were connected in a peer-to-peer fashion (no Internet routing is involved during the exchange of audio packets). With reference to Fig. 2, we remark that the TURN server on the top side of the figure only acted as a plain traffic relay, without intervening in the packet exchange.

The utilized NMP system is based on the Elk Audio OS (a low-latency audio operating system optimized for embedded systems [8]) and an ad-hoc hardware device that translates analogue audio signals into IP-packets for network transport and vice versa. The system enables deterministic processing for high-precision packet pacing and timestamping, as well as logging of received IP packet latency, jitter, and packet loss. It produces a protocol data unit comprising 64 audio samples (each sample requiring 16 bits) for each audio channel. To optimize for latency, the UDP is utilized for transport, without including any audio redundancy or retransmission-scheme at the application layer. Since two audio channels are involved, the total protocol data unit size is ≈ 272 bytes. The device works with a sampling frequency of 48 kHz, and the packet transmission rate is one packet every $64/(48 \cdot 10^3) \approx 1.33$ ms. A required data rate per box of approximately 2, 7, and 11 Mbit/s in both uplink and downlink was measured for NMP systems comprising, respectively, two, three, and four NMP devices. Therefore, the total bandwidth (for both uplink and downlink) was 4 Mbit/s in settings with two boxes, 21 Mbit/s with three boxes, and 44 Mbit/s with four boxes.

The codecs used for analog-to-digital conversion as well as digital-to-analog conversion introduced a delay of 0.5 ms each in the respective two NMP devices (d_{ADC} and d_{DAC}). The time taken for packetization ($d_{\text{packetization}}$) and depacketization ($d_{\text{depacketization}}$) was negligible. The time introduced by the audio host at the sender and receiver device ($d_{\text{audio_buffer_snd}}$ and $d_{\text{audio_buffer_rcv}}$) was related to the audio buffer utilized,

and amounted to ≈ 1.33 ms (i.e., 64 samples at sampling rate of 48 kHz). Therefore, the main delay components in the NMPs are due to over-the-air transmissions, backhaul routing, processing, as well as the jitter buffer size. The latter ($d_{\text{jitter_buffer}}$) was set to 512 samples (i.e., ≈ 10.66 ms at the sampling rate of 48 kHz). Therefore, as illustrated in Fig. 1, the deterministic delay due to the functioning of the NMP system amounted to 14.32 ms. This left a latency budget for the network transmission (d_{network}) of up to 15.68 ms in order to avoid exceeding the total latency tolerable by musicians.

The two UEs used to create extra background traffic consisted of the same CPE type as the other four UEs, and were connected to one laptop each. An additional UE consisted of a 5G-enabled smartphone (model Axon 10 Pro 5G by ZTE). The first laptop acted as a receiver for the downlink traffic generated by a server placed inside the CN. The second laptop acted as a receiver for the uplink traffic generated by the smartphone. The traffic (either UDP or TCP according to the experimental conditions detailed in Section III-B) was implemented by a server-client architecture based on the `iperf3` software for network traffic generation and performance tests.

2) *Radio Access Network*: The RAN was provided by a base station working in the 5G SA mode, which comprised an antenna-based device and a BBU. The antenna-based device (ZTE QCell R8149) received and transmitted wireless signals (5G NR) from/to the CPEs. It was configured to operate in the 3GPP frequency band n78 (from 3.3 to 3.4 GHz), using a bandwidth of 100 MHz, and a TDD configuration. The QCell was connected to a ZTE V9200 BBU via a 1-m optical cable. The BBU was connected to an MEC server (ZTE ZXRRAN U9003) through a 2-m optical cable. The MEC acted as a TURN server, i.e., as a relay of the audio traffic between the peers.

3) *5G Core Network*: The CN was located in the same building as the BBU, about 10 m apart, and connected via a fiber optic cable. We recall that we counted on a 5G SA deployment for our measurements, meaning that all signaling passes through a 5G common core, which includes the access and mobility management function (AMF), the session management function (SMF) and user plane function (UPF), respectively, with the control-plane and user-plane packet and service gateways. The proportional fair scheduler applied no packet dropping policies, and the traffic was routed without giving priority to any kind of packets.

We remark that all the hardware employed in our tests is commercially available and has *not* been modified in any ways for the purposes of this experiment. Similarly, the software employed in the ZTE cellular radio equipment is standard and has not been modified or optimized in order to run our experiments. For example, the scheduling algorithm that manages radio traffic is the proportional-fair scheduler, a de-facto standard in cellular technology to date. The above helps make our study reproducible.

B. Evaluation Procedure

We assessed the performance of the deployed architecture under different conditions, including both ideal conditions

TABLE I
TESTED EXPERIMENTAL CONDITIONS

	Number of boxes	UDP traffic	TCP traffic
Condition 1: "2 boxes"	2	✗	✗
Condition 2: "2 boxes + UDP"	2	✓	✗
Condition 3: "2 boxes + TCP"	2	✗	✓
Condition 4: "3 boxes"	3	✗	✗
Condition 5: "3 boxes + UDP"	3	✓	✗
Condition 6: "3 boxes + TCP"	3	✗	✓
Condition 7: "4 boxes"	4	✗	✗
Condition 8: "4 boxes + UDP"	4	✓	✗
Condition 9: "4 boxes + TCP"	4	✗	✓

without interfering traffic in the same cell, and in worst-case scenarios, including concurrent background traffic that saturates the available bandwidth. Table I provides a synopsis of the test scenarios. For each condition, we continuously transmitted audio for 10 min and 30 s from one endpoint to the other(s), and vice versa, while measuring the performance of the IP connection via the logging system of each NMP device. Three recordings were performed for each condition. We retrieved high-precision measurements of four performance metrics considered for the analysis as follows.

- 1) *Latency*: One-way latency in milliseconds, calculated as the round-trip time between two nodes divided by two (under the assumption that the time of the outbound and inbound communication was the same).
- 2) *Packet Loss Ratio*: The ratio between lost and transmitted packets within a given analysis window.
- 3) *Missed Packets*: The number of lost packets within the analysis window.
- 4) *Max Number of Consecutive Missed Packets*: The maximum number of consecutively lost packets within the analysis window.

Such metrics were computed on windows of ≈ 2.33 s. Each analysis window contained 1750 packets of 64 samples. We discarded the first 30 s of recording to remove any effect due to the handshaking of the devices. This led to an analysis of 450.000 packets for each box (i.e., 10 min) in each recording, leading to a total of 1.350.000 packets for each box for each condition (as there were three recordings per condition). We computed the mean, standard deviation, minimum, and maximum of each of the four performance metrics by merging the log data recorded at each box in each experiment condition.

Concerning the one-way latency measurement, this included the actual delay introduced by the network as well as the contribution due to the jitter buffer (i.e., ≈ 10.66 ms). The round-trip latency computation is achieved by associating to every transmitted packet a time stamp and a sequence number. Once the receiving node receives a packet from the sender, it piggybacks information about the received packet into the next outgoing audio packet. This information is then used by the original sending node to compute the round-trip latency. From this measure the one way latency is computed dividing by two. Different from other NMP systems (e.g., [16]), the one involved in the present study does not necessitate

extra hardware (e.g., GPS) or a shared clock to synchronize the involved nodes: synchronization is carried out via statistical inference, thanks to a patent-pending algorithm of the system manufacturer. Notably, the latency measurement is robust because it is carried out systematically for all transmitted packets, resulting in a large number of data points. The measurement method is not fully accurate for measuring the one-way latency, but it is surely an optimal tradeoff that is possible to achieve in the absence of a shared clock. However, the method is highly accurate when measuring the round-trip latency.

IV. RESULTS

Table II shows the results concerning the considered statistical measures on the four metrics (latency, packet loss ratio, missed packets, and maximum number of consecutive missed packets) for all conditions. An analysis of variance (ANOVA) was performed on different linear mixed effect models, one for each metric. Specifically, each model had the metric and condition as fixed factors, and the NMP device as a random factor. Post hoc tests were performed on the fitted model using pairwise comparisons adjusted with the Tukey correction.

Regarding the analysis on latency, a significant main effect was found for factor condition ($F(29183) = 1488.3$, $p < 0.001$). The post hoc tests revealed that the latency was lower for condition 1 compared to conditions 2 and 3, as well as conditions 4 and 7; it was lower for condition 4 compared to 5 and 6; it was also lower for condition 7 compared to conditions 8 and 9; all comparisons were significant at $p < 0.001$. These results indicate that latency significantly increased with the number of boxes (without traffic), and the addition of traffic (both UDP and TCP) significantly increased the latency compared to any conditions where the boxes constituted the only sources of traffic.

Fig. 4 uses box plots to convey the mean and standard deviation of the latency distribution over all conducted experiments. Triple asterisks connect conditions for which comparisons are relevant at $p < 0.001$. The numbers show the two main trends discussed above, whereby an increasing number of boxes or the presence of background traffic (regardless of whether the traffic is TCP or UDP) contributes to increasing latency.

For the most saturated case with four boxes, we also observe the expected result that TCP connections are more lenient toward the UDP traffic from the NMP devices. Conversely, background UDP traffic does not pose any limit on the transmit rate, and causes a higher latency increase.

As far as packet loss ratio, missed packets, and maximum number of consecutive missed packets are concerned, no significant main effect was found. Fig. 5 illustrates the bar plots for the three above metrics, showing that all of them are approximately independent of the experimental conditions. Because the RAN resources were saturated in the presence of background traffic, we observe that the proportional fair scheduler used in the 5G RAN allocates a fair amount of bandwidth to all UEs (including those offering non-NMP traffic), and errors span both desired audio packets and background traffic.

TABLE II
MEAN, STANDARD DEVIATION, MINIMUM, AND MAXIMUM OF THE FOUR
INVESTIGATED METRICS FOR EACH CONDITION. THE UTILIZED NMP
SYSTEM INVOLVED A JITTER BUFFER WITH SIZE OF ≈ 10.66 MS

	Mean	SD	Min	Max
CONDITION 1: "2 boxes"				
Latency (ms)	22.27	0.58	21.08	27.42
Packet loss ratio	0.0055	0.0065	0	0.0582
Missed packets	9.7	11.53	0	102
Max number of consecutive missed packets	2.63	8.14	0	93
CONDITION 2: "2 boxes + UDP"				
Latency (ms)	22.78	0.57	22.05	26.28
Packet loss ratio	0.0057	0.0092	0	0.056
Missed packets	10.03	16.1	0	98
Max number of consecutive missed packets	3.16	10.35	0	80
CONDITION 3: "2 boxes + TCP"				
Latency (ms)	23.68	0.61	23.02	28.93
Packet loss ratio	0.006	0.0081	0	0.046
Missed packets	10.63	14.32	0	82
Max number of consecutive missed packets	3.5	10.28	0	78
CONDITION 4: "3 boxes"				
Latency (ms)	22.36	0.51	20.85	25.84
Packet loss ratio	0.006	0.0076	0	0.0685
Missed packets	10.63	13.44	0	120
Max number of consecutive missed packets	2.95	7.71	0	93
CONDITION 5: "3 boxes + UDP"				
Latency (ms)	22.75	0.42	21.96	24.61
Packet loss ratio	0.0061	0.0075	0	0.0394
Missed packets	10.81	13.15	0	69
Max number of consecutive missed packets	2.33	3.75	0	43
CONDITION 6: "3 boxes + TCP"				
Latency (ms)	22.83	0.57	21.71	27.33
Packet loss ratio	0.0061	0.0083	0	0.0662
Missed packets	10.67	14.56	0	116
Max number of consecutive missed packets	2.83	7.77	0	99
CONDITION 7: "4 boxes"				
Latency (ms)	22.62	0.5	21.11	28.48
Packet loss ratio	0.0062	0.00711	0	0.0697
Missed packets	10.94	12.44	0	122
Max number of consecutive missed packets	3.3	7.96	0	102
CONDITION 8: "4 boxes + UDP"				
Latency (ms)	23.21	0.51	21.82	27.31
Packet loss ratio	0.0064	0.0074	0	0.0782
Missed packets	11.2	13.06	0	137
Max number of consecutive missed packets	2.87	5.93	0	94
CONDITION 9: "4 boxes + TCP"				
Latency (ms)	23.02	0.43	21.80	26.57
Packet loss ratio	0.0064	0.0075	0	0.0862
Missed packets	11.21	13.2	0	151
Max number of consecutive missed packets	3.18	7.38	0	99

From this result, we conclude that 5G NR radio communications were correctly configured to balance traffic, that the scheduler worked properly by allocating a fair amount of

bandwidth to all UEs, and that errors spanned both desired audio packets and background traffic. The latter was proven by visual inspection of the logs of the `iperf3` application during the experiments and showed expected patterns (e.g., more frequent losses when background traffic operates along with a larger number of boards, or if UDP is used).

Fig. 6 shows the evolution of the four performance metrics over time (10 min), recorded at one of the boxes, for the conditions *4 boxes* (left panels), *4 boxes + UDP traffic* (middle) and *4 boxes + TCP traffic* (right), respectively. These were the most complex conditions investigated, as they involved all boxes as well as the saturating traffic streams. We observe latency peaks as well as bursts of consecutively missed packets. This situation, however, is common to all conditions, and statistical analysis yielded no significant differences.

Fig. 7 illustrates the cumulative density function of the four metrics for the conditions *4 boxes*, *4 boxes + UDP traffic* and *4 boxes + TCP traffic*. Concerning latency, we observe that for condition *4 boxes* 99.3% of the packets incur a delay of 24 ms or less, while for condition *4 boxes + UDP traffic* and *4 boxes + TCP traffic* the percentages are at 92.8% and 98.1%, respectively. Regarding missed packets, the figure shows that for condition *4 boxes* 99% of the lost packets amount to up to 62, while for condition *4 boxes + UDP traffic* and *4 boxes + TCP traffic* the number is 66 and 62, respectively. As far as the maximum number of consecutive missed packets is concerned, for condition *4 boxes* 99% of the bursts amount to up to 44 packets, while for condition *4 boxes + UDP traffic* and *4 boxes + TCP traffic* the number is 27 and 32, respectively.

We searched for possible correlations between latency and the other three measures in all conditions' results (grouping the results for all boxes in the same condition). For this purpose we utilized Pearson's correlation tests. For all sessions we identified significant correlations at $p < 0.01$, but their strength was always weak (up to $r < 0.3$).

V. DISCUSSION

Regarding latency, the results of our tests showed that the implemented IoMusT system guaranteed, in all experimental conditions, the latency requirements needed to ensure a realistic musical interplay (i.e., 30 ms). The measured end-to-end latency was below 24 ms on average, and never exceeded 29 ms. We achieved such delays through proper configuration of ZTE's equipment to work at 5G NR numerology 1 in the n78 (3.3–3.4 GHz) band, implying a subcarrier spacing of 30 kHz and radio frame length of 500 μ s. We remark that, per the discussion in Section I, we did not resort to an explicit UR-LLC setup. In any event, we observed no exceeding delays across the RAN. Instead, we remark that the CN was optimized for uplink/downlink communications, and not for peer-to-peer communications occurring across the NMP devices. Delays within the CN itself can thus happen to be large due, e.g., to the multiple interrogations of UE location registers before forwarding traffic in downlink. Optimizing these aspects is part of our future work.

As far as reliability is concerned, packet losses occurred with a probability of less than 10^{-2} on average, with irregular

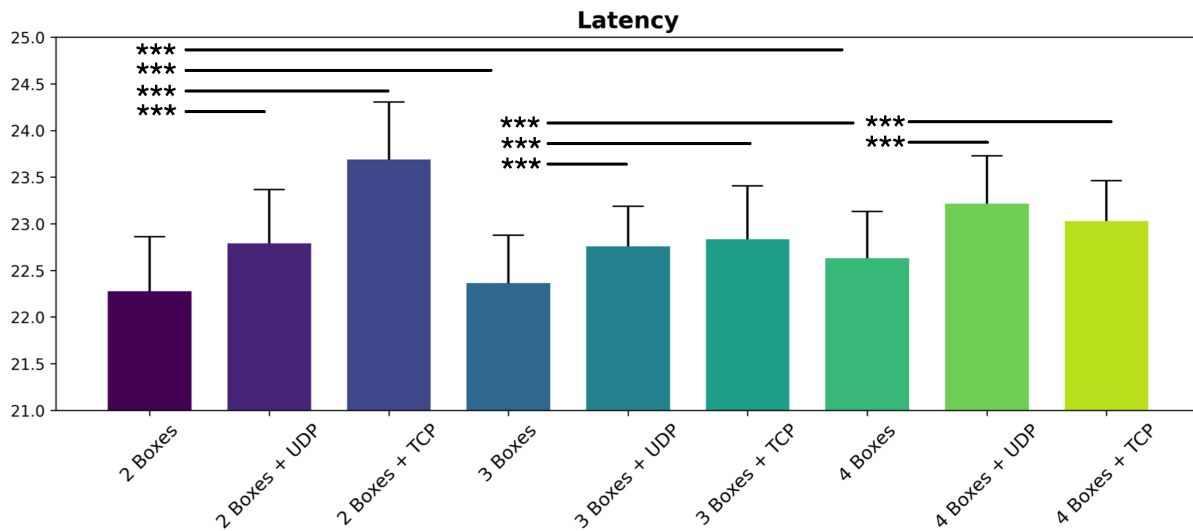


Fig. 4. Mean and standard deviation of the latency for all conditions, with indication of the relevant statistically significant pairs. Legend: *** = $p < 0.001$. The sole network latency can be retrieved by subtracting the duration of the jitter buffer (≈ 10.66 ms) from the reported data.

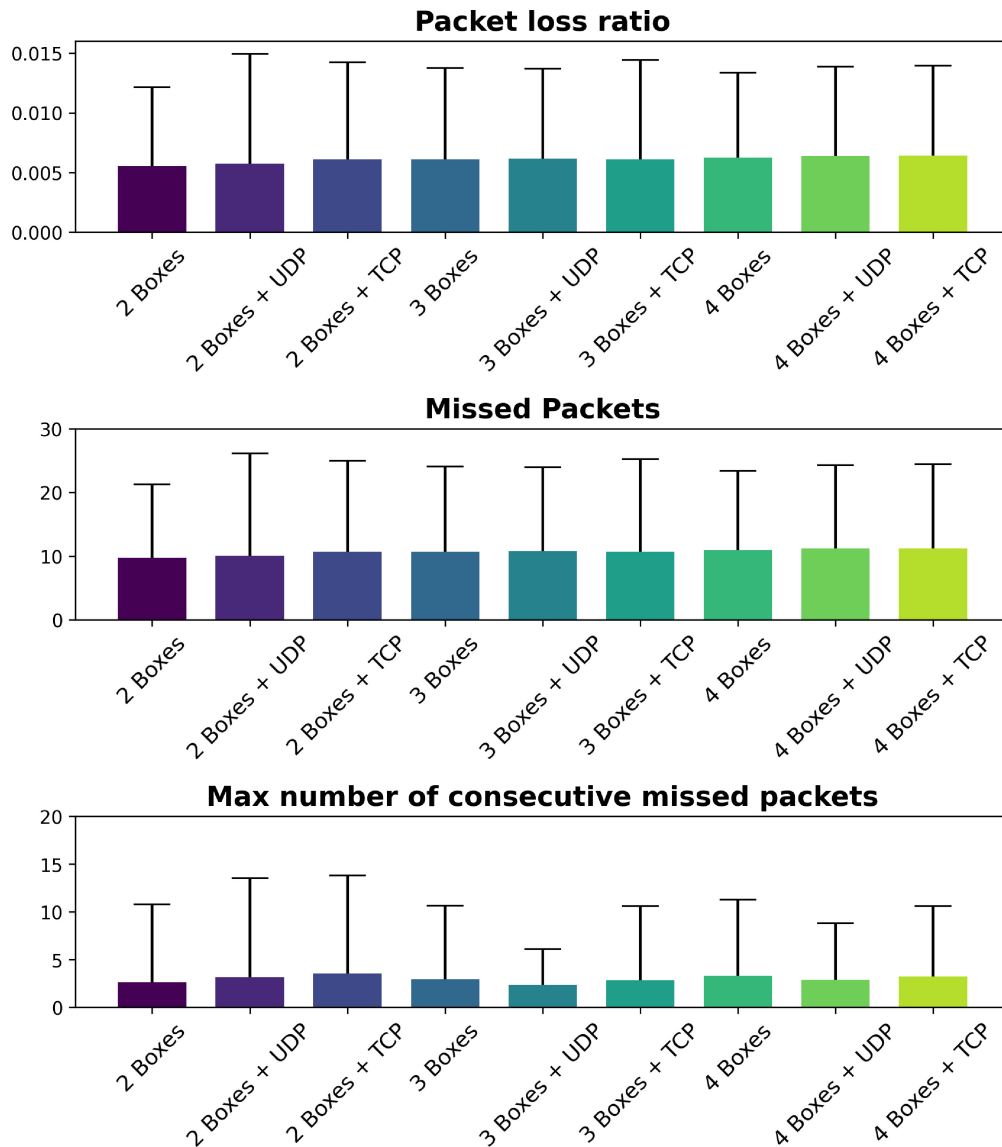


Fig. 5. Mean and standard deviation of packet loss ratio, missed packets, and maximum number of consecutive missed packets for all conditions.

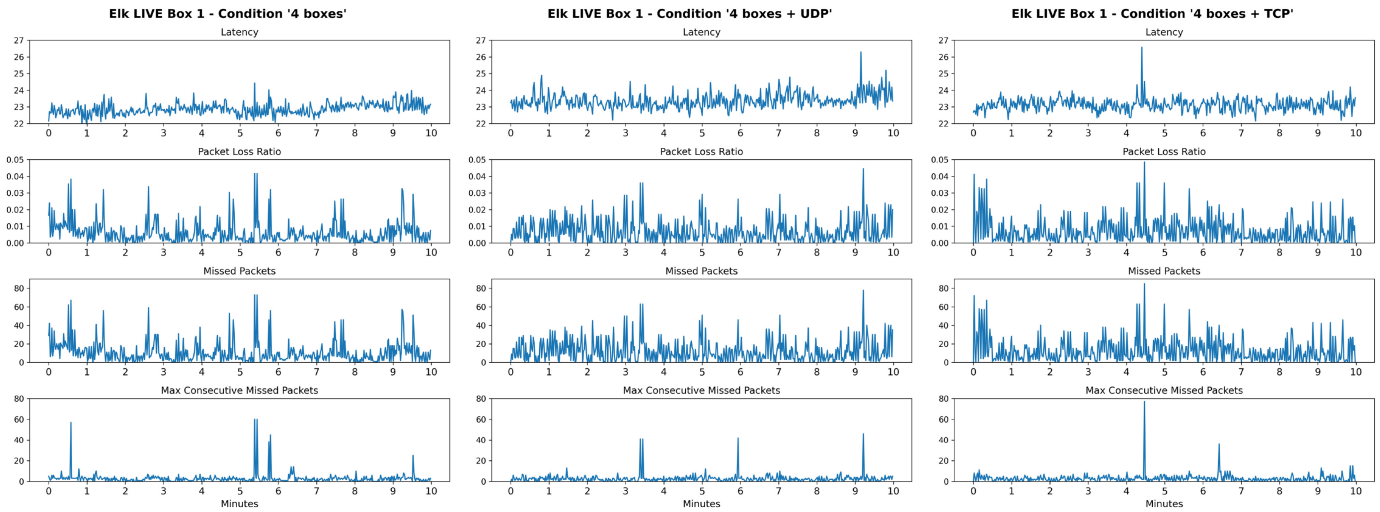


Fig. 6. Evolution of the four performance metrics over time (10 min), recorded at one of the boxes, for the most complex conditions: 4 boxes (left), 4 boxes + UDP traffic (middle), and 4 boxes + TCP traffic (right).

bursts of up to 151 consecutive packet losses in some cases. This data requires the adoption of efficient packet error concealment methods able to reconstruct (with no additional latency) the parts of the audio signals that are missing. Retransmission mechanisms or audio redundancy schemes could also be set in place, although they were not activated for the presented experiments.

From Fig. 6, we observe that latency and packet loss significantly fluctuated over time (see the latency spikes and packet loss bursts). However, our in-depth investigations suggested that packet loss may not be correlated to latency. This may indicate that packet loss and latency originate from different network operations, e.g., that latency is not necessarily due to loss recovery attempts via retransmissions at the radio link level. This result is in accordance with the findings reported in [16] for a public 5G SA network involving two nodes.

To investigate the source of packet loss in more depth, we performed measurements directly on the ZTE equipment, yielding a block error ratio (BLER) of about 0.08. This figure is fully in line with 5G specifications, but slightly higher than the expected BLER of 0.05 or less from proprietary ZTE trials. The reason for this discrepancy is attributed to the configuration of 5G transport blocks for throughput maximization instead of resilience against interference. We also conjecture that other ZTE QCells deployed in the ZIRC area and operating in the same band could sporadically cause interference to the QCell used in our experiments.

Notably, our results showed that latency increased with the number of nodes and with the presence of background traffic. Nevertheless, the metrics related to reliability did not significantly vary with the complexity of the conditions. The Elk Live NMP system uses 44 Mbit/s in both downlink and uplink when four nodes are involved. This value is much smaller than the available bandwidth (1000 Mbit/s in downlink and 270 Mbit/s in uplink). The network dynamically adapts the radio resources to be allocated to all connected nodes according to the proportional fair scheduling principle, and can therefore support the NMP service also in the presence of congestion due to

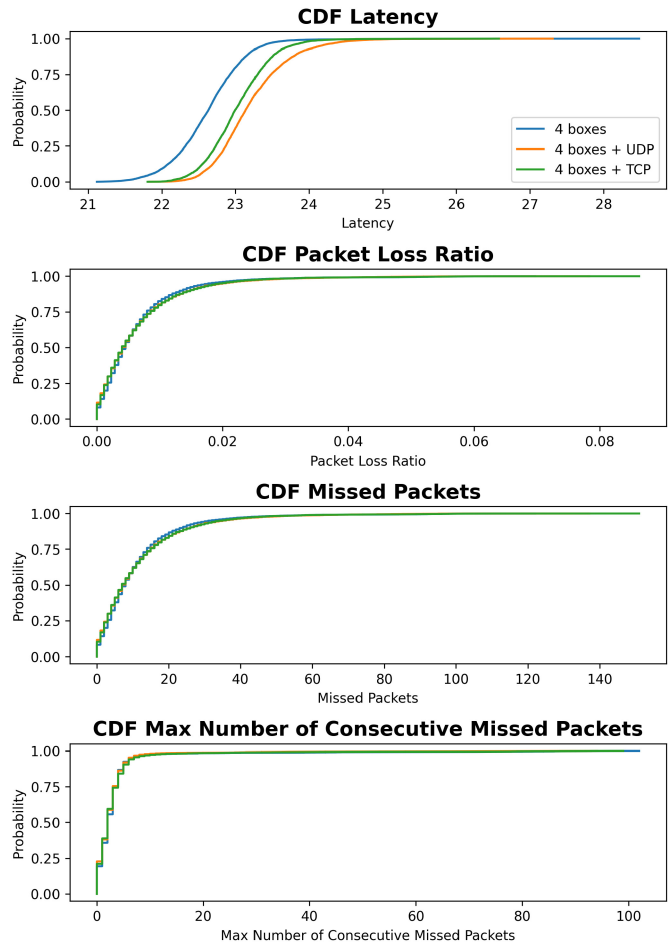


Fig. 7. Cumulative density function for all metrics in the most complex conditions: four boxes, four boxes + UDP traffic, and four boxes + TCP traffic.

background traffic. This adaptation, however, also causes a small latency increase.

In our deployment, we configured the jitter buffer to a size corresponding to 10.66 ms. However, by looking at the subplots in Fig. 7, the buffer size could have been increased to

15 or 16 ms and still yield a total latency lower than the 30 ms threshold recommended for NMP (see Fig. 1) in most of the cases. This increase would have enabled the inclusion of several packets that were otherwise discarded in our measurements, since they arrived after the maximum delay allowed by the jitter buffer). Nevertheless, this would not have affected the reliability in terms of packet radio losses.

It is worth noting that our study involves a scenario with four NMP endpoints co-located in the same room and connected to the same base station. From the point of view of radio access performance and mutual interference, this likely represents a worse case than a typical NMP deployment, where performers are distributed across a larger metropolitan area, and possibly served by different gNBs or by different sectors of the same gNB [12]. Moreover, we did not consider the case where a WAN, such as the Internet bridges multiple endpoints. The presence of a WAN would have limited the duration of the jitter buffer for long distances (so as to satisfy the overall 30-ms latency requirement), and would have affected the reliability performance.

Our results provide various insights for the design and the configuration of an NMP system involved in a 5G IoMusT deployment. First, they indicate the need for retransmission mechanisms when dealing with a wireless link with a broad available bandwidth but yielding suboptimal reliability, such as the one encountered in our experiments. Second, an efficient packet loss concealment algorithm (working at zero latency) is required, especially to deal with consecutive lost audio packets. Such algorithm could be placed not only at the receiver side but also on the MEC. Third, our findings suggest that budgeting a sufficient transport delay over a WAN requires further progress with the design of the RAN hardware, including support for higher numerologies, which contribute to reducing radio access and transmission latency. Fourth, the presence of concurrent traffic (especially if intense, such as the one generated in our experiments) contributes to increasing end-to-end delays. The 3GPP 5G standard provides the concepts needed for a design able to support such performances: in addition to the use of an MEC, slicing mechanisms would allow to decouple resources allotted to NMPs from those allotted to other types of traffic. However, further measurements are needed to precisely identify the requirements of a 5G slice for musical interactions.

VI. CONCLUSIONS

This article presented and evaluated a 5G-based IoMusT system designed to support NMPs. Our setup included up to four musicians, and provides a more realistic situation than the scenarios involving two networked musical devices typically investigated in the NMP literature. Our evaluation focused on the latency and reliability of digital audio packet exchanges over the 5G network, which are KPIs for NMP quality-of-experience requirements. In particular, the network performance was assessed both in ideal conditions and in worst-case conditions, i.e., respectively, without and with background TCP and UDP traffic contending for RAN resources against audio traffic.

Our results revealed that latency proportionally increased with the number of nodes and with the presence of background traffic, whereas reliability metrics did not vary with the complexity of the conditions. In particular, the average latency was below 24 ms for all conditions, whereas packet losses occurred on average with a probability of less than 10^{-2} . The presence of sporadic spikes was observed for all latency and reliability metrics. Latency peaks and, especially, long bursts of consecutive lost packets represent problematic situations for the strict QoS requirements that need to be ensured for NMPs. Packet loss ratios resulted to be uncorrelated with latency: this indicates that they originate from different causes. In the considered experimental conditions, the CN seemed to impose significant transit delays to audio packets. We are collaborating with the ZIRC research center to relieve such delays, and improve the speed of peer-to-peer communications among networked musical instruments.

A continuous stream of reliable and low-latency communications, such as those needed for NMPs, are challenging to be supported by the fifth generation of mobile networks. This type of QoS is vastly different from that of traditional mobile broadband applications. Our findings suggest that current 5G network designs need to improve in terms of latency and reliability in order to properly support NMPs, especially when involving a WAN between the end users. The 5G standard has provisions for dedicated slicing and MEC mechanisms that are yet to be properly explored for the case of musical interactions. Future investigations toward these directions could prove that 5G is fully capable of supporting NMPs and, as a consequence, that it is a fundamental enabler of the IoMusT paradigm.

ACKNOWLEDGMENT

The authors are indebted with Daniele Ciambone and Domenico Puntillo (ZTE), and Stefano Zambon (Elk), for their excellent and timely technical support.

REFERENCES

- [1] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.
- [2] C. Rinaldi, F. Franchi, A. Marotta, F. Graziosi, and C. Centofanti, "On the exploitation of 5G multi-access edge computing for spatial audio in cultural heritage applications," *IEEE Access*, vol. 9, pp. 155197–155206, 2021.
- [3] L. Gabrielli and S. Squartini, *Wireless Networked Music Performance*. Singapore: Springer, 2016.
- [4] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An overview on networked music performance technologies," *IEEE Access*, vol. 4, pp. 8823–8843, 2016.
- [5] "Study on audio-visual service production; stage 1 (release 17)," 3GPP, Sophia Antipolis, France, Rep. TSG-SA 22.827, Rev. 17.1.0, Dec. 2019.
- [6] N. R. Fram, V. Goudarzi, H. Terasawa, and J. Berger, "Collaborating in isolation: Assessing the effects of the Covid-19 pandemic on patterns of collaborative behavior among working musicians," *Front. Psychol.*, vol. 12, Jul. 2021, Art. no. 674246.
- [7] J. Cáceres and C. Chafe, "JackTrip: Under the hood of an engine for network audio," *J. New Music Res.*, vol. 39, no. 3, pp. 183–187, 2010.
- [8] L. Turchet and C. Fischione, "Elk audio OS: An open source operating system for the Internet of Musical Things," *ACM Trans. Internet Things*, vol. 2, no. 2, pp. 1–18, 2021.
- [9] C. Drioli, C. Allocchio, and N. Buso, "Networked performances and natural interaction via LOLA: Low latency high quality A/V streaming system," in *Proc. Int. Conf. Inf. Technol. Perform. Arts, Media Access, Entertain.*, 2013, pp. 240–250.

- [10] A. Carôt, C. Hoene, H. Busse, and C. Kuhr, "Results of the fast-music project—Five contributions to the domain of distributed music," *IEEE Access*, vol. 8, pp. 47925–47951, 2020.
- [11] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, "Internet of Musical Things: Vision and challenges," *IEEE Access*, vol. 6, pp. 61994–62017, 2018.
- [12] M. Centenaro, P. Casari, and L. Turchet, "Towards a 5G communication architecture for the Internet of Musical Things," in *Proc. IEEE Conf. Open Innov. Assoc. (FRUCT)*, 2020, pp. 38–45.
- [13] F. Cheli and S. Giordano, "Service parameters identification for adaptive networked music performance," in *Proc. Global Inf. Infrastruct. Netw. Symp.*, 2022, pp. 94–98.
- [14] "Remote music performance by Ericsson and MIND music labs at mobile world congress 2019." Oct. 2022. [Online]. Available: <https://elk.audio/remote-music-performance-by-ericsson-and-mind-music-labs-at-mwc19/>
- [15] "Project 5G-TOURS deliverable D4.2: First touristic city use case results." Oct. 2022. [Online]. Available: <http://5gtours.eu/documents/deliverables/D4.2.pdf>
- [16] J. Dürre et al., "In-depth latency and reliability analysis of a networked music performance over public 5G infrastructure," in *Proc. Audio Eng. Soc. Convent.*, 2022, pp. 1–10.
- [17] A. Carôt, M. Dohler, S. Saunders, F. Sardis, R. Cornock, and N. Uniyal, "The world's first interactive 5G music concert: Professional quality networked music over a commodity network infrastructure," in *Proc. Sound Music Comput. Conf.*, Jun. 2020, pp. 407–412.
- [18] L. Chettri and R. Bera, "A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 16–32, Jan. 2020.
- [19] *System Architecture for the 5G System (5GS)*, 3GPP TS 23.501, Feb. 2023. [Online]. Available: <https://www.3gpp.org/DynaReport/23501.htm>
- [20] A. Weissberger. "Another opinion: 5G fails to deliver on promises and potential." Feb. 2023. [Online]. Available: <https://techblog.comsoc.org/2022/12/10/another-opinion-5g-fails-to-deliver-on-promises-and-potential/>
- [21] R. F. Sari, R. Harwahu, and R.-G. Cheng, "Load estimation and connection request barring for random access in massive C-IoT," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6539–6549, Jul. 2020.
- [22] F. Jabbarvaziri, N. M. Balasubramanya, and L. Lampe, "HARQ-based grant-free NOMA for mMTC uplink," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8372–8386, May 2021.
- [23] L. Miuccio, D. Panno, and S. Riolo, "Joint control of random access and dynamic uplink resource dimensioning for massive MTC in 5G NR based on SCMA," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 5042–5063, Jun. 2020.
- [24] M. R. Chowdhury and S. De, "Queue-aware access prioritization for massive machine-type communication," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 15858–15873, Sep. 2022.
- [25] X. Pei, H. Qian, H. Wang, and K. Kang, "An improved random access scheme using directional beams for 5G massive machine-type communications," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 8913–8924, Jun. 2022.
- [26] B.-H. Lee, H.-S. Lee, S. Moon, and J.-W. Lee, "Enhanced random access for massive-machine-type communications," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 7046–7064, Apr. 2021.
- [27] D. Goswami, M. Dzaferagic, H. Siljak, S. S. Das, and N. Marchetti, "Analysis of temporal robustness in massive machine type communications," *IEEE Internet Things J.*, vol. 10, no. 8, pp. 6915–6927, Apr. 2023.
- [28] Y. Zhao, X. Chi, L. Qian, Y. Zhu, and F. Hou, "Resource allocation and slicing puncture in cellular networks with eMBB and URLLC terminals coexistence," *IEEE Internet Things J.*, vol. 9, no. 19, pp. 18431–18444, Oct. 2022.
- [29] P. Yang, X. Xi, T. Q. S. Quek, J. Chen, X. Cao, and D. Wu, "RAN slicing for massive IoT and bursty URLLC service multiplexing: Analysis and optimization," *IEEE Internet Things J.*, vol. 8, no. 18, pp. 14258–14275, Sep. 2021.
- [30] Y. Huang, S. Li, C. Li, Y. T. Hou, and W. Lou, "A deep-reinforcement-learning-based approach to dynamic eMBB/URLLC multiplexing in 5G NR," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6439–6456, Jul. 2020.
- [31] F. Librino and P. Santi, "Resource allocation and sharing in URLLC for IoT applications using shareability graphs," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10511–10526, Oct. 2020.
- [32] J. Yun, Y. Goh, W. Yoo, and J.-M. Chung, "5G multi-RAT URLLC and eMBB dynamic task offloading with MEC resource allocation using distributed deep reinforcement learning," *IEEE Internet Things J.*, vol. 9, no. 20, pp. 20733–20749, Oct. 2022.
- [33] T. N. Weerasinghe, I. A. M. Balapuwaduge, and F. Y. Li, "Priority-based initial access for URLLC traffic in massive IoT networks: Schemes and performance analysis," *Comput. Netw.*, vol. 178, Sep. 2020, Art. no. 107360.
- [34] Y. Ruan, G. Nie, W. Ni, H. Tian, and J. Ren, "Efficient traffic scheduling for coexistence of eMBB and URLLC in industrial IoT networks," in *Proc. IEEE WCNC*, 2022, pp. 1431–1436.
- [35] G. Interdonato, S. Buzzi, C. D'Andrea, L. Venturino, C. D'Elia, and P. Vendittelli, "On the coexistence of eMBB and URLLC in multi-cell massive MIMO," 2023, *arXiv:2301.03575*.
- [36] "5G technology in Europe: Ultra-reliable low latency communication." EU Parliament. Feb. 2023. [Online]. Available: <https://map.sciencemediahub.eu/5g#m=6/1388.88067/659.5,p=62>
- [37] "5G-PPP and 5G-IA: The European 5G annual journal 2020." Feb. 2023. [Online]. Available: <https://bscw.5g-ppp.eu/pub/bscw.cgi/d356008/Fu%20%20Annu%20Journ%202020.pdf>
- [38] F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji, and V. C. M. Leung, "Enabling massive IoT toward 6G: A comprehensive survey," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 11891–11915, Aug. 2021.
- [39] D. Amaxilatis et al., "Advancing experimentation-as-a-service through urban IoT experiments," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2563–2572, Apr. 2019.
- [40] T. B. da Silva, R. P. Dos Santos Chaib, C. S. Arismar, R. da Rosa Righi, and A. M. Alberti, "Toward future Internet of Things experimentation and evaluation," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 8469–8484, Jun. 2022.
- [41] "IoT categorization: Exploring the need for standardizing additional network slices." Alliance for Telecommunications Industry Solutions. Feb. 2023. [Online]. Available: https://access.atis.org/apps/group_public/download.php/51129/ATIS-I-0000075.pdf
- [42] X. Jiang et al., "Low-latency networking: Where latency lurks and how to tame it," *Proc. IEEE*, vol. 107, no. 2, pp. 280–306, Feb. 2019.
- [43] N. Bouillot and J. R. Cooperstock, "Challenges and performance of high-fidelity audio streaming for interactive performances," in *Proc. Int. Conf. New Interfaces Musical Exp.*, 2009, pp. 135–140.
- [44] C. Perkins, O. Hodson, and V. Hardman, "A survey of packet loss recovery techniques for streaming audio," *IEEE Netw.*, vol. 12, no. 5, pp. 40–48, Sep./Oct. 1998.
- [45] M. Fink and U. Zölzer, "Low-delay error concealment with low computational overhead for audio over IP applications," in *Proc. Int. Conf. Digit. Audio Effects*, 2014, pp. 309–316.
- [46] P. Verma, A. Mezzay, C. Chafe, and C. Rottondi, "A deep learning approach for low-latency packet loss concealment of audio signals in networked music performance applications," in *Proc. 27th Conf. Open Innov. Assoc. (FRUCT)*, 2020, pp. 268–275.
- [47] M. Sacchetto, Y. Huang, A. Bianco, and C. Rottondi, "Using autoregressive models for real-time packet loss concealment in networked music performance applications," in *Proc. Int. Conf. Audio Mostly*, 2022, pp. 203–210.
- [48] C. Chafe, J. Caceres, and M. Gurevich, "Effect of temporal separation on synchronization in rhythmic performance," *Perception*, vol. 39, no. 7, pp. 982–992, 2010.
- [49] S. Farner, A. Solvang, A. Sæbo, and U. Svensson, "Ensemble hand-clapping experiments under the influence of delay and various acoustic environments," *J. Audio Eng. Soc.*, vol. 57, no. 12, pp. 1028–1041, 2009.
- [50] P. Driessen, T. Darcie, and B. Pillay, "The effects of network delay on tempo in musical performance," *Comput. Music J.*, vol. 35, no. 1, pp. 76–89, 2011.
- [51] C. Bartlette, D. Headlam, M. Bocko, and G. Velikic, "Effect of network latency on interactive musical performance," *Music Percept.*, vol. 24, no. 1, pp. 49–62, 2006.
- [52] A. Sawchuk, E. Chew, R. Zimmermann, C. Papadopoulos, and C. Kyriakakis, "From remote media immersion to distributed immersive performance," in *Proc. ACM SIGMM Workshop Experiential Telepresence*, 2003, pp. 110–120.
- [53] C. Rottondi, M. Buccoli, M. Zanoni, D. Garao, G. Verticale, and A. Sarti, "Feature-based analysis of the effects of packet delay on networked musical interactions," *J. Audio Eng. Soc.*, vol. 63, no. 11, pp. 864–875, 2015.

- [54] R. Hupke, D. Jan, N. Werner, and J. Peissig, "Latency and quality-of-experience analysis of a networked music performance framework for realistic interaction," in in *Proc. Audio Eng. Soc. Conv.*, 2022, Art. no. 10546.
- [55] A. F. Khalifeh, A.-K. Al-Tamimi, and K. A. Darabkh, "Perceptual evaluation of audio quality under lossy networks," in *Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, 2017, pp. 939–943.
- [56] M. Fink, M. Holters, and U. Zölzer, "Comparison of various predictors for audio extrapolation," in *Proc. Int. Conf. Digit. Audio Effects*, 2013, pp. 1–7.
- [57] M. Zorzi and R. Rao, "On the statistics of block errors in bursty channels," *IEEE Trans. Commun.*, vol. 45, no. 6, pp. 660–667, Jun. 1997.
- [58] L. Bonati, M. Polese, S. D'Oro, S. Basagni, and T. Melodia, "Open, programmable, and virtualized 5G networks: State-of-the-art and the road ahead," *Comput. Netw.*, vol. 182, Dec. 2020, Art. no. 107516.



Luca Turchet (Senior Member, IEEE) received the first master's degree (summa cum laude) in computer science from the University of Verona, Verona, Italy, in 2006, the second master's degree in classical guitar and third master's degree in composition from the Music Conservatory of Verona, Verona, in 2007 and 2009, respectively, and the fourth master's degree in electronic music from the Royal College of Music of Stockholm, Stockholm, Sweden, in 2015, and the Ph.D. degree in media technology from Aalborg University Copenhagen, Copenhagen, Denmark, in 2013.

He is an Associate Professor with the Department of Information Engineering and Computer Science, University of Trento, Trento, Italy. His scientific, artistic, and entrepreneurial research has been supported by numerous grants from different funding agencies, including the European Commission, the European Institute of Innovation and Technology, the European Space Agency, the Italian Ministry of Foreign Affairs, and the Danish Research Council. He is Co-Founder of the music-tech company Elk.

Dr. Turchet is the Chair of the IEEE EMERGING TECHNOLOGY INITIATIVE ON THE INTERNET OF SOUNDS and the Founding President of the Internet of Sounds Research Network. He serves as an Associate Editor for *IEEE ACCESS* and the *Journal of the Audio Engineering Society*, and has been Guest Editor of the *IEEE Communications Magazine*, the *Personal and Ubiquitous Computing Journal*, the *Journal of the Audio Engineering Society*, *Frontiers in VR*, and *Digital Creativity*.



Paolo Casari (Senior Member, IEEE) received the Ph.D. degree in information engineering from the University of Padua, Padua, Italy, in 2008.

He is an Associate Professor with the Department of Information Engineering and Computer Science, University of Trento, Trento, Italy. He was on leave with the Massachusetts Institute of Technology working on underwater communications and networks, in 2007. He is currently the Principal Investigator of the NATO SPS project SAFE-UComm. Previously, he coordinated the NATO SPS project ThreatDetect, and was the Scientific Manager of the EU H2020 RECAP and SYMBIOSIS projects. From 2015 to 2019, he was with the IMDEA Networks Institute, Madrid, Spain, where he led the Ubiquitous Wireless Networks Group. His research interests include many aspects of networked communications, such as channel modeling, network protocol design, localization, simulation, and experimental evaluations.

Prof. Casari received two best paper awards. He regularly collaborates with the organizing committee of several international conferences. He serves on the editorial board of the *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS*, served as an Associate Editor for *IEEE TRANSACTIONS ON MOBILE COMPUTING* from 2018 to 2023, is currently the Co-Guest Editor of two special issues for Elsevier's *Computer Communications* and *Computer Networks journals*, and has been the Co-Guest Editor of a special issue of *IEEE ACCESS* on underwater acoustic communications and networking.