

On the Impact of 5G Slicing on an Internet of Musical Things System

Luca Turchet¹, Senior Member, IEEE, and Paolo Casari², Senior Member, IEEE

Abstract—Considering the specific Quality-of-Service requirements of Internet of Musical Things (IoMusT) deployments, 5G networks are deemed as a key enabling technology to support IoMusT applications as needed, e.g., in networked music performance (NMP) settings. Slicing, in particular, is a promising feature of 5G networks, as it could help dedicate specific resources to NMP traffic, so as to schedule NMP communications with higher priority and thus guarantee swift delivery delays and higher reliability. Moreover, handling NMP traffic via a slice residing on a geographically closer multiaccess edge computing (MEC) server would help significantly reduce WAN transit times for NMP traffic. In this article, we conduct the first tests with slicing in 5G networks supporting IoMusT applications (to the best of our knowledge). We conduct our tests by decoupling WAN transit times from the radio access performance (evaluated in terms of latency, error rate, number of lost packets, and maximum length of a packet loss burst). Our results show that slicing, even when introducing slight extra delays due to UPF implementation or hardware specifications, is a potentially excellent choice to save on WAN transit times for NMP traffic while delivering the same reliability as the operator’s core network equipment.

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

I. INTRODUCTION

NETWORKED music performance (NMP) systems are hardware and software technologies aiming at enabling geographically displaced musicians to play together over wireless and/or wired networks [1], [2]. During the COVID-19 pandemic, NMP systems received boosted attention and demand from professional and amateur musicians, who were prevented from conducting a variety of musical activities in presence, from classes to rehearsals and performances [3].

To date, NMP remains a challenging application over wireless networks. The main reason is that NMP has strict Quality-of-Service (QoS) requirements that must be guaranteed in order to ensure realistic collaborative interactions between musicians over distant locations. Not only must end-to-end audio

communications occur through the network with low latency (namely, below 30 ms) but audio transfers must be very reliable.

To ensure high audio quality, packet loss rates must be sufficiently low to prevent musicians from perceiving quality drops in the signal heard. Typically, the minimization of latency is prioritized over reducing the bandwidth consumption in NMP systems. As a consequence, audio compression algorithms are usually avoided, to prevent them from introducing extra latency. For the same reason, loss-prone but fast Internet transport protocols, such as UDP, are preferred.

Relevant examples of NMP systems, either at the commercial or at the experimental level, are Elk LIVE [4], JackTrip [5], LOLA [6], and fast-music [7]. Initially, the majority of these technologies consisted of software applications executable on general-purpose machines. Today, the most advanced NMP systems leverage dedicated hardware platforms, specifically designed to minimize analog-to-digital and digital-to-analog conversions of the audio signals produced and received by the connected musicians, as well as audio processing and buffering delays.

All devices that support NMPs are a fundamental component of the emerging paradigm of the Internet of Musical Things (IoMusT), an extension of the Internet of Things (IoT) concept to the musical domain. The IoMusT vision refers to the network of “Musical Things.” Each Musical Thing is a computing device embedded in a physical object dedicated to the production and/or reception of musical content [8]. According to the IoMusT vision, future musical interfaces will embed intelligence and communication capabilities. The enabling technologies for the IoMusT are thus high-performance embedded systems for audio sampling and processing, along with reliable low-latency connectivity options.

From the point of view of communications, one of the expected enablers for the IoMusT is the widespread availability of fifth-generation cellular networks (5G), the latest generation of mobile networks standardized by the 3rd Generation Partnership Project (3GPP) [9], [10]. 5G networks have been conceived to overcome a number of shortcomings of previous-generation 4G networks [11]. This includes the support for higher bandwidth data communications, faster transmission scheduling via higher numerologies,¹ a more

Manuscript received 31 May 2024; accepted 30 June 2024. Date of publication 2 July 2024; date of current version 25 September 2024. This work was supported in part by the European Space Agency under Grant 4000132621/20/NL/AF, and in part by the MUR PNRR PRIN 2022 Grant funded by Next Generation EU under Grant 2022CZWWKP. (Corresponding author: Luca Turchet.)

Luca Turchet is with the Department of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy (e-mail: luca.turchet@unitn.it).

Paolo Casari is with the Department of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy, and also with CNIT, 43124 Parma, Italy (e-mail: paolo.casari@unitn.it).

Digital Object Identifier 10.1109/JIOT.2024.3422287

¹In the 5G system, the term “numerology” refers to the configuration of subcarrier spacing (SCS) and, therefore, to the duration of a transmission slot. For example, numerology $\mu = 0$ refers to a 15-kHz SCS and a slot length of 1 ms; for higher values of μ , the SCS increases by 2^μ and the slot length decreases by the same factor. For example, numerology $\mu = 2$ implies a 60-kHz SCS and a slot length of 250 μ s.

flexible core network (CN), including virtualized network functions and edge computing, as well as lower radio access network (RAN) latency [12]. These features provide the ideal networking substrate for low-latency and highly reliable applications such as NMPs [13].

5G is envisioned to support a wide range of services, each having a diverse set of QoS requirements. Satisfying such requirements entails a flexible and highly programmable networking infrastructure, and clear interfaces to allocate and relieve computing and communication resources. These objectives are achieved through slicing, a technology that allows multiple, virtualized, and independent networks to be instantiated and co-exist on top of a common physical infrastructure [14], [15].

Network slicing represents a key technology to enable multiple services with diverse requirements to share infrastructure resources. However, to date, the use of slicing in real IoMusT deployments has been largely overlooked, especially in combination with edge computing. It remains unclear whether a slicing-enabled, 3GPP-compliant 5G architecture offers feasible key performances indicators (such as latency and reliability) for the IoMusT, and if yes, to what extent. Furthermore, it is not clear how many concurrent NMPs can be supported over the same base station, for instance, when the members of two ensembles play together using two distinct but simultaneous NMP sessions. The impact of networking and processing delays in different components of the 5G system [e.g., the CN and multiaccess edge computing (MEC)] has been scarcely investigated in the NMP context.

To bridge these gaps, in this article, we investigate and compare three 5G architectures where up to eight emulated musicians are connected to the same 5G base station. The first architecture assumes that the NMP devices, the RAN, and the CN are co-located. The second that they are co-located with a MEC server, and the third that the CN and MEC server are remote, but at different location, with the MEC being closer to the NMP devices than the CN itself. The latter represents a typical deployment case for 5G networks, where MEC servers are deployed closer to users of specific network and computing functions in order to yield lower access delay to such functions. All architectures were investigated with and without slicing as well as with and without concurrent traffic. In summary, our main contributions in this work are as follows.

- 1) We measure the performance of a 5G network in terms of key performance indicators (KPIs) related to latency and reliability, when supporting one or two NMP sessions.
- 2) We evaluate how 5G performance varies as a function of the number of IoMusT nodes, with and without slicing, and in the presence of background traffic.
- 3) We specifically evaluate the case of two NMP sessions under a single base station, with and without slicing.
- 4) We discuss whether measured 5G performance is sufficient to cover the requirements of dense IoMusT deployments.

While slicing has been already investigated in a number of scenarios (e.g., unmanned aerial vehicle [16] or healthcare [17]), to the best of the authors' knowledge, this is the first

study that evaluates slicing in an IoMusT deployment. Existing studies on 5G deployments to support NMPs leveraging private or public networks [18], [19], [20] have consistently shown that not every single feature specified in 5G standards is available in state-of-the-art 5G-standalone (5G-SA) networks. This suggests that only features with the most promising market viability will be implemented as the market requirements and implementation pressure increases and motivates further investments for network updates. As a consequence, the potential of 5G cellular systems remains largely unexpressed in musical settings, and a systematic evaluation of the 5G network performance in realistic IoMusT scenarios remains an open research question. Although several studies in the literature investigate networking systems at large in the context of the IoMusT [21], we identify a gap on the effects of network slicing in musical use cases. The study in this article sheds some initial light on these aspects, hoping to trigger further investigations on optimized architectures to support NMP in 5G networks.

II. RELATED WORK

A. 5G Networks

5G networks may be public or private. Public 5G networks are intended for use by general citizens (e.g., with millions of subscribers on a given nationwide network). In contrast, private 5G networks are installed at the premises of a single organization, typically within a specific area. An end-to-end 5G network (private or public) is typically composed of three elements [22].

- 1) *RAN*: The network infrastructure that includes radio base stations and bridges the connection between mobile radio network devices and the CN.
- 2) *CN*: The central part of a network that provides services to users connected via the access network and conveys IP packets to external networks such as the Internet.
- 3) *User Equipment (UE)*: Any device directly used by an end user to communicate; this includes mobile smartphone appliances, communication systems embedded in low-power edge devices, massive IoT communication devices, as well as NMP equipment.

Furthermore, the 5G standard promoted by 3GPP has provisions for the following features.

- 1) *Slicing*: The arrangement of a network into a set of logically separated, self-contained, and independent networks, each of them dedicated to a different purpose or services with different QoS requirements.
- 2) *MEC*: An architecture that provides cloud computing capabilities at the edge of the network (i.e., close to the base station), making it possible to reduce latency, ensure highly efficient network operations and service delivery, and improve the customer experience; notably, by deploying services and caching content at the network edge, the CN is alleviated of further congestion and it is possible to serve local purposes in an efficient way.
- 3) *User Plane Function (UPF)*: A virtual network function responsible for user packet routing and forwarding, packet inspections, and QoS handling; the UPF may be located on the MEC or on the CN.

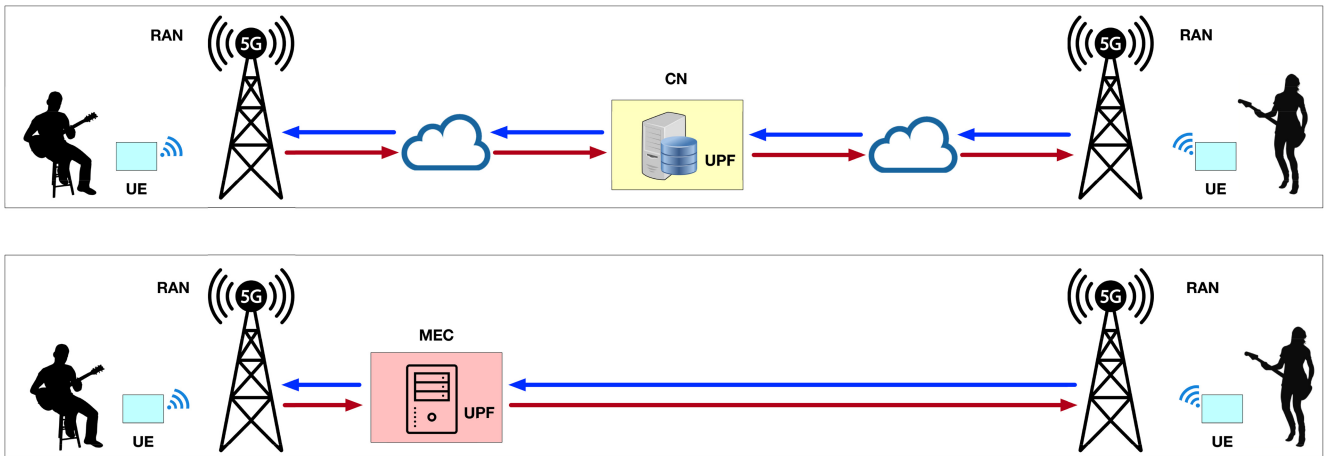


Fig. 1. Schematic representation of the components of a 5G architecture for NMPs, with MEC (bottom) and without MEC (top).

Fig. 1 illustrates two 5G architectures, with and without MEC, that can support an NMP including two musicians.

B. Key Performance Indicators for NMP Systems

The aim of NMP systems is to provide connected musicians with the same conditions as acoustic/instrumental on-site performances [2]. Such conditions imply very strict QoS requirements, which translate in the following KPIs: 1) end-to-end latency less than 30 ms; 2) low and constant jitter (i.e., the variation of latency); and 3) high audio quality (i.e., minimal packet losses in order to avoid noticeable audio dropouts) [23], [24]. Achieving such KPIs ensures that the musicians maintain a stable tempo, play synchronously, and enjoy a satisfactory auditory perception, leading to a high-quality interaction experience [1, Ch. 3].

The 30-ms latency threshold has been determined experimentally (for a review see [2]), and corresponds to the time it takes for a sound wave to propagate for about 10 m in air. Such a distance is typically assumed to be the maximum displacement tolerable by members of a musical ensemble, while ensuring a stable interplay in the absence of external synchronization cues.

Concerning jitter, NMPs require that the latency should not fluctuate significantly; otherwise, the synchronization among musicians is negatively affected, and artifacts, such as audio glitches, are introduced in the audio stream [2]. To compensate for the varying transmission latency of individual packets, NMP systems usually comprise a jitter buffer of configurable size at each receiver: the larger the buffer, the higher the introduced latency. Once the jitter buffer is set in place, latency becomes constant. Choosing the size of the jitter buffer is not a trivial task, which demands careful consideration, and entails a tradeoff between the gain in audio quality and the longer latency.

Regarding reliability (i.e., the capability of guaranteeing successful message transmissions within a defined latency bound), at the time of writing, there is still no consensus in the scientific community on a minimum threshold value. Dürre et al. [18] proposed that potentially realistic packet error ratios for NMPs range from 10^{-6} up to 10^{-4} . Nevertheless,

a measure of reliability should not be based solely on the probability of packet losses, as it also depends on the time span of packetized audio samples (which depends on the number of audio samples and the sampling frequency). This consideration represents a crucial aspect, especially when applying packet loss concealment methods that aim to reconstruct the missing signal without introducing additional delays [25], [26], [27], [28]. In general, the relationship between packet loss, its distribution over time, and perceived audio quality has not been fully determined to date. Only a few studies have focused on such a complex matter [29], [30]. However, there is consensus that consecutive packet losses cause the most harmful impact on the perceived 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.

C. NMPs Over 5G

To date, only a few 5G architectures have been investigated to support NMPs. Centenaro et al. [10] outlined a communication architecture for an IoMusT system based on the 5G framework, introducing two prominent use cases (NMP and fast access to a server) and introducing a system model. Some initial experiments involving early 5G hardware have typically targeted feasibility rather than an in-depth statistical analysis of the 5G network's actual latency and reliability performance for NMP applications (see [31], [32]). In general, most existing experiments have focused on special-purpose 5G architectures involving only a few endpoints (two or three), such as the one reported by Carôt et al. in [33].

More recently, Dürre et al. [18] have performed a long-term collection of latency and packet error measures for a simulated NMP over a 5G public infrastructure connecting two endpoints. Their measurements confirm that maintaining adequate QoS is a required condition to enable NMP over cellular 5G in a consistent and flexible way. Along the same line, our previous study [19] deployed two 5G architectures involving two musical endpoints: 1) a private 5G standalone network with edge computing infrastructure and 2) a pre-commercial public 5G nonstandalone (NSA) network. Latency

and reliability measurements showed that public cellular 5G standalone architectures with edge computing support are required for realistic real-time musical interactions. In 2022, the 5G festival [34] provided a facility to test musical and video interactions on a metropolitan scale. The setup included private 5G networks offering both SA and NSA service, public 5G-NSA service, and layer-2 metropolitan links with guaranteed throughput and latency performance. The network was organized into a 1-hub, 2-spoke architecture and included a virtual MEC server hosting the 5G common service platform, as well as remote caching. This remarkable setup shows a network latency of 30–60 ms, above perceptual thresholds for remote performers. In a later work [20], we described a private 5G IoMusT deployment with four endpoints and analyzed its performance when supporting NMPs, with a focus on conditions including or not including background wireless traffic in support of other applications. Results showed that latency increased with the number of nodes and with the presence of background traffic, whereas the reliability did not vary with the complexity of the conditions. To the best of authors' knowledge, no study has investigated the use of 5G slicing in an NMP setting thus far.

III. MATERIALS AND METHODS

In this article, we consider both co-located and remote deployments. In part, the methodology was inspired by that reported in our previous study [20], where we measured the latency and reliability performances of a co-located 5G architecture supporting NMPs. For the co-located architectures, we focused on the assessment of the sole wireless link. In absence of a real deployment using multiple 5G base stations over a sufficiently broad geographical area, for the remote architectures, we combined the measurements of the NMP system conducted on a WAN with those related to the co-located architectures. Decoupling the measurements on the RAN from those on the WAN allowed us to transfer the measurements of the co-located architectures to a realistic, distributed NMP architecture.

In particular, for the measurements on the WAN, we leveraged the dedicated communication infrastructure interconnecting Italian Universities (GARR: <https://www.garr.it/>). This selection was made in order to realistically simulate the behavior of the wired network of a 5G mobile operator, which is expected to have better performances compared to the best effort infrastructure of the Internet.

A. Apparatus

Four 5G architectures (illustrated in Fig. 2 for two endpoints and described below) were investigated using 5G radio and CN equipment. The test environment was arranged in an indoor space of the ZTE Italia Innovation & Research Center (ZIRC) located in the city of L'Aquila, Italy. The equipment considered included a base station installed on the ceiling of the experiment room, about 3 m away from ten UEs placed on a table (see Fig. 3).

- 1) *Co-located CN*: This architecture consists of a set of 5G-enabled UEs that interact with the same base station.

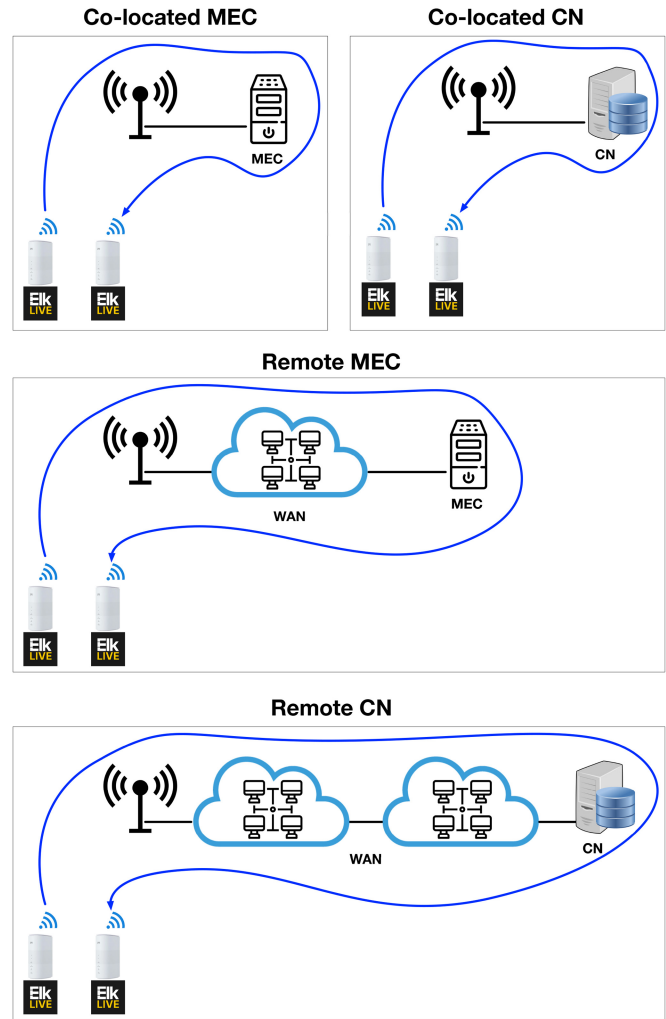


Fig. 2. Schematic representation of the four architectures investigated and the flow of packets between a transmitter and a receiver NMP device.

CN servers implemented TURN functionalities to enable direct communications between the UEs, and hosted all core functions including the UPF. Therefore, for this architecture, the slicing was not enabled.

- 2) *Co-located MEC*: This architecture comprises the same UEs and base station of the co-located CN architecture, but we enable slicing and let the MEC server host the UPF.
- 3) *Remote CN*: This architecture compounds the co-located CN architecture with a dedicated WAN, which simulates the connection of the base station to a CN server.
- 4) *Remote MEC*: This architecture combines a WAN with dedicated performance with the co-located MEC architecture. Specifically, we investigated the case of WAN connecting the base station to the MEC server.

The above arrangement makes it possible to decouple the performance of the RAN and WAN, and thus to assess the impact of these two network sections separately. For example, having to route packetized audio traffic through a WAN will incur longer transit delays, but will give access to typically high-end CN servers to host network and computing functions, providing state-of-the-art processing delay. Conversely,

TABLE I
BANDWIDTH CONSUMPTION FOR BOTH UPLINK AND DOWNLINK (IN Mb/s) FOR THE NUMBER OF BOXES IN SESSION 1 (S1) AND SESSION 2 (S2)

Number of boxes	Number of audio streams	Throughput per box [Mbit/s]	Total throughput [Mbit/s]
2 (in S1)	2	≈ 2.5	≈ 5
3 (in S1)	6	≈ 7.5	≈ 23
4 (in S1)	12	≈ 11.5	≈ 46
5 (in S1)	20	≈ 15.5	≈ 78
6 (4 in S1 + 2 in S2)	14	≈ 11.5 for S1, ≈ 2.5 for S2	≈ 52
7 (4 in S1 + 3 in S2)	18	≈ 11.5 for S1, ≈ 7.5 for S2	≈ 70
8 (4 in S1 + 4 in S2)	24	≈ 11.5 for S1, ≈ 11.5 for S2	≈ 93



Fig. 3. Picture of the simulation environment of the 5G architecture, showing the base station, the nine CPEs, the eight NMP devices, the eight headphones, the soundcard, the 13 laptops, and the smartphone.

keeping the traffic local to the closest MEC server reduces the WAN transit delay, even though MEC servers may not deliver the same computational power as CN servers. We comment on these aspects further below.

All the hardware employed in our experiments is commercially available and has not been modified in any ways for the purposes of this experiment.

1) *User Equipment*: Our experimental setup involved ten UEs. Eight were dedicated to the NMP system, acting 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. Each of the eight UEs used for audio transfer consisted of a ZTE MC801A1 customer premise equipment (CPE, i.e., a 5G/WiFi/Ethernet router) connected via Ethernet to an audio/network interface device providing a peer-to-peer NMP system. For this purpose, we used the Elk LIVE NMP system, where each node consists of a hardware device running the Elk Audio OS, a low-latency audio operating system optimized for embedded systems [4].

The musical audio streams were emulated via a dedicated software coded in the Pure Data real-time audio programming language. This enabled us to automate the experiment sessions compared to the involvement of real musicians. The eight audio signals corresponded to recordings of eight musicians playing together but recorded separately (namely, two electric

basses, two drums, two keyboards, and two electric guitar players). Such recordings were played back synchronously via a dedicated laptop, and routed to a soundcard with eight analog output channels (RME Fireface UFX II). Using audio cables, we connected each such output channel to the input of one NMP device. Each device independently mixed the local input sound stream with the streams received from the other devices (one to seven depending on the experimental conditions). The mixed signal could then be heard from headphones plugged in each device.

In more detail, each device transforms the incoming analog audio signals into IP packets when acting as a sender, and vice versa when acting as a receiver. The device enables deterministic processing times for high-precision packet pacing and timestamping, logs received IP packet latency and jitter values, and tracks any losses. Each device in the NMP system was set to work with a sampling frequency of 48 kHz. All devices introduced a deterministic delay which amounts to 14.32 ms. This includes 0.5 ms for analog-to-digital and 0.5 ms for digital-to-analog conversions; ≈ 1.33 ms for the audio buffer utilized by the audio host for input and output (i.e., 64 samples at a sampling rate of 48 kHz); ≈ 10.66 ms for the set jitter buffer size. This left a latency budget of up to 15.68 ms to network transit in order to avoid exceeding the total latency tolerable by musicians (i.e., 30 ms).

The NMP system transmits 2-channel stereo signals. For each channel, the device produces a protocol data unit (PDU) comprising 64 audio samples at 16 bits/sample. The UDP protocol is employed at the transport layer without including any forward error correction (FEC) or automatic repeat request (ARQ) scheme to protect the stream. As two audio channels are involved, and considering timestamping data and the UDP header length, the total PDU size is 272 bytes, and the packet transmission rate is one packet every $64/(48 \cdot 10^3) \approx 1.33$ ms. Table I summarizes the measured bandwidth consumption for both the uplink and the downlink, as a function of the number of devices involved. The NMP system allows for running multiple sessions where up to five musicians play together. In the experimental conditions, one and two sessions were involved. In each session, the audio streams transmitted by a box are $n-1$, where n is the number of boxes used, for a total of $n(n-1)$ streams.

The background traffic was created using two UEs, one of which consisted of a laptop and the same CPE as the other UEs. The second UE consisted of a 5G-enabled ZTE Axon 10S Pro 5G smartphone. The reference architecture is

composed of two client–server communications. The laptop connected to the CPE acted as a receiver for the downlink traffic generated by a server placed inside the CN or MEC, depending on the chosen architecture. The smartphone acted as a sender for the uplink traffic toward a server, also hosted in the CN or MEC. The UDP traffic was implemented by a server–client architecture based on the `iperf3` software for network traffic generation and performance tests. It is worth noting that we used only UDP, which provides a more aggressive form of contending traffic with respect to, e.g., bandwidth-throttled TCP streams [20].

2) *Radio Access Network*: Radio access was provided by a 5G SA base station, comprising an antenna-based device and a baseband unit (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 time-division duplexing (TDD). The QCell was connected to a ZTE V9200 BBU via a 1-m optical cable. The maximum available bandwidth was measured via ZTE’s proprietary data rate metering software, yielding 1000 Mb/s in downlink and 270 Mb/s in uplink.

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. The CN included eight servers, of which 3 were devoted to computing and network function hosting, including the UPF. A standard proportional fair scheduler was used, without any nonstandard priority settings.

4) *MEC*: The MEC server (ZTE ZXRRAN U9003) was located in the same building as the BBU. The MEC was connected to the BBU through a 2-m optical cable. The role of the MEC was to act as a TURN server, and relay the audio traffic among the NMP devices. In particular, the TURN server forwards the data flows from one box to the other $n - 1$ boxes in the session. The MEC encompassed five servers, of which only one dedicated to the UPF deployment. This configuration was selected to replicate real-world deployments.

5) *Slicing*: Activating slicing in our setup has a twofold effect: on the one hand, the MEC server hosts both the TURN server and the UPF, network function, making it possible to avoid having to route all traffic through the CN; on the other hand, it reserves resources for the audio streams of the UEs dedicated to the NMP. Specifically, we configured two end-to-end slices: one was optimized for the NMP, and configured to guarantee a higher priority in resource scheduling via the 5G QoS Identifier (5QI). The second slice was configured for non-NMP communications, and thus collected all background UDP traffic. The slicing setup guarantees resources to the NMP devices and avoids that external traffic intervenes by subtracting resources.

In the absence of slicing, both the NMP UEs and the background traffic generators resided on the same network slice, and shared resources equally at the same priority level.

6) *WAN*: As discussed above, the need to route traffic through the mobile operator’s WAN connection until the operator’s CN increases the overall end-to-end packet delivery delays during the NMP sessions. In order to reproduce these delays realistically while honoring the generally high degree of

reliability of mobile operator WANs, we consider the GARR network as a use case [35]. The GARR consortium manages high-performance, state-of-the-art connectivity among all Italian universities, and is part of the broader GEANT pan-European data network for research and education. Having endpoints in all major Italian cities and relying on dedicated fiber-optic links across such cities, the GARR network is an ideal representative of commercial-grade WAN performance for our evaluation.

For our setup, we make the following assumption. The musicians participating to the NMP session are located in Trento and Padua, in north-eastern Italy. The CN servers are located in Milan, Italy. The MEC server is assumed to reside in Padua as well, as a service point for north-eastern Italy as a whole.

B. Measurement Procedure for the Co-Located Architectures

The experiments carried out were organized into separate groups with different experimental conditions. Our aim with these experiments was to evaluate the correct functioning of the application and the efficiency of the 5G network both under nonsaturated and saturated conditions, with and without the use of network slicing.

The experimental setting involved one or two NMP sessions, obtained by varying the number of boxes from 2 to 8 (see Table I). This led to a total of 28 settings (7 box configurations \times 2 slicing configurations \times 2 traffic configurations). Three recordings were performed for each configuration. In each recording, the boxes continuously transmitted audio for 5 min and 30 s from one endpoint to the other(s), and vice versa. The logging system of each box measured the performance of the IP connection by observing packets in windows of predefined length and focusing on the following four metrics.

- 1) *Latency*: One-way latency in milliseconds, calculated as the round-trip time (RTT) between two nodes divided by two (under the assumption that the time of the outbound and inbound communication was the same); this measurement included the actual delay introduced by the network as well as the contribution due to the jitter buffer (i.e., ≈ 10.66 ms).
- 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) *Maximum Number of Consecutive Missed Packets*: The maximum number of consecutively lost packets within the analysis window.

Following the approach reported in [20], these metrics were computed on analysis windows of ≈ 2.33 s containing 1750 packets, each carrying 64 audio samples. The first 30 s of recording were discarded to remove any handshaking and service setup delays. This led to an analysis of about 225 000 packets for each box (i.e., 5 min) in each of the three recordings, leading to a total of about 676 000 packets for each box in each experimental setting. For each of the four performance metrics, the mean, standard deviation, minimum,

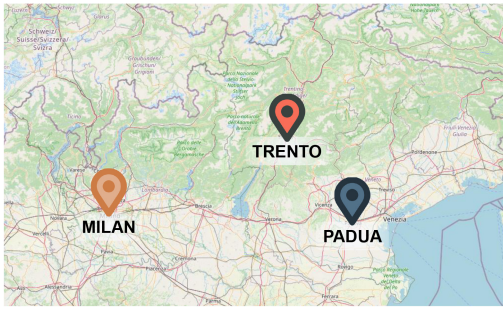


Fig. 4. Map of north-eastern Italy, showing the three cities of Trento, Padua, and Milan. (Map courtesy of OpenStreetMap.)

and maximum were computed by merging the log data from at each box in each experimental setting.

C. Accounting for WAN Transit Delays in Remote Architectures

To account for WAN transit delays across the cities of Trento, Milan, and Padua, we performed end-to-end RTT across endpoints of the GARR network in the three cities around 12 noon on 13 February 2024, a time of day that exhibits peak traffic. With reference to the map in Fig. 4, and using road distances as a rough proxy for network cabling lengths, we have the following data.

- 1) Trento \longleftrightarrow Padua: 180 km (RTT: 4.55 ms).
- 2) Trento \longleftrightarrow Milan: 220 km (RTT: 7.57 ms).
- 3) Padua \longleftrightarrow Milan: 240 km (RTT: 6.30 ms).

Moreover, as the GARR network is a dedicated infrastructure, we assume that its contribution to packet losses is negligible with respect to the 5G RAN.

In order to assess the performance of remote architectures, we then compound the co-located architecture measurements with WAN network performance. Starting with the total end-to-end latency, for the remote-CN architecture, we sum the radio access latency measurements of the co-located CN architecture with the time it takes for packets to transit from Trento to Padua via Milan (on average 13.87 ms). Conversely, for the remote-MEC architecture, we consider the flow of traffic to not involve the CN in Milan, but rather take place through the UPF hosted in the Padua-located MEC server. Therefore, we consider only the Trento–Padua delay (on average 4.55 ms). The loss performance of the network (packet loss ratio, number of missed packets, and maximum consecutive number of missed packets) is assumed to depend only on the RAN both in the remote-CN and in the remote-MEC case.

IV. RESULTS

A. RAN Performance Assessment

The characteristics of the radio link connecting the musical devices to the 5G base station have been first assessed using ZTE’s proprietary management system. The following KPIs of interest were considered:

- 1) *Modulation and Coding Scheme*: between 24 and 27 [36, Table 5.1.3.2-2];

- 2) *Block Error Rate*: between 7% and 8%, in accordance with the 3GPP standard mandating values less than 10%;
- 3) *Resource Blocks*: equal to 270, the maximum possible value for a bandwidth of 100 MHz;
- 4) *Noise and Interference Power*: measured equal to $-110/105$ dBm, with an interference level on the order of 5 dBm.

B. Latency and Reliability Metrics

Fig. 5 shows the mean and standard error for the co-located architectures, by considering each combination of box, slicing, and traffic setting.

We searched for possible correlations between the four metrics and the number of audio streams involved (see Table I). For this purpose, we utilized Pearson’s correlation test. Regarding latency, for all slicing and traffic conditions, we identified significant and strong correlations (all at $p < 0.001$ and with $r > 0.8$). Concerning the packet error ratio and missed packets metrics, a significant and medium-strength correlation with the number of audio streams was found only in the condition without traffic, for both the slicing and no-slicing conditions (in each case, the correlation was observed at $p < 0.05$ and with $r < 0.6$). Regarding the maximum number of consecutive missed packets, no significant correlations were found with the number of audio streams in any of the conditions.

Fig. 6 shows the mean and standard error for the co-located architectures, by considering the slicing and traffic conditions for all box settings together. This allowed us to analyze and compare the general behavior of the two architectures.

An analysis of variance (ANOVA) was performed on different linear mixed effect models, one for each of the four metrics. Specifically, each model had the metric and traffic or slicing condition as fixed factors, and the NMP device as a random factor.

For the condition with traffic, a significant main effect was found for factor architecture ($F(1, 37007) = 3524.8, p < 0.001$). Moreover, for the condition with slicing, a significant main effect was found for factor traffic ($F(1, 33817) = 1961, p < 0.001$), i.e., competing UDP traffic caused in the co-located MEC architecture a higher latency compared to when the traffic was absent. Analogously, for the condition without slicing, a significant main effect was found for factor traffic ($F(1, 38992) = 8627.7, p < 0.001$), i.e., the UDP traffic caused in the co-located CN architecture a higher latency compared to when the traffic was absent.

Conversely, no significant main effect was found for any of the three reliability metrics. This indicates that the packet error ratio, the missed packets, and the maximum number of consecutive missed packets were similar for the co-located MEC architecture and the co-located CN architecture. Moreover, the traffic did not have a significant influence on reliability.

Regarding the analysis on latency, for the condition without competing traffic, a significant main effect was found for factor slicing ($F(1, 35799) = 17019, p < 0.001$), i.e., the co-located MEC architecture lead to higher latencies than the co-located CN architecture. We remark that this is in accordance to a

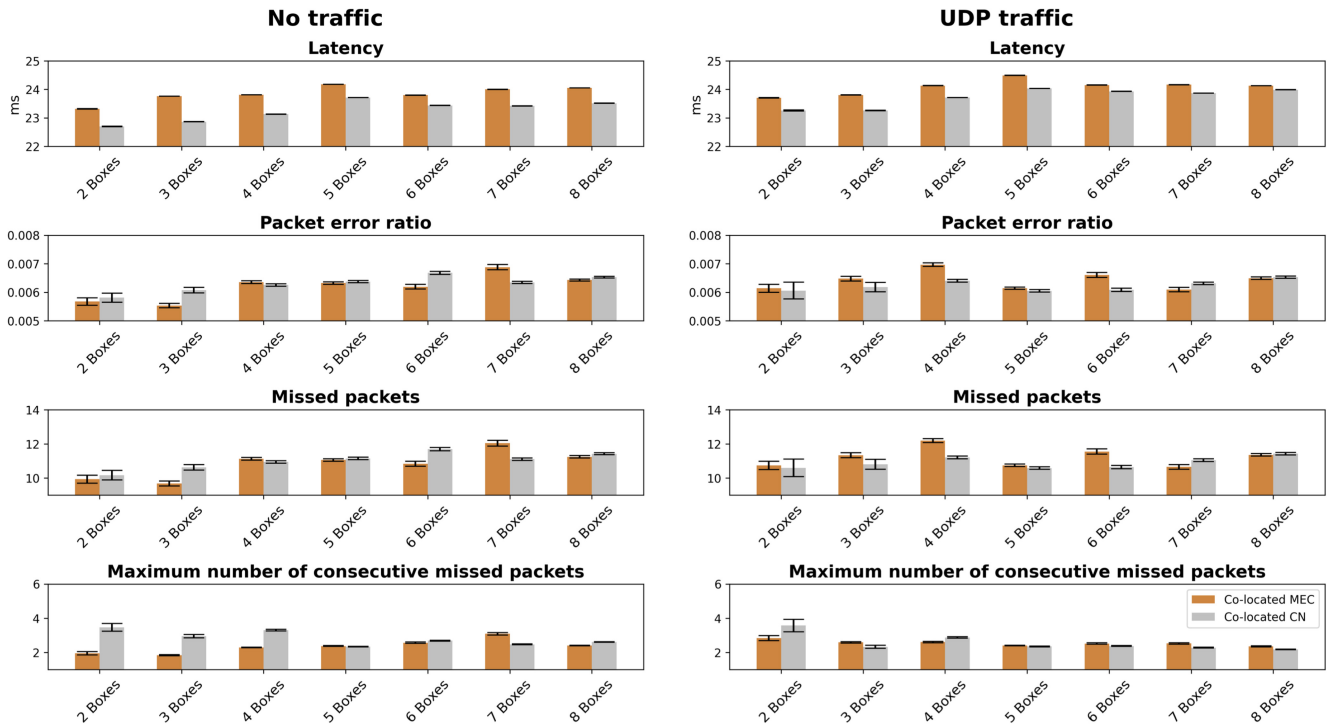


Fig. 5. Mean and standard error for the co-located architectures (by each box number, slicing, and traffic condition).

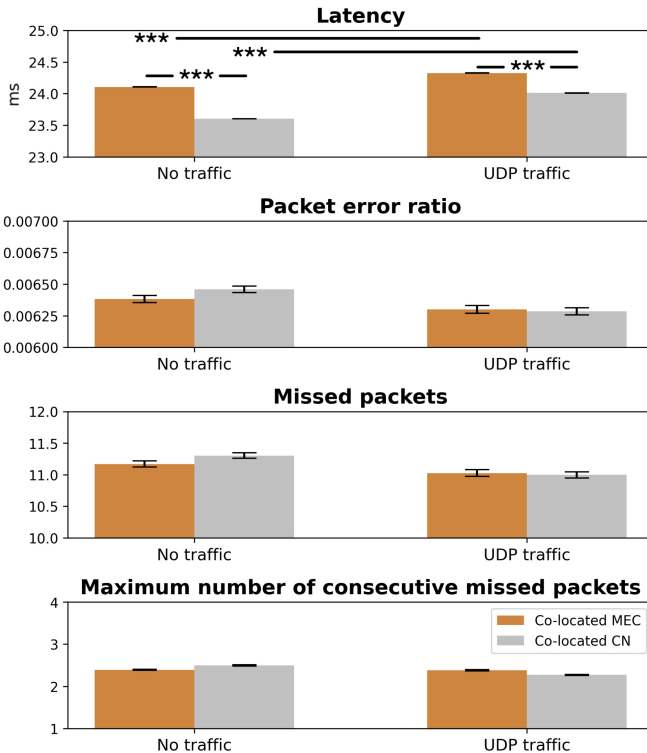
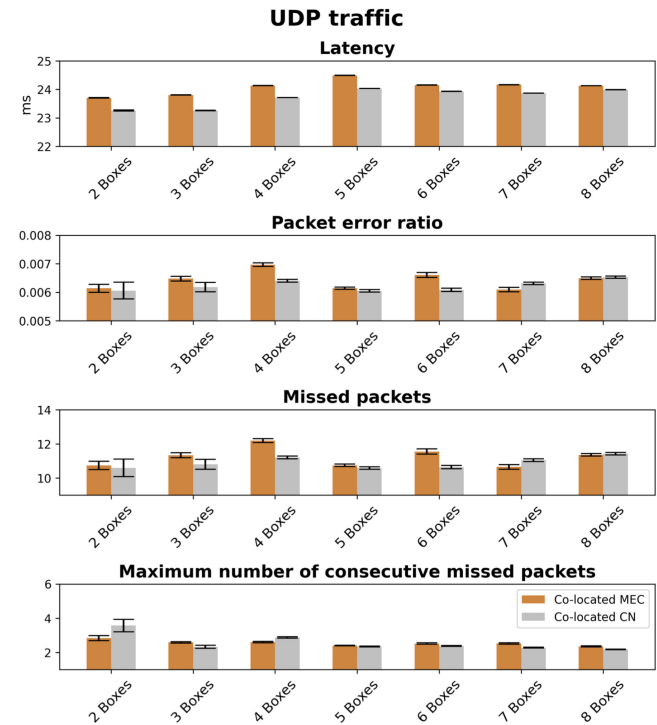


Fig. 6. Mean and standard error for the co-located architectures (by the slicing and traffic conditions for all box conditions together).

typical CN versus MEC-hosted network function setup, as MEC servers can be reasonably expected to provide less computational resources than CN servers. Therefore, execution



times on MEC servers are slightly larger and cause a globally larger latency.

The situation becomes starkly different when we consider WAN transit times for the remote CN and the remote MEC architectures. According to the data in Section III-C, end-to-end delays for the remote CN case amount on average to about 13.87 ms, whereas for the remote MEC case they amount to about 4.55 ms. The difference among the two figures is 9.32 ms. While this figure depends on the relative distance between the NMP devices and remote CN/MEC servers, realistic mobile network deployments are likely to offer shorter transit delays to territorial MEC servers rather than to the central CN of the mobile network operator. Whenever such transit times are higher than the extra delay caused by MEC hosting and processing, it becomes much more convenient to create a dedicated slice for the NMP devices and refer their traffic to a UPF instantiated on the closest MEC server.

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

V. SUMMARY AND DISCUSSION

The objective of our work in this article is to test a realistic NMP deployment, supported by 5G and slicing. Slicing makes it possible to isolate the resources allocated to a set of devices as opposed to sharing all radio resources among all mobile UEs. To the best of our knowledge, this article is the first to

provide quantitative results on the effectiveness of slicing in support of NMPs over 5G deployments. In our setting, the slice dedicated to non-NMP traffic had a lower priority, causing it to adapt to the bandwidth left available after serving NMP traffic. With slicing active, NMP traffic was always served by the UPF of the MEC server, whereas contending traffic was directed to the UPF residing in the CN.

To keep the tests significant in the absence of a dedicated commercial 5G standalone infrastructure supporting slicing, we focus on a co-located CN and MEC infrastructure for radio access tests, while considering extra WAN transit delays from the dedicated GARR infrastructure.

The results of the experiments conducted on the co-located architectures showed that slicing introduced a small but statistically significant latency increase compared to the co-located CN without slicing. This occurred both in the absence and in the presence of contending UDP traffic. In terms of reliability, slicing did not significantly improve the metrics under consideration. This counter-intuitive result can be explained by the following observations.

Because WAN transit delays are not present in our radio access performance tests using co-located infrastructure, the slightly higher latency obtained after activating slicing depends on the architecture and equipment setup. The MEC server managing the NMP traffic is not meant to be as powerful as the CN equipment. Moreover, the UPF running on the MEC is a slimmer version of the UPF running on the CN, as it is designed to manage a smaller data plane than the central node, resulting in slightly but noticeably higher processing delays.

Finally, the data traffic generated by each NMP device amounts to 11.5 Mb/s, in the most complex conditions (see also Table I). Therefore, the effects of slicing the NMP traffic through the UPF hosted on the MEC can be seen mostly in terms of WAN transit delay reduction rather than reliability improvements. In fact, the addition of contending UDP traffic caused the latency to increase both in the presence and in the absence of slicing, but did not affect reliability metrics (see Fig. 6). Significantly, latency and reliability results were found to be uncorrelated irrespective of slicing and contending traffic conditions. This outcome is in agreement with other previous studies conducted on NMP systems connected over 5G networks [18], [19], [20] and suggests that these two KPIs are driven by different root causes.

Moreover, our results show that latency increases with the number of audio streams involved and with the presence of background traffic, whereas the reliability did not substantially vary in the same context. This result is in agreement with the findings reported in our previous study [20], which considered a different architecture without slicing. Furthermore, irregular spikes were found for all latency and reliability metrics, which can significantly reduce the QoS perceived by the users of NMP applications. This aspect is also in common with other studies in [18], [19], [20], and [33].

Taken together, our findings suggest that concepts, such as MEC and slicing in 5G, have the concrete potential to become technology enablers for the IoMusT. Nevertheless, by assessing of the current limitations of slicing-enabled 5G networks in support of NMP sessions, we revealed that

improvements at the hardware and software level are necessary to reduce latency and improve the reliability of the transmitted audio streams, especially when a high number of nodes is involved. This view is also shared with Mohjazi et al. [21], who foresee the solution of the found issues in the provisions of 6G. In general, our results are important for the design of future networks as well as of time-sensitive and mission-critical applications beyond the musical case, such as the Metaverse.

VI. CONCLUSION

This article investigated different architectures supporting NMPs over 5G networks, both with and without slicing. In our setting, slicing was used to allocate resource exclusively to the radio links of the NMPs, while at the same time instantiating the user-plane function on a MEC server, assumed to be geographically closer to the NMP participants than the CN of the mobile operator.

Our results show that slicing, even when introducing slight extra delays due to UPF implementation or hardware specifications, is a potentially excellent choice to save on WAN transit times for NMP traffic while delivering the same reliability as the operator's CN equipment.

NMPs mandate specific QoS requirements related to low-latency and high-reliability communications among all connected musical devices involved. Thus, slicing becomes a key feature to meet QoS requirements, establish a guaranteed data rate, confine contending radio traffic to different radio resources so that they do not interfere with the latency- and loss-sensitive NMP communications, and in general upkeep the required QoS level for an effective musical interaction.

In future work, we plan to investigate the slicing performances in a real public 5G SA network. Moreover, we plan to exploit virtual network functions residing on the MEC in order to mix audio streams from all musical devices, before redistributing them back. This has the advantage of saving bandwidth in the uplink phase, as each NMP device would need to send only one audio stream to the MEC rather than multiple audio streams (one for each other device). Finally, we plan to study the 5G performance at the audio-visual level, where each musician transmits not only the audio streams but also the video streams.

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] L. Gabrielli and S. Squartini, *Wireless Networked Music Performance*. Singapore: Springer, 2016.
- [2] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An overview on networked music performance technologies," *IEEE Access*, vol. 4, pp. 8823–8843, 2016.
- [3] 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.

- [4] L. Turchet and C. Fischione, "Elk audio OS: An open source operating system for the Internet of Musical Things," *ACM Trans. Internet of Things*, vol. 2, no. 2, pp. 1–18, 2021.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] F. Cheli and S. Giordano, "Service parameters identification for adaptive networked music performance," in *Proc. Global Inf. Infrastruct. Netw. Symp.*, 2022, pp. 94–98.
- [10] M. Centenaro, P. Casari, and L. Turchet, "Towards a 5G communication architecture for the Internet of Musical Things," in *Proc. IEEE Conf. Open Innovat. Assoc. (FRUCT)*, 2020, pp. 38–45.
- [11] L. Vignati et al., "Is music in the air? Evaluating 4G and 5G support for the Internet of Musical Things," *IEEE Access*, vol. 12, pp. 38081–38101, 2024.
- [12] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.
- [13] A. Baratè, G. Haus, L. A. Ludovico, E. Pagani, and N. Scarabottolo, "5G technology and its applications to music education," in *Proc. Multi Conf. Comput. Sci. Inf. Syst., (MCCSIS) Int. e-Learn.*, 2019, pp. 65–72.
- [14] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network slicing in 5G: Survey and challenges," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 94–100, May 2017.
- [15] P. Caballero, A. Banchs, G. de Veciana, X. Costa-Pérez, and A. Azcorra, "Network slicing for guaranteed rate services: Admission control and resource allocation games," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6419–6432, Oct. 2018.
- [16] A. E. Garcia et al., "Performance evaluation of network slicing for aerial vehicle communications," in *Proc. IEEE ICC Workshops*, 2019, pp. 1–6.
- [17] E. Kapassa et al., "An innovative ehealth system powered by 5G network slicing," in *Proc. IOTSMS*, 2019, pp. 7–12.
- [18] J. Dürre et al., *In-Depth Latency and Reliability Analysis of a Networked Music Performance Over Public 5G Infrastructure*, Audio Eng. Soc., New York, NY, USA, 2022.
- [19] L. Turchet and P. Casari, "Assessing a private 5G SA and a public 5G NSA architecture for networked music performances," in *Proc. 4th Int. Symp. Internet Sounds*, 2023, pp. 1–6.
- [20] L. Turchet and P. Casari, "Latency and reliability analysis of a 5G-enabled Internet of Musical Things system," *IEEE Internet Things J.*, vol. 11, no. 1, pp. 1228–1240, Jan. 2024.
- [21] L. Mohjazi, B. Selim, M. Tatipamula, and M. A. Imran, "The journey toward 6G: A digital and societal revolution in the making," *IEEE Internet Things Mag.*, vol. 7, no. 2, pp. 119–128, Mar. 2024.
- [22] 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.
- [23] 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.
- [24] N. Bouillot and J. R. Cooperstock, "Challenges and performance of high-fidelity audio streaming for interactive performances," in *Proc. Int. Conf. New Interfaces Music. Exp.*, 2009, pp. 135–140.
- [25] 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, Oct. 1998.
- [26] 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.
- [27] P. Verma, A. Mezza, 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 Innovat. Assoc. (FRUCT)*, 2020, pp. 268–275.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] "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/>
- [32] "Project 5G-TOURS deliverable D4.2: First Touristic city use case results," document G.A.:856950, 5 GPPP Heidelberg, Germany, Oct. 2022. [Online]. Available: <http://5gtours.eu/documents/deliverables/D4.2.pdf>
- [33] 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.*, 2020, pp. 407–412.
- [34] K. Katsaros et al., "5G festival-re-inventing live collaborative performances," in *Proc. IEEE Future Netw. World Forum (FNWF)*, 2022, pp. 7–12.
- [35] P. Bolletta et al., "Considering the next generation of GARR network," WhitePaper, Rome, Italy, Accessed: May 31, 2024. [Online]. Available: <https://www.garr.it/it/documenti/3474-garr-white-paper-maggio-2017>
- [36] "5G; NR; Physical layer procedures for data; (Release 16), Version 16.2.0;" 3GPP, Sophia Antipolis, France, Rep. TS 38.214, Accessed: May 31, 2024.



Luca Turchet (Senior Member, IEEE) received the master's degree (summa cum laude) in computer science from the University of Verona, Verona, Italy, in 2006, the master's degrees (summa cum laude) in classical guitar and in composition from the Music Conservatory of Verona, Verona, in 2007 and 2009, respectively, the master's degree (summa cum laude) in electronic music from the Royal College of Music, 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.

Dr. Turchet is the Chair of the IEEE Emerging Technology Initiative on the Internet of Sounds and the General Chair of the IEEE International Symposium on the Internet of Sounds.



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

He is an Associate Professor with the Department of Information Engineering and Computer Science, University of Trento, Trento, Italy. He was with the University of Padova from 2008 to 2014, and with IMDEA Networks, Madrid, Spain, from 2015 to 2019. His research interests include all aspects of networked communications.

Prof. Casari received two best paper awards. He serves on the editorial board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.