IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 December 2023, accepted 3 March 2024, date of publication 7 March 2024, date of current version 15 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3374641

RESEARCH ARTICLE

Is Music in the Air? Evaluating 4G and 5G Support for the Internet of Musical Things

LUCA VIGNATI[1], **GIOVANNI NARDINI**[2], **(Member, IEEE)**, **MARCO CENTENARO**[3], **(Member, IEEE)**, **PAOLO CASARI**[1], **(Senior Member, IEEE)**, **SANDRA LAGÉN**[4], **(Senior Member, IEEE)**, **BILJANA BOJOVIC**[4], **STEFANO ZAMBON**[5], **AND LUCA TURCHET**[1], **(Senior Member, IEEE)**

[1]Department of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy
[2]Department of Information Engineering, University of Pisa, 56126 Pisa, Italy
[3]Athonet S.r.l., 36050 Vicenza, Italy
[4]Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), 08860 Barcelona, Spain
[5]Elk, 11357 Stockholm, Sweden

Corresponding author: Luca Vignati (luca.vignati@unitn.it)

This work was supported in part by European Space Agency under Grant 4000132621/20/NL/AF, in part by Italian Ministry for University and Research (MIUR) under the Initiative ''Departments of Excellence'' under Grant Law 232/2016, in part by Centre Tecnològic de Telecomunicacions de Catalunya (CTTC) funded by MCIN/AEI/10.13039/501100011033 and ''European Regional Development Fund (ERDF) A Way of Making Europe'' under Grant PID2021-126431OB-I00, in part by Spanish Government under Grant TSI-063000-2021-56/57 6G-BLUR, and in part by Generalitat de Catalunya under Grant 2021 SGR 00770.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethical Committee of the University of Trento.

ABSTRACT The full potential of the Internet of Musical Things (IoMusT) paradigm can be fully unleashed only in the presence of widespread, reliable wireless connectivity, allowing musicians to connect their smart instruments (almost) anywhere they are. For this reason, we propose a realistic, end-to-end communication architecture for a IoMusT system based on public fifth-generation (5G) mobile networks that considers a networked music performance use case, and we introduce a model for the resulting system. We define high-level service requirements and key performance indicators for the network's connect-compute architecture. We evaluate our solution via system-level simulations using the well-known 5G-LENA/ns3 and Simu5G/OMNeT++ frameworks. We found that operating 6 or more IoMusT devices over a fourth-generation (4G) network results in a worst-case latency well over 20 ms for more than 90% of the packet transmissions, while losing more than 10% of packets. On the other hand, operating the same number of devices over a 5G network reduces the latency significantly. After testing our findings both in single-cell and in multi-cell scenarios, assuming a transient in the upgrade of mobile network infrastructures from 4G to 5G, we consider a E-UTRA-NR dual connectivity (EN-DC) scenario, where a 5G base station serves a subset of the users of a 4G cell. Furthermore, we conducted a user study where musicians were asked to assess their playing experience during simulated 4G- and 5G-based networked music performances. The results of the simulations and of the user study consistently indicate that the use of 5G technology improves performance significantly, and advocate the need for a 5G framework to fully support the IoMusT.

INDEX TERMS Telecommunication network performance, ultra reliable low latency communication, numerical simulation, 4G mobile communication, 5G mobile communication.

I. INTRODUCTION

A. GENERAL CONTEXT

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 Internet of Things (IoT) concepts to the musical domain. This yields a global vision commonly termed the IoMusT [1]. According to this vision, future musical instruments and interfaces (Musical Things) will embed intelligence and communications capabilities [2]. These devices will enable not only distributed music performances, but also the active

The associate editor coordinating the review of this manuscript and approving it for publication was Hassan Tariq Chattha.

involvement of an event's audience. The latter may utilize innovative multi-sensory interfaces, both to enrich their own listening experience and to participate actively in the music creation process [3], [4], [5]. An increasing body of literature dealing with IoT-based musical devices [6], [7], [8], [9], [10], communication architectures and protocols [3], [11], [12], [13], [14], [15], [16], packet loss recovery methods [17], [18], synchronization mechanisms [19], [20], studies involving distributed musicians [21], [22], as well as discussion papers [23], [24] confirms the growing interest of the IoT community about the Internet of Musical Things.

The IoMusT may revolutionize the traditional concept of musical interaction in many ways, with an impact on both synchronous and asynchronous interactions between musical stakeholders (e.g., performers, audiences, composers, teachers, students) in various contexts (e.g., education, performance, composition, among others). The so-called networked music performance (NMP) systems, which enable geographically dispersed musicians to play together [25], [26], are prominent components of the IoMusT. NMP-based services are going to be more and more integrated into musical things by leveraging embedded systems dedicated to networked audio processing tasks.

B. WHAT MAKES THE IOMUST DIFFERENT FROM TRADITIONAL AUDIO STREAMING SERVICES?

As far as connectivity is concerned, both wired and wireless networks can connect musical things and support IoMusT traffic. The former option is of course the most effective option performance-wise, however, it can be problematic in certain scenarios, such as when assembling/disassembling musical devices on stage or when dealing simultaneously with different musical tools and resources [27], [28]. Wireless networks provide better opportunities and potential, especially a much smoother user experience in terms of the instrument (self)-configuration, the seamless integration and usage of multiple musical devices and resources, freedom of movement for the musicians, and ubiquitous use of musical devices and associated services. Regardless of the type of connectivity, an effective remote and distributed music performance entails extremely strict quality of service (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 imperceptible signal dropouts) [29]. Satisfying these key performance indicators (KPIs) makes it possible for the performers to play synchronously, maintain a stable tempo and, more generally, experience high-quality interactions [30, Ch. 3].

The above requirements, combined with the continuous and periodic nature of the associated network streams, make IoMusT applications stand out from other audio consumption services relying on streaming over any sufficiently broadband network. Conveying audio chunks sequentially for listening on a device is a comparatively established task, especially when the audio track is pre-recorded. In this case, the listener's device typically allows sufficient playout buffering space to cover delays and temporary outages (or for wireless networks, higher error rates due to low signal coverage).

By way of contrast, supporting potentially many Musical Things that interact in real time and exchange audio streams produced on the spot is a completely different problem. Here, each Musical Thing requires that its own digital audio output be delivered to all other involved Musical Things, and locally mixed with their own audio output. The transfer through the wireless channel must be thus extremely reliable, fast, and should experience little if any outage. Connectivity interruptions may still happen, but should be rare, so that low-complexity error correction schemes can compensate for them. Alternatively, machine learning-based algorithms should be able to fill in comparatively longer gaps within an interrupted audio stream, a task that is also best executed if errors are rare and not exceedingly bursty. In addition to the above, Musical Things may have limited computational resources, especially if they are based on highly embedded computing boards, calling for audio services to be preferably located at the edge of the network.

C. 5G AS AN ENABLER OF THE IOMUST

While there exist some preliminary experiments using cellular networks as a connectivity provider for IoMusT scenarios (see, e.g., [31], [32], [33], [34]), the potential of wireless cellular systems in this context remains largely uninvestigated, especially as far as the multiple users scenarios in realistic conditions are concerned.

In cellular networks, smart musical instruments (SMIs) (as well as other kinds of Musical Things) can be seen as a completely new class of user equipments (UEs) [2]. Thanks to cellular radio's plug-and-play concept, SMIs can achieve end-to-end connectivity with minimal configuration efforts both on the device side and on the network side, assuming to exploit the publicly available network infrastructure. In particular, there are significant expectations that the 5G System (5GS) [35] can become a key enabler for IoMusT scenarios. The standardization body of cellular networks, the Third Generation Partnership Project (3GPP), developed the 5GS as the successor of the 4G evolved packet system. As such, the 5GS brings many novelty aspects in both the radio access network (RAN) and the core network.

As for the RAN, 5G's so-called New Radio (NR) air interface [36], [37] introduces a redesigned and flexible orthogonal frequency-division multiple access structure, both in the time domain and in the frequency domain. This is achieved through the introduction of short transmission time intervals and additional subcarrier spacing options. As a shorthand for different compatible parameter configurations, the NR standard defines three so-called numerologies μ, where $\mu = 0$, configures the air interface to be compatible with the Long-Term Evolution (LTE) standard. Instead, higher values such as $\mu = 1$ and $\mu = 2$ reshape transmitted signals by reducing the symbol time while increasing the bandwidth occupation [36], [38]. Among other consequences, increasing the numerology μ enables faster transmissions both by allotting additional bandwidth to each scheduled UE, and by making scheduling decisions more often over time. Moreover, reduced processing times at both the UE and the base station, grant-free transmissions [39], antenna diversity, and multi-connectivity [40], make it possible for the 5G NR to meet the latency-reliability constraints [29], and to become a relevant enabler for networked musical interactions [41], [42].

As for the core network, the 5G service-based architecture seamlessly integrates multi-access edge computing (MEC) platforms into the 5G core administrative domain [43]: this makes it possible for the MEC host to interact with the core network, negotiate traffic and workload routing policies, as well as provide or exploit value-added services [44], [45], [46], [47], [48]. Because the IoMusT is a highly time-critical application, the MEC's role is as key as the RAN's or the transport network's: in fact, the MEC is a perfect candidate location to receive and mix synchronous audio streams, as well as implement more advanced or machine learning-based functions. An example of the latter includes filling audio gaps originating from, e.g., bursts of wireless transmission errors, unrecoverable packet losses, and severely out-of-order packet deliveries.

D. MOTIVATION OF THIS STUDY

Three very common misconceptions arise when relating 5G technology to IoMusT scenarios.

Common understanding #1: *5G will support augmented reality (AR)/virtual reality (VR), of which the IoMusT is just an easier variant.*

Actual situation: IoMusT traffic patterns and requirements are very different from those of AR and VR applications. IoMusT devices are operated by interactive performers that generate a continuous flow of musical data. The network needs to (*i*) deliver them without errors (lest an excessive use of audio breakage concealment algorithms makes the transmission errors perceivable by the performers and audience); and (*ii*) make sure all interacting performers receive the audio data within a very limited time span (lest it becomes impossible for them to synchronize and correctly maintain a stable tempo). If the audio hauling is organized well, the uplink and downlink bandwidth requirements are thus the same, as is the rate of data packets being sent and received.

Comparatively, AR and VR data flows mostly downlink from the base station to the users wearing a visor, and serve to manage how the visor offers an interactive experience to the human in augmented or fully virtual environments. In multiple-access networks without guaranteed bandwidth and packet loss performance such as 5G networks, centralized servers typically receive compressed sensor data and process it in order to identify relevant interactions in the virtual world. The corresponding interaction data is sent back to the users' VR visors, so that the software therein can give the appropriate visual and haptic feedback to the user. Because most of the graphical processing takes place not remotely but in local devices, flows are not continuous and the amount of transferred data may be even less than the audio flows of IoMusT deployments. This setup is conceived to minimize data transfer at the expense of heavier local processing and rendering at the user's visor, and is designed to meet the real-time requirements of currently available visors and networks.

Common understanding #2: *5G will support both massive IoT and industrial IoT services; IoMusT is just a trivial example of either of the two, depending on the specific use case.*

Actual situation: On the one hand, the massive IoT service enabled by 5G satisfies the requirements of a huge amount of end devices requesting access to the network to send sporadic traffic in uplink, with no significant QoS requirements in terms of reliability and latency. On the other hand, 5G-powered industrial IoT is designed to support end devices employed in localized industrial processes, thus demanding extremely high reliability and the lowest possible delivery delay.

For IoMusT, however, a peculiar working point for the mobile network needs to be established, as typically:

- the number of involved end devices may be similar to or larger than industrial IoT but lower than massive IoT;
- packet reliability needs to be much higher than massive IoT, but not necessarily as high as in industrial IoT;
- packet latency needs to be bounded with respect to massive IoT, but does not need necessarily to be pushed to the minimum like in industrial IoT.
- IoMusT data flows are continuous and periodic, with very short periods, which differentiates the IoMusT from other (albeit mission-critical) industrial IoT deployments.

In other words, IoMusT *represents a quite different tradeoff* between the performance requirements of massive and industrial IoT in 5G, thus it cannot be straightforwardly mapped to either of them.[1] It is also worth remarking that a simplistic worst case analysis, whereby a 5G network configured for industrial IoT (IIoT) is deployed to accommodate IoMusT traffic, is not applicable in our case. In fact, we aim at investigating the potential of the publicly available network infrastructure to serve IoMusT users with minimal configuration efforts.

Common understanding #3: *Legacy 4G technology already delivers sufficiently high performance for supporting IoMusT.*

Actual situation: Traditionally, the air interface latency is the bottleneck of the overall packet delivery delay in mobile networks. In this regard, 5G's NR improves the legacy LTE air interface in several respects: higher numerologies pair up with evolved capabilities of UEs and base stations, so as to

---

[1] We will elaborate extensively on the use case for IoMusT in Section III-A.

reduce the latency experienced with 4G. Therefore, as we will show later in this paper, 4G does not possess the technical features to meet the requirements of IoMusT.

The cellular network ecosystem undergoes a relentless evolution, as telecommunication operators eagerly compete to be the first that offer faster connections, improved coverage, or new services to their customers. However, adopting new technologies requires huge investments. For this reason, deployment plans span several years, and typically prioritize densely populated areas. 5G's deployment will make no exception: several companies have been installing their first 5G cell towers, but 5G will represent no more than 3.5% of the total mobile network connections in 2023, according to the Cisco Global Mobile Data Traffic Forecast Update [49]. Moreover, so far most of the 5G connectivity was granted by non-standalone (NSA) deployments, i.e., those providing 5G radio access to a 4G core network (CN) [50]. These facts confirm a very important point: 5G and 4G networks will co-exist for many years, before standalone (SA) deployments with fully 3GPP-compliant 5G RAN and CN become the de-facto standard for operational 5G networks.

For the above reasons, 3GPP standardized EN-DC, whereby 5G base stations (BSs) (the so-called next-generation NodeB (gNB) in 3GPP jargon) are deployed along with an incumbent 4G system, thus providing 5G access alongside 4G evolved universal terrestrial radio access (E-UTRA). This configuration allows network operators to deliver localized but faster 5G access, while maintaining broader coverage at lower data rates via 4G.

The main research questions. The above discussion suggests that, to date, it remains uncertain whether 5G-enabled music performances will be open as a massive access application, or rather if they will only be supported as special services with specific QoS requirements. In fact, there does not exist any comprehensive performance evaluation of large-scale, public 5G networks in serving IoMusT users; the few IoMusT tests available to date are limited in many aspects. For example, early demonstrations such as [32] focus on the NMP system architecture, and employ a single 5G connection between a single device and a co-located 5G base station. Therefore, they neglect multiple-access and inter-cell interference, as well as non-IoMusT air traffic. Other focused demonstrations such as [31] create many-to-many traffic patterns, where an IoMusT users send unicast data flows to (and receives from) all other users. This reduces the scalability of the system as the number of performers increases.

Our main research questions in this paper are aligned with the currently open problems in 5G support for IoMusT, namely:

1) Will 5G be sufficient to cover the requirements of IoMusT deployments?
2) Can we quantify the characteristics of the service that 5G will provide to IoMusT users in terms of metrics of interest for musical performances, e.g., performer-to-performer delay (which relates to the feasibility of the performance itself) and packet loss probability (which relates to the quality of the sound perceived by the connected performers and audience)?
3) Do intermediate 5G deployments offer sufficient resources to support IoMusT users during the transition from 4G to 5G?

E. CONTRIBUTIONS

Our main purpose in this paper is to answer the above questions by using validated and community-accepted simulation tools to reproduce multi-IoMusT user 5G scenarios that would exceed the capabilities of current commercial deployments. We aim to show that 5G will be a fundamental enabler of the IoMusT paradigm, that will overcome the packet latency and reliability limitations of current 4G systems [51]. In doing so, simulation enables us to surpass the limitations of current IoMusT deployments (e.g., bandwidth availability, number of deployed devices, specifications of each device), and test scenarios as complex as are likely to appear in future IoMusT contexts.

The simulators employed in this study are ns-3 [52] and OMNeT++ [53]. These frameworks are by far the most used discrete event simulation tools to evaluate large-scale and complex network and communication technologies. Both of them feature specifically developed extensions that cover 5G systems, namely 5G-LENA [54] for ns-3, and Simu5G [55] for OMNeT++. Using community-validated simulators makes it possible to realistically reproduce large-scale 5G scenarios and yields representative quantitative results. To ensure this, we avoid beyond-5G (B5G) or 6G solutions, which are still under discussion and have not reached a standardization stage yet.

As will become clearer later, 5G-LENA and Simu5G have been developed independently, and thus provide different sets of functionalities. By working with both of them, we can thus cover a larger spectrum of realistic 5G network features. Moreover, as a non-secondary outcome of our research, we discuss how the main differences between ns-3 and OMNeT++ affect the evaluation of IoMusT scenarios, in the hope to encourage the 5G-and-beyond research community to adopt either tool for their own evaluations related to IoMusT deployments.

In our previous conference paper [56], we focused on requirement analysis and on a high-level architectural description. In this paper, we cover recent advances in the field, and then move the focus towards the performance of different cellular architectures and on a *quantitative* comparison of how 4G and 5G networks support the requirements of the IoMusT. In particular, we assess the (baseline) performance of 4G systems, quantify the benefit provided by 5G, and identify the most suitable configuration for a 5G-enabled IoMusT both in the long run and during the expectedly long transition phases between 4G and 5G.

A summary of our key contributions is as follows:

- We review high-level IoMusT requirements and translate them into network-level system requirements;

- We outline a feasible 5G network configuration including the most relevant IoMusT components and their interactions;
- We define relevant KPIs to evaluate IoMusT system performance;
- We provide a quantitative performance evaluation in realistic scenarios using two community-validated simulation frameworks; we configure each of them to replicate our defined scenarios and to embed typical response times for key system components.
- We provide a realistic vision of the transitional stages of 5G network deployments in support of the IoMusT, and specifically discuss the role and impact of the 5G EN-DC architecture.

The remainder of this paper is organized as follows:

- Section II introduces the architecture of the envisioned 5G enabled IoMusT, as well as the relevant KPIs and high-level service requirements.
- Section III discusses the relevant works related to 5G, NMP and IoMusT.
- Section IV presents our characterization of the IoMusT service coverage, including the details on how simulations are set up and the simulation results.
- Section V covers the performance evaluation of the IoMusT service in a mixed 4G/5G scenario using EN-DC. Again, details on how simulations are set up are provided as well as the simulation results.
- Section VI describes a user study where musicians were asked to assess their playing experience during simulated 4G- and 5G-based networked music performances.
- In Section VII, we present our conclusions on this subject and possible future work.

II. RELATED WORK

A. NETWORKED MUSIC PERFORMANCE ENABLERS

Candidate wireless communication standards for NMP span both short- and long-range systems. As Gabrielli and Squartini discuss in their survey [30, Ch. 4], both proprietary audio-specific solutions and IEEE 802.11a/b/g/n/ac and such technologies as IEEE 802.11af may support NMP. Short-range technologies operate mainly in the 2.4 GHz, unlicensed frequency band.

Public IEEE standards from the IEEE 802.11 family present issues when applied to NMP. Specifically, i) they suffer from high interference from coexisting traffic in unlicensed frequency bands, ii) impose a comparatively large channel access latency, which increases for larger network sizes, and iii) lead to multiple-access interference, e.g., due to well-known issues such as the presence of hidden terminals. Even if the IEEE 802.11af standard operates on sub-GHz frequency bands with good propagation properties (previously used for television broadcast), such sub-GHz bands are unevenly available in different countries, making it impossible to achieve a fully interoperable NMP technology. Proprietary solutions also present issues for NMP, as they usually support just unidirectional transmissions.

Because wireless technologies present such limitations, most existing NMP frameworks rely on wired networks [26, Tab. 3]. The "LOw LAtency audiovisual streaming system" (LOLA) is one of the most advanced solutions in this respect. LOLA is specifically conceived to enable networked interaction distributed performing arts [57], [58]. However, reliance on wired infrastructure and using specialized hardware limits LOLA's scalability and cost-effectiveness.

IoMusT applications require embedded digital audio platforms that can sample and process audio data within a minimal amount of time [59], [60]. Among other recent advances in this field, we mention the Bela board [61], which provides a Beaglebone Black cape for audio processing under low latency constraints. A typical drawback of embedded systems for audio processing is that they support limited connectivity options. An exception in this respect is Elk's Audio OS, an embedded Linux-based operating system that achieves sub-millisecond audio processing latency [62]. Besides these optimizations, Elk's Audio OS also supports local and remote connectivity options using a number of well-known standard networking protocols.

B. STANDARDIZATION EFFORTS IN THE IOMUST FIELD

In a dedicated technical report [63], 3GPP addresses the potential of 5G-enabled audiovisual content distribution and service provisioning. They describe several use cases, each requiring different proportions of advanced signal processing (e.g., high-quality audio/video acquisition and mixing) and network distribution services (e.g., audio/video stream dispatching).

A specific use case deals with audio streaming in live performances [63, §5.2], as would take place in live on-stage events, taking place in front of an audience. This is the most relevant use case for the IoMusT paradigm in 3GPP's official documents. In this use case, UE-type audio sources (e.g., microphones) generate multiple audio streams, which are mixed and returned to the musicians via additional wireless UEs, such as in-ear monitors. The architecture and capabilities of the 5G infrastructure enable this use case. The report also lists a number of KPIs and system parameters to achieve effective audio streaming during live performances [63, Tab. 5.2.1-1], in order to provide a satisfactory quality-of-experience for both the performers and the audience.

In addition to the above, the European Telecommunications Standards Institute (ETSI) has recently formed an industry specification group working on augmented reality scenarios and applications. The purpose of this group is to define a framework to inter-operate devices, systems, and services in this scenario [64]. It is highly likely that initiatives such as the above can give additional impulse to the whole extended reality ecosystem, including virtual reality. For example, new kinds of musical experiences can be encompassed by envisioning the use of virtual avatars or virtual objects [1], [65], enabled by the real-time, 5G-mediated inter-working between IoMusT and augmented/virtual reality devices.

While the above documents cover a few IoMusT use cases, the list is not exhaustive, and there are several more that can be of interest to the IoMusT community. Presenting different scenarios for one such new use case is among the objectives of our work in this paper.

C. SYSTEM-LEVEL 5G NETWORK SIMULATION

Simulations represent effective and low-cost tools for evaluating the performance of networked systems. In the context of wireless networks, we typically distinguish between *link-level* simulators (e.g., the TU-Wien Vienna 5G framework [66]) and *end-to-end system-level* simulators. The former accurately reproduce physical layer functionalities such as antenna designs and transmission schemes in order to measure such quantities as the signal-to-interference-plus-noise ratio (SINR), the spectral efficiency, etc. End-to-end system-level simulators, instead, rely on simplified yet comprehensive physical-layer models to reproduce the whole network protocol stack, up to the application logic. This makes it possible to simulate complete network deployments, possibly at large scale, and evaluate higher-layer metrics such as the application throughput, end-to-end latency, and so forth. Assessing whether 5G networks can support the IoMusT, requires us to evaluate application-level metrics, and their interplay with technology-specific parameters. Hence, end-to-end simulators are the most appropriate tools for this purpose.

Examples of end-to-end 5G system-level simulators are 5G-air-simulator, 5G-LENA, and Simu5G. 5G-air-simulator [67] is a standalone simulator that models the whole NR stack, but lacks a model for EN-DC. Moreover, simulation workflow automation is not as developed, which makes large-scale simulation campaign management more challenging. 5G-LENA [54] complements the well-known ns-3 framework (https://www.nsnam.org) by providing a detailed model of the physical and medium-access control (MAC) layer of 5G NR, including millimeter wave and bandwidth parts management, for NSA 5G deployments. The simulator has been initially developed in collaboration with several companies, and later released open-source under the GPLv2 license. 5G-LENA is compliant to 3GPP's technical specifications, and simulates many details of 5G communications and networking protocols, including 3GPP spatial channel and propagation models, 3GPP antenna array models, multiple numerologies, FDD and TDD, uplink/downlink slot formats, beamforming, as well as standard-compliant access schemes, including OFDMA and TDMA scheduling. Also, extensions of 5G-LENA to simulate NR V2X (vehicular-to-everything) communications [68] and NR-U (NR in unlicensed) technologies [69] are available.

Simu5G [55] is a model library for the OMNeT++ simulator (https://omnetpp.org), which started as the result of a joint research project carried out by Intel Corporation. Simu5G's modular architecture provides a large set of LTE and NR functionalities that includes support to both FDD and TDD, multiple numerologies, carrier aggregation, EN-DC and ETSI specifications-compliant MEC, SA and NSA 5G network deployments, as well as mixed 4G/5G scenarios. Protocol implementations cover all layers of the LTE and NR protocol stacks, from Packet Data Convergence Protocol (PDCP) to the physical layer.

III. 5G-ENABLED INTERNET OF MUSICAL THINGS ARCHITECTURE AND KEY PERFORMANCE INDICATORS

The communication architecture we envision for a 5G-enabled IoMusT includes several key components. With reference to Fig. 1, the most important elements are:

- *UE-type musical things*, that include audio I/O hardware, a processing system (e.g., Elk's Audio OS), and a 5G communication device;
- a *5GS* enabling low-latency packet delivery and highly reliable wireless communications. The system combines two macro-components: i) a next-generation radio access network (NG-RAN), formed of multiple 5G base stations – the so-called gNBs – and ii) a 5G core network. The latter transfers (digital) audio traffic among musical things. The transfer can be mediated by audio application services, such as in-network stream processing and content caching;
- *cloud computing platforms*, which host such application services. These platforms may be located either in the remote cloud (e.g., a centralized data center) if they perform latency-tolerant tasks, or at the edge of the network if they perform latency-critical tasks. Because the IoMusT demands real-time audio processing and transfer, in this paper we will focus on edge-based service provisioning.

A. USE CASE DESCRIPTION

In this work, we consider a key use case for the IoMusT, namely *low-latency NMP (LL-NMP)*. In this use case, a distributed set of IoMusT users rehearse and play together without being co-located (whence the NMP focus), while UE-type SMIs are networked together by the 5GS. Fig. 2 summarizes the typical data flows among IoMusT users. Each source of audio streams (e.g., each performer) transmits a flow of packetized digital audio samples through the 5GS. These flows reach an edge-located MEC server running audio application services such as re-synchronization/mixing, or other advanced functions to conceal missing stream sections caused by transmission errors or out-of-order packet delivery [17], [70]. The output of the audio services running on the MEC server is a single, synchronized, and mixed audio stream, which is returned to each performer through the 5GS, so that the delivery happens within a maximum tolerable delay. Completing the above time-sensitive tasks in due time makes it possible to maintain a smooth musical interaction. In some cases the MEC server may be optional, e.g., when there exist only two performers and no other

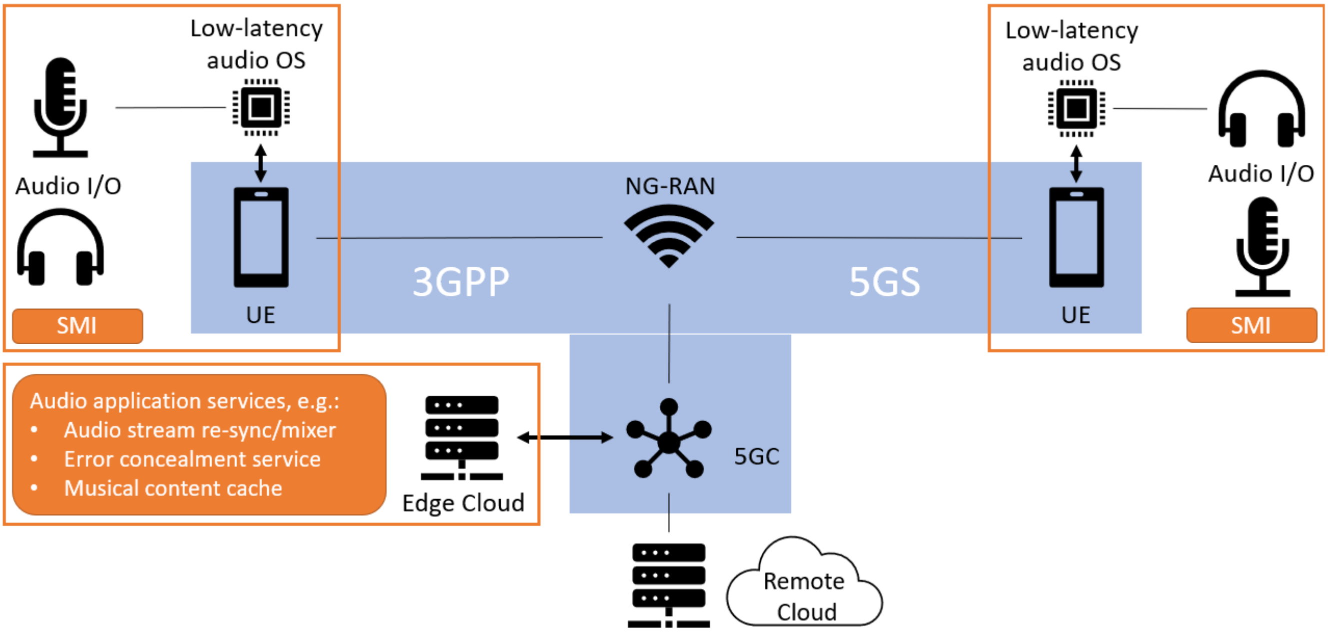


FIGURE 1. 5G-based IoMusT communication architecture overview, entailing UE-type SMIs, 3GPP 5GS, and cloud computing platforms (remote and edge). (Adapted from [56].)

processing or storage services are required. In such cases, the 5GS straightforwardly routes all source-generated audio streams to all IoMusT users.

We remark that the above configuration is very different from state-of-the-art NMP setups [62], which require each of the N_{UE} UEs involved in an NMP to send a separate audio stream to every other UE. In such setups, data traffic rates grow as $\mathcal{O}(N_{\text{UE}}^2)$. Using a properly located MEC server running a jitter and remixing buffer, instead, requires a single upstream flow from each UE to the MEC server, and a single (mixed) flow from the MEC server to each UE, reducing the traffic rate growth to $\mathcal{O}(N_{\text{UE}})$.

B. KEY PERFORMANCE INDICATORS

1) LATENCY

Several components concur to the end-to-end audio communication latency. The time span of interest here starts from the epoch an SMI generates some (analog) audio signal, and ends when the mixed audio stream containing contributions from all performers is delivered to a musician. We identify the following key delay components:

$$\mathcal{D} = \tau_{\text{audio,upstream}} + \tau_{\text{tx,uplink}} + \tau_{\text{transport}} + \tau_{\text{proc}} + \tau_{\text{tx,downlink}} + \tau_{\text{audio,downstream}}, \quad (1)$$

where

- $\tau_{\text{audio,upstream}}$ is the time that a musical thing takes to acquire and digitize an analog audio signal. Note that the so-called *transfer interval* [63], which regulates the traffic pattern towards the 5G module, is deterministic;
- $\tau_{\text{tx,uplink}}$ encompasses the UE processing delay, the transmission time, and the processing time at the gNB side;
- $\tau_{\text{transport}}$ is the delay component due to the transmission of the audio stream from the gNB serving the transmitting UE towards the gNB connected to the receiving UE. This component also includes the delivery delay towards any MEC server that processes the data before sending it to the destination;
- τ_{proc} is the time that the MEC server requires in order to process incoming audio streams (e.g., to re-synchronize/mix streams sent by different musical things, or to retrieve and serve any cached audio contents);
- $\tau_{\text{tx,downlink}}$ is the counterpart of $\tau_{\text{tx,uplink}}$, and includes the gNB processing delay, the transmission time, and the UE processing delay. Due to the different direction of the transmission (downlink vs. uplink), in general it holds that $\tau_{\text{tx,downlink}} \neq \tau_{\text{tx,uplink}}$;
- $\tau_{\text{audio,downstream}}$ accounts for the operations that the receiving musical thing (e.g., an SMI) must perform to serve an analog audio signal to the musician. Note that this delay depends on whether a MEC server intervenes or not to synchronize/mix audio streams. In the absence of the MEC server, each receiving musical thing would perform the above operations independently, so that typically $\tau_{\text{audio,downstream}} \neq \tau_{\text{audio,upstream}}$.

2) RELIABILITY

This KPI generally measures how long a given system performs its intended function under well-defined conditions.

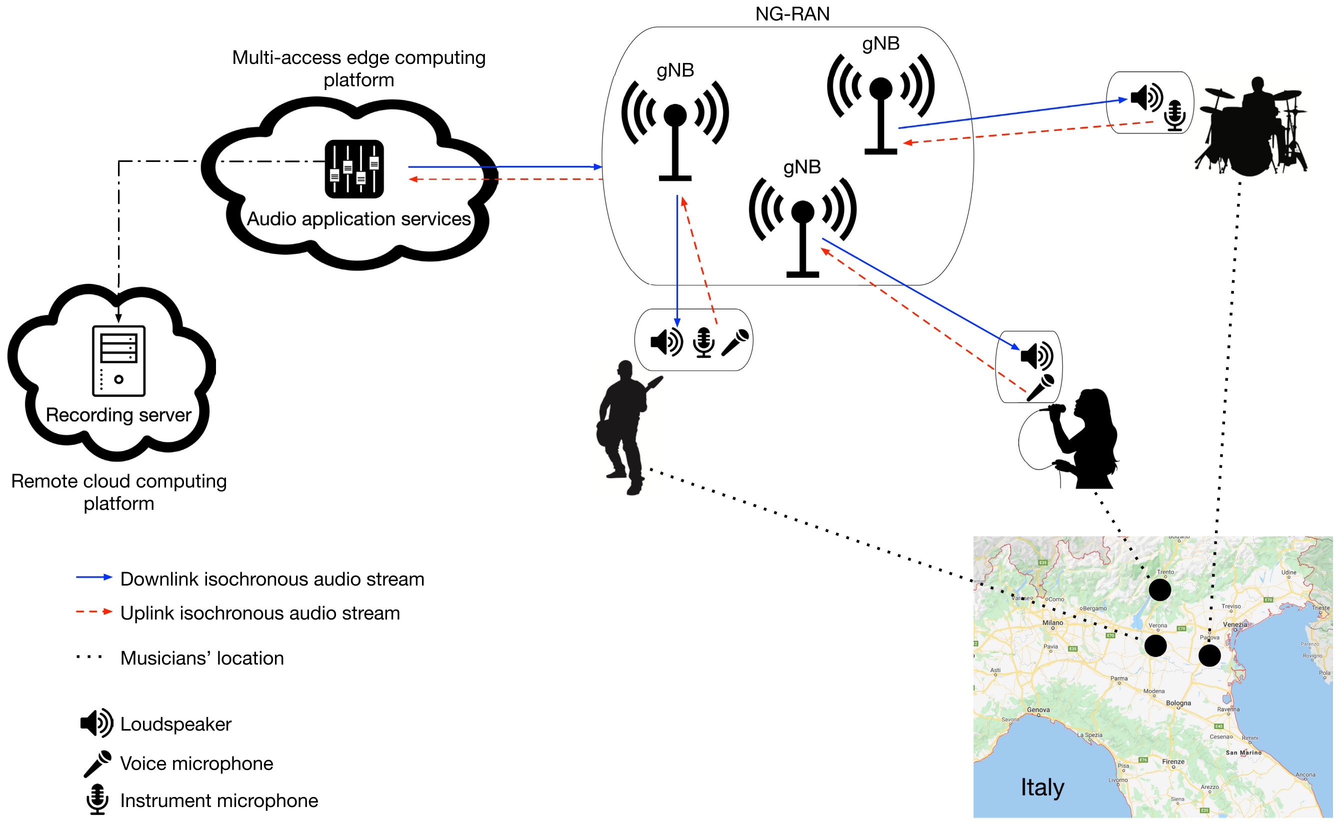


FIGURE 2. **Illustration of the envisioned LL-NMP use case. In this sample scenario, three players are distributed in a regional area of Northern Italy and exploit UE-type microphones and loudspeakers to perform together. The 5GS infrastructure is in charge of managing the generated audio flows (singing voices and musical instruments signals). A MEC platform located in the proximity of the NG-RAN, and connected to it through the 5G core can provide additional processing or advanced services (e.g., adding extra effects or processed versions of pre-recorded tracks). A remote cloud platform may also carry out latency-tolerant tasks, e.g., recording and storing the musical session. (Adapted from [56].)**

This concept is coupled with that of *availability*, which instead is a measure of the percentage of time the system is in an operable state. We stress that a reliable system has also high availability, while a highly available system may not be reliable. Thus, in the following, we refer to the concept of reliability as a KPI.

In terms of network-layer communications, the reliability is typically measured in terms of network-layer packets which reach another system entity within the time constraint required by the desired service, relative to the total number of network-layer packets sent [71, §3.1]. We identify several components that concur to the overall reliability as follows:

$$\mathcal{R} = p_{\text{succ,uplink}} \cdot p_{\text{succ,transport}} \cdot p_{\text{succ,comp}} \cdot p_{\text{succ,downlink}}, \tag{2}$$

where

- $p_{\text{succ,uplink}}$ is the success probability of the uplink transmission;
- $p_{\text{succ,transport}}$ is the success probability of the packet forwarding across the backhaul transport network;
- $p_{\text{succ,comp}}$ represents the probability of error-free data processing at the computing platform side;
- $p_{\text{succ,downlink}}$ is the success probability of the downlink transmission.

In practice, we can assume that $p_{\text{succ,transport}} \rightarrow 1$ as the backbone transport network is highly reliable. However, typically $p_{\text{succ,comp}} \neq 1$ for several possible reasons, e.g.: not all audio data to be mixed is received at the MEC server; error concealment techniques fail to compensate for audio gaps; computational resources at the serving MEC server are heavily loaded and unable to complete computational tasks timely. Thus, the reliability of a 5G-enabled IoMusT can be approximated as follows:

$$\mathcal{R} \simeq p_{\text{succ,uplink}} \cdot p_{\text{succ,comp}} \cdot p_{\text{succ,downlink}}. \tag{3}$$

In other words, we can safely assume that the overall reliability is due to the reliability of uplink/downlink radio links, and the dependability of the edge computing platform.

3) SERVICE COVERAGE

In 3GPP's technical documents, the term "service area" commonly refers to a geographic region where a 3GPP communication service is accessible [71]. For 5G-enabled IoMusT, the concept of service area more closely concerns the topology and transport delay of the network that connects the performers, rather than their geographic distance. For this reason, in the following we will denote the accessibility of IoMusT services as "service coverage", and we will characterize this KPI thoroughly by means of the simulation results presented in Sections IV and V.

First, we discuss an estimate of the audio communication time. We consider a low-latency audio operating system that produces a protocol data unit (PDU) comprising 32 audio samples (each requiring 24 bits, i.e., 3 bytes) for each audio channel, plus some redundancy that helps compensate for packet losses (88 bytes, including the UDP header). Since two audio channels are considered in a stereo setting, the total PDU size is $32 \cdot 3 \cdot 2 + 88 = 280$ bytes. For a sampling frequency of 48 kHz, the packet transmission rate is one packet every $32/(48 \cdot 10^3) \approx 0.67$ ms. For instance, this is in line with the operational parameters of recent commercial platforms [72], and provides the value of $\tau_{\text{audio,upstream}}$. Therefore, we can safely assume that $\tau_{\text{audio,upstream}} \ll \mathcal{D}$. We may also assume that $\tau_{\text{audio,downstream}}$ has the same order of magnitude as $\tau_{\text{audio,upstream}}$. Therefore, the main delay components are due to over-the-air transmissions, backhaul routing, and processing. While the operations of an SMI on audio streams are inherently local tasks, the transmission of audio samples over the air requires the multiplexing of traffic from multiple users, which places the burden on the RAN and transport networks. In particular, multiple IoMusT users co-located in the same geographical area could potentially lead the 5GS to worse latency ($\tau_{\text{tx,uplink}}$ and $\tau_{\text{tx,downlink}}$) and reliability performance ($p_{\text{succ,uplink}}$ and $p_{\text{succ,downlink}}$). An increasing background load of the gNBs that also serves IoMusT devices would have consequences on the computing platform, as well. In both the envisioned use cases, an edge computing platform [43] may be instrumental to achieving low-latency IoMusT user interaction. With reference to Fig. 2, it is reasonable to assume that each MEC host will serve a certain geographical area, thus if the amount of IoMusT users in that area increases, the computing load will increase as well. An increasing offered load leads to a higher latency component $\tau_{\text{transport}}$, but may also impact the reliability term $p_{\text{succ,comp}}$.

Based on the above observations, a limited number of IoMusT users would result in $\tau_{\text{tx,uplink}}$, τ_{proc}, $\tau_{\text{tx,downlink}}$, $\tau_{\text{audio,downstream}} \rightarrow 0$. As a result, $\mathcal{D} \simeq \tau_{\text{transport}}$, thus almost the entire latency budget may be employed to cover the round-trip time between each user and the serving cloud computing platform, which performs audio processing (e.g., syncing, mixing, error concealment, caching). In this respect, we remark that the geographical distance between each IoMusT user and the server depends on the backhaul network topology. Therefore, the definition of *proximity* among performers relates to *traffic routing* delay, rather than to geographical distance [73]. Thus, in the absence of *closely or fully integrated* administrative network domains, a given use case may be feasible or not.

According to the above reasoning, we can measure the IoMusT service coverage based on the following two parameters:

- the delay components related to the network uplink, downlink and transport, i.e., not including the downstream and upstream audio processing times at the SMIs;
- the probability of error related to communication components, i.e., assuming a sufficiently large MEC processing capacity, that ensures error-free data processing.

The following subsection discusses practical limits for these two metrics that would lead to an acceptable service for LL-NMP applications.

C. HIGH-LEVEL SERVICE REQUIREMENTS

According to several studies, rhythmic synchronization among multiple performers is optimal as long as the acoustic delay remains below 20-30 ms. Skilled musicians may even tolerate an absolute maximum of 50 ms of delay, without affecting the music performance [26]. For the LL-NMP use case, we may therefore set the total delay budget to an intermediate value of

$$\mathcal{D} \leq 20 \text{ ms}. \tag{4}$$

As far as the packet reliability $\mathcal{R}$ is concerned, air interface reliability lower bounds between $1 - 10^{-3}$ and $1 - 10^{-9}$ typically satisfy the so-called *Tactile Internet* paradigm [74], of which the IoMusT is part. However, such service guarantees typically hold only under well-defined conditions and depend on many factors, e.g., the varying offered traffic load. In any case, as also seen for the latency, the upper bounds on the reliability value should be a function of the audio application. In particular, assuming that satisfactory error concealment techniques are applied at the application level in the MEC server and that we transmit packets of duration less than 1 ms, we may relax the reliability constraint to

$$1 - \mathcal{R} < 10^{-2}. \tag{5}$$

IV. SERVICE COVERAGE CHARACTERIZATION

The IoMusT service coverage strongly depends on the considered scenario and network deployment conditions, including the coverage of the 5GS and the location and instantaneous load of MEC servers. Therefore, the evaluation of service coverage is an interesting open research question.

In the following, we start filling this gap by performing system-level simulations using the well-known ns-3 software, and specifically the modules LENA [75] (for 4G communications) and 5G-LENA [54] (for 5G communications). The ns-3 framework is a community-supported, availed, and continually maintained open-source simulation engine, which enables us to obtain relevant results in the absence of a deployment with real devices. 5G-LENA is the natural evolution of LENA, that was initially developed to implement the radio access and core networks of 4G, and which has been extended to implement the fundamental 5G PHY-MAC features in line with NR specifications. With our analysis, we evaluate the likelihood that the 4G or 5G network provides users with sufficiently low delay and probability of reception error, so as to satisfy the main KPIs for an effective IoMusT service. Ultimately, we wish to evaluate the extent to which the current 4G and the upcoming 5G mobile network technologies can support LL-NMP.

A. SIMULATION SETUP

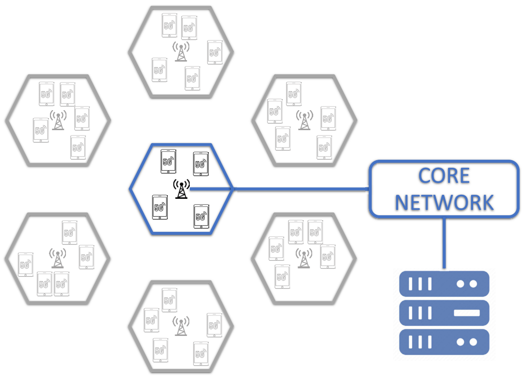
For our simulation study, we implement two simulation scenarios.

- *Scenario 1* represents a single-cell distribution of performers. Here, multiple IoMusT users connect locally to the same gNB. The gNB is located amidst six other cells, which generate interfering traffic. This scenario tests the stress on the RAN in minimal transport delay conditions.
- *Scenario 2* represents a relatively lower-density but more complex multi-cell performer distribution. Here, multiple IoMusT users are distributed over a wider area and connected to different gNBs, which forward the audio flows to a common MEC server. A tier of nine cells surrounds the central cells where the performers are located. Each UE in these cells and the gNBs themselves, therefore, constitute a source of interfering traffic.

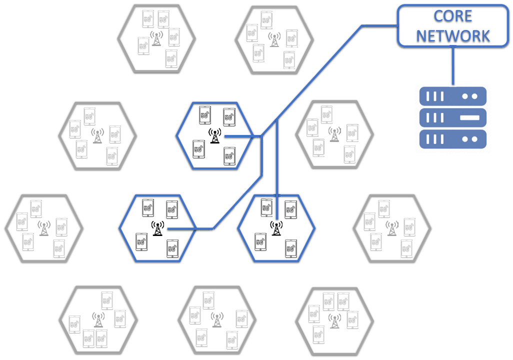
Fig. 3 illustrates the above scenarios.

We deploy N_{BS} cellular base stations ($N_{BS} = 1$ for scenario 1, whereas $N_{BS} = 3$ for scenario 2). Depending on whether we are testing the 4G or the 5G network architecture, these base stations are evolved NodeBs (eNBs) or gNBs, respectively. In either case, base station antennas are installed at a height $h_{\mathrm{BS}} = 25$ m. We randomly generate the locations of the N_{UE} SMIs within a circle of diameter $L = 150$ m, centered on each base station. Moreover, in scenario 2, the UEs are randomly spread throughout the deployed base stations. We assume that all UEs are static.

We use the Urban Macro (UMa) wireless channel model, featuring a path-loss component that depends on the distance between each UE and the base station, and a shadowing component that follows a log-normal distribution with standard deviation $\sigma_{\mathrm{dB}} = 8$ dB. In both scenarios, we adopted a Monte-Carlo approach, repeating the experiments 60 times, each with a different random number generator seed.



(a) Scenario 1



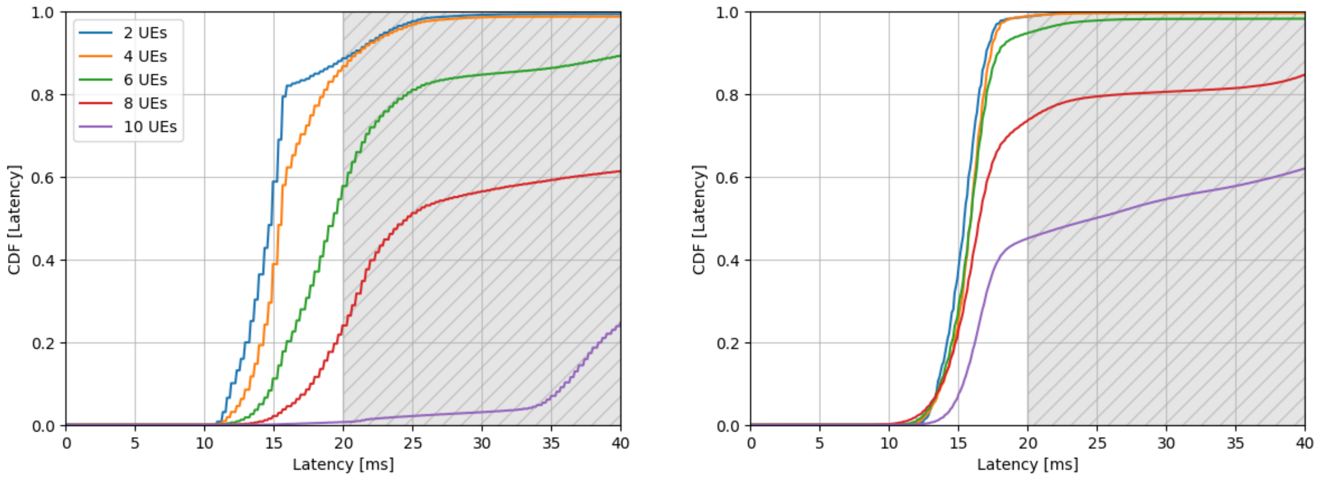
(b) Scenario 2

FIGURE 3. **Representation of the two envisioned service coverage scenarios. For Scenario 2, we assume that the cells (each identified with a different color) are spaced apart to resemble the deployment in Fig. 2.**

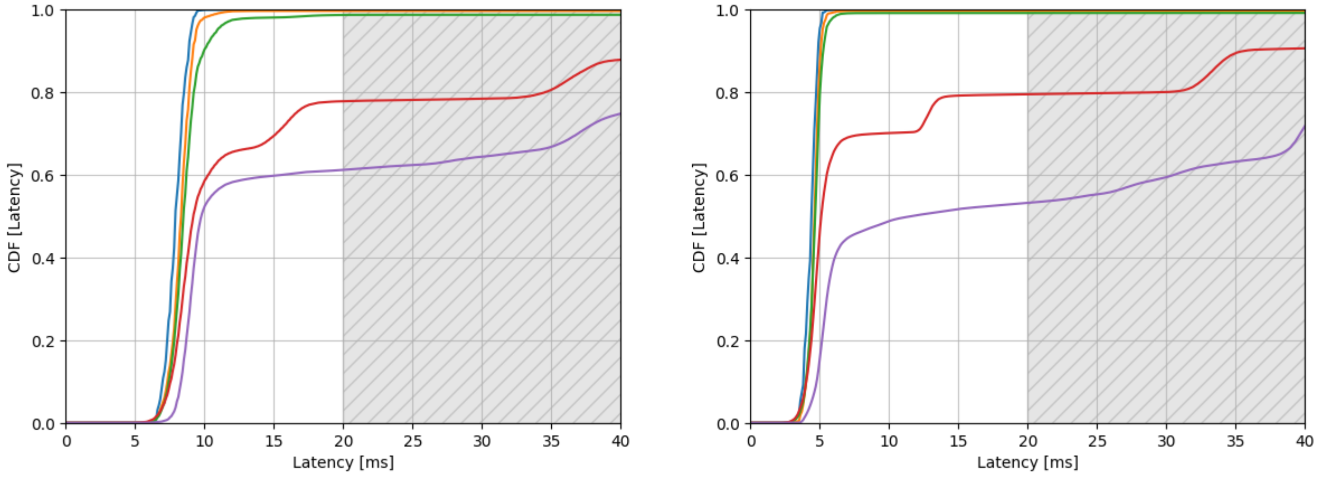
Each UE generates User Datagram Protocol (UDP) traffic with a payload size of 280 Bytes and a maximum transmission unit size of 1500 Bytes. A MEC server is deployed after the packet session anchor of the core network and represents the destination of the packets generated by the various SMIs. In this way, we emulate the application that performs audio mixing and error concealment at the edge of the radio access network. τ_{proc} is the sum of the time needed for synchronizing the audio streams and the time needed to perform audio mixing and error concealment. Since audio mixing and error concealment are comparatively low-cost operations, we focus on the time spent on synchronizing the audio streams. To accurately simulate the non-trivial effects of the synchronization procedure on the overall performance, we wrote an ad hoc ns-3 application that implements the following logic:

- Each inbound audio packet waits for the arrival of all other packets that originated in the same time slot from a member of the same NMP session.
- If all packets from the same time slot and NMP session are available, they will be mixed into one packet and sent to all the members of the NMP session. If the first packet to arrive from a specific time slot keeps waiting for more than $\theta_{MEC} = 10$ milliseconds, the server mixes the available packets from that time slot (regardless of whether all of them have arrived or not) and sends the resulting mixed audio data packet to all members of the NMP session.

We also assume that the MEC server is located in close proximity to the RAN infrastructure, thus $\tau_{\mathrm{transport}} \ll \mathcal{D}$. Moreover, both the audio processing at the MEC platform and the packet transport are assumed error-free, that is, $p_{\mathrm{succ,comp}} = 1$ and $p_{\mathrm{succ,transport}} = 1$. To increase realism, interfering traffic was added to both scenarios in the form of additional UEs supported by additional base stations. These base stations surround the ones supporting



(a) 4G (b) 5G, numerology $\mu = 0$



(c) 5G, numerology $\mu = 1$ (d) 5G, numerology $\mu = 2$

FIGURE 4. **Scenario 1: packet delivery latency for 4G and 5G cellular architectures. (a) 4G; (b) 5G numerology 0, (c) 5G numerology 1, and (d) 5G numerology 2. The larger bandwidth and tighter transmission scheduling of the 5G architecture enable faster packet delivery even as the number of users increases. Key: see panel (a).**

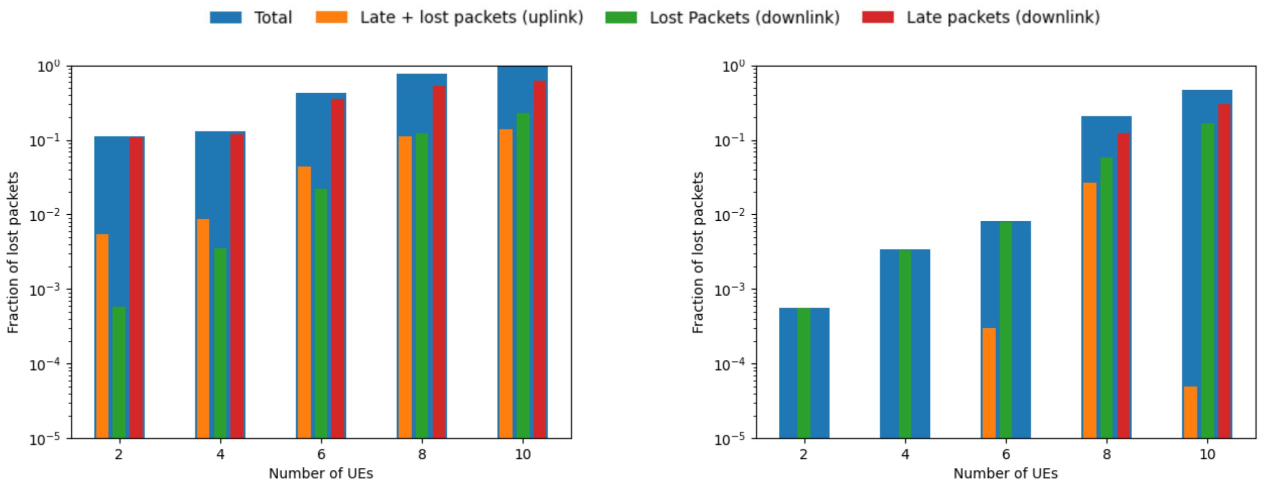
regular NMP traffic, as shown in panels (a) and (b) of Fig. 3. Each additional base station supports 5 UEs, and each UE sends a 25 bytes packet every millisecond. The complete list of simulation parameters is provided in Table 1. Because our study is the first to propose a large-scale evaluation of different IoMusT simulation scenarios, in the following we compare different configurations of the 5G radio access parameters. Notably, the only experimental performance evaluation of a multi-user networked music performance deployment [34] focuses on a single-cell co-located deployment, and does not include an evaluation of multi-cell scenarios or EN-DC configurations.

B. SIMULATION RESULTS

1) SCENARIO 1

We start with Fig. 4, which presents the cumulative distribution function (CDF) of the end-to-end packet delivery latency (i.e., the probability that the latency is less than the value indicated in the abscissa) for 4G and 5G cellular systems. Such latency value conveys the time required for a musical thing to send data to the MEC host and receive back a mixed audio stream. For 5G, we consider the three different numerologies $\mu = \{0, 1, 2\}$ introduced in the NR standard [36]. In these graphs, the best-case CDFs are as steep as possible, and transition from 0 to 1 as close as possible to the left side of each panel. As per the discussion in Section III-C, desirable delay values are below 20 ms. For representation purposes, here we extend the scope of the abscissa up to 40 ms, and mark the region exceeding 20 ms using a grey background.

Fig 4a shows that, in 4G systems, latency exceeds 10 ms irrespective of the number of users in the network. This delay includes all components along the communication and processing chain, including the radio links and the MEC



(a) 4G

(b) 5G, numerology $\mu = 2$

FIGURE 5. Scenario 1: packet reliability comparison among 4G and 5G with numerology 2.

TABLE 1. Simulation parameters.

	Parameter	Value
Traffic model	UDP payload size UDP MTU $\tau_{\text{audio,upstream}}$	280 B 1500 B 0.67 ms
Interfering traffic model	UDP payload size UDP MTU $\tau_{\text{audio,upstream}}$	25 B 1500 B 1 ms
Delay	τ_{proc}	Variable, max 10 ms
Channel model	FDD uplink frequency FDD downlink frequency Uplink bandwidth Downlink bandwidth Channel type MAC scheduler eNB/gNB TX power UE TX power eNB/gNB noise figure UE noise figure 4G error model 5G error model	1930 MHz 2120 MHz 20 MHz 20 MHz 3GPP TR 38.901 (UMa) Proportional fair 30 dBm 10 dBm 5 dB 9 dB LenaErrorModel NrEesmIrT1
Deployment (All)	h_{BS} h_{UE} L θ_{MEC}	25 m 1.5 m 150 m 10 ms
Deployment (Scenario 1)	— NMP Traffic — N_{BS} N_{UE} — Interfering Traffic — N_{BS} N_{UE}	 1 {2, 4, …, 10} 6 5 per BS
Deployment (Scenario 2)	— NMP Traffic — N_{BS} N_{UE} — Interfering Traffic — N_{BS} N_{UE}	 3 {2, 4, …, 20} 9 5 per BS

processing delay. In the simplest case with only two IoMusT devices, the maximum delay is already above the 20 ms threshold. As the number of users increases, the average and maximum delays grow, and the statistical dispersion of the delay also increases. For example, with 6 devices, 85% of the delivery delays exceed 15 ms, and the maximum delay is well above 20 ms, cf. Eq. (4). As the number of devices increases further, the percentage of packets lost or more than 20 ms late increases dramatically. With 8 UEs, almost 80% of packets are late or lost, while nearly no packets arrived before 20 ms with 10 UEs.

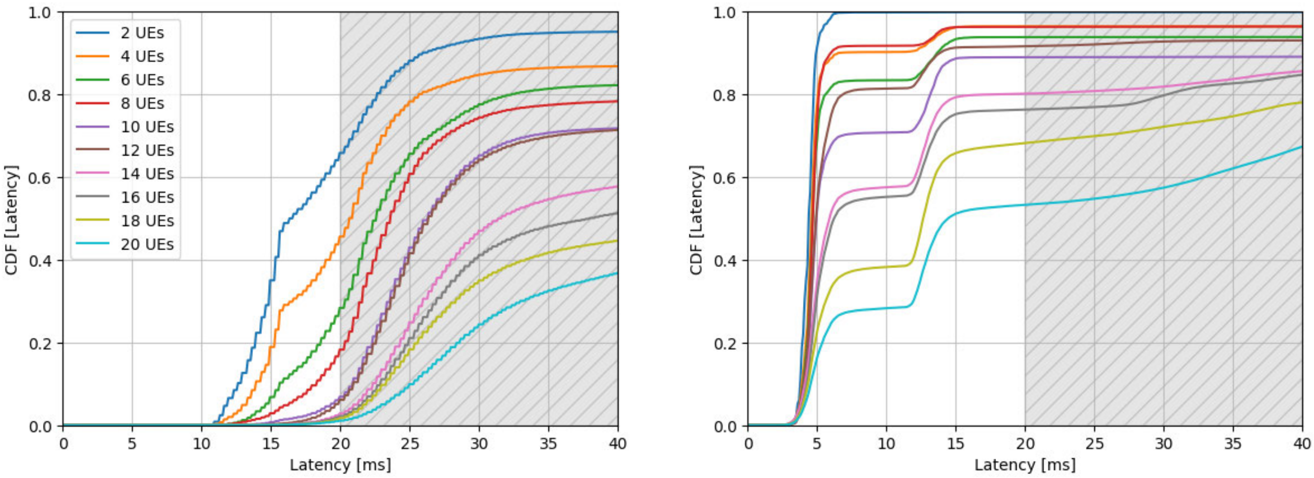
Conversely, the 5G architecture is more effective at delivering packets. The baseline 5G configuration with numerology $\mu = 0$ (Fig. 4b), while improving over 4G, is still not sufficiently robust to support more than 4 users. By comparing panels (b) to (d) in Fig. 4, we observe that increasing the numerology value from $\mu = 0$ to $\mu = 2$ improves the statistics of the latency. With numerology 0, the CDF shows a dispersion between 10 and 20 ms, showing degrading performance with the increase of the number of users. With 6 IoMusT devices, about 5% of the delay values exceed 20 ms. The performance improves significantly with numerologies 1 and 2, which reduce both the statistical dispersion of the latency (as seen from the steeper CDF curves) as well as both the minimum and maximum latency values. With numerology 2, less than 1% of the packets exceed 20 ms of delay for 6 users.

Regarding the reliability performance, Fig. 5 shows the fraction of lost to transmitted packets using the 4G (5a) and 5G (5b) cellular architectures and shows that a 4G system cannot satisfy the reliability constraint in Eq. (5) for any number of users while 5G manages to support up to at least 6 users per cell at $\mu = 2$, before exceeding downlink losses and late packets make networking issues perceivable.

2) SCENARIO 2

We ran a second batch of simulations to evaluate the performance in Scenario 2, with $N_{\text{BS}} = 3$, and up to 20 UEs.

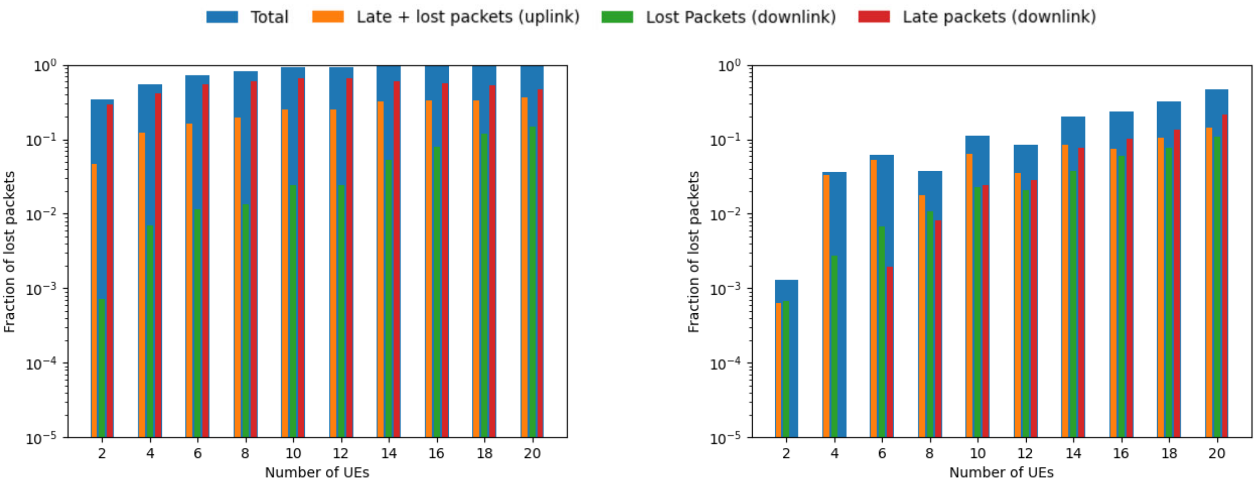
The latency performance results are provided in Fig. 6 for a 4G system and 5G system with $\mu = 2$, as the latter



(a) 4G

(b) 5G, numerology $\mu = 2$

FIGURE 6. Scenario 2: latency comparison between 4G and 5G with $\mu = 2$. Key: see panel (a).



(a) 4G

(b) 5G Num-2

FIGURE 7. Scenario 2: packet reliability comparison among 4G and 5G with numerology 2. Key: see panel (b).

configuration yields the best performance. Because users are now distributed across different cells, we tested the system for up to 20 UEs. While the overall trend of the curves for both 4G and 5G confirms the insights from scenario 1, we also observe a few differences. First, interference limits the performance of the whole system more heavily than in scenario 1. Even though users are distributed across different cells, neighboring cells interfere with one another, leading to an overall worse performance than in scenario 1 for the same number of users. Such interference comes not only from the outer tier of interfering cells, but also from the inner cells where the IoMusT users are located.

Second, as a consequence of interference and packet losses, timeouts occur more often at the the edge server. In this case, the server waits for the maximum edge delay budget, $\theta_{MEC} = 10$ ms. This is observed, e.g., from the "plateau" seen in all curves of Fig. 6b from about 5 ms to about 12.5 ms. As expected, the number of packets received within 20 ms decreases for increasing number of UEs, from > 99% with 2 UEs down to ≈90% with up to 10 UEs. For these numbers, the CDFs in Fig. 6b show that ≈80% of all packets are received within 5 ms. The remaining ones are either lost or delayed beyond the maximum delay of 20 ms. Still, 5G offers a much better performance than 4G, where no more than 65% of the packets are received within 20 ms, even with 2 UEs.

Similar considerations regarding the packet reliability can be derived from the results in Fig. 7, where the 4G network cannot satisfy the reliability target for any number of UEs. Conversely, 5G achieves the reliability target for 2 UEs and exceeds it mildly for 4 to 8 UEs. In all cases, errors or late packets in the uplink portions constitute the most significant cause of packet losses.

V. PERFORMANCE EVALUATION IN A MIXED 4G/5G SCENARIO USING EN-DC

We now proceed to evaluate a different scenario including the co-located deployment of 4G and 5G cells. This scenario is representative of the continuous transition phase between a fully 4G and a fully 5G-SA network. The standard-compliant configuration that enables the co-existence of 4G and 5G is named EN-DC, and its connect-compute architecture is illustrated in Fig. 8. Here, the 4G evolved NodeBs (eNBs) and the gNB operate at two different frequencies, whereby the former implements a macro cell operating at frequency f_{macro} and the latter implements a small cell operating at f_{micro}. Moreover, both the eNBs and the gNBs rely on a 4G mobile core network, i.e., an Evolved Packet Core (EPC). In particular, each eNB and gNB may play the role of *master node* and *secondary node*, respectively. As a consequence, the eNB is connected to the EPC via the S1 interface. Moreover, the gNB is not directly connected to the EPC: rather, both user plane and data plane traffic are forwarded to the master node via the X2 interface [76, Section 4]. As for the user plane, the UE can leverage different radio bearer types to exploit 4G or 5G radio resources in order to convey/receive multiple streams of information to/from the mobile network [76, Section 9].

For the mobile core network topology, we consider a distributed S/PGW deployment [77]. Here, the Serving Gateway (SGW) and packet data network gateway (PGW) virtual network functions (VNFs) (which deal with data traffic aggregation and QoS enforcement before routing the traffic to mobile network operator's services or the Internet) are deployed at an edge site, as close as possible to the serving RAN infrastructure. This aligns the EN-DC deployment with the single-technology 4G or 5G scenarios considered in Section IV. In this context, the MEC host connects to the (local) PGW over the SGi interface.

Because ns-3 and 5G-LENA do not yet support the simulation of EN-DC scenarios, we resort to the Simu5G package [55], an end-to-end simulator for 4G and 5G networks based on the popular OMNeT++ framework.

A. EN-DC SIMULATION SCENARIOS AND SETUP

The ns-3/5G-LENA and OMNeT++/Simu5G simulators were born and evolved as very different packages, and some differences between their outputs are to be expected. Therefore, as a first check before conducting EN-DC simulations, we configure Simu5G to replicate the Scenario 1 as introduced in the ns-3/5G-LENA-based simulations of Section IV (including, e.g., the placement of the cells and the use of a realistic packet mixing application). This provides common grounds for an initial comparison between single-technology simulation results.

However, for the subsequent EN-DC scenario, we need to set some extra parameters. Under the assumption that SMIs are mostly concentrated in a given area and that 5G coverage serves as a high-performance connectivity spot, we configure a secondary gNBs as a micro BS, radiating a power of 30 dBm. To avoid cross-tier interference, we let the gNBs use a different carrier frequency than the macro eNBs. The distance between the gNBs and their corresponding eNBs is 100 m. In the EN-DC configuration, SMIs and UEs attach to the eNB or gNB that yields the strongest received power.

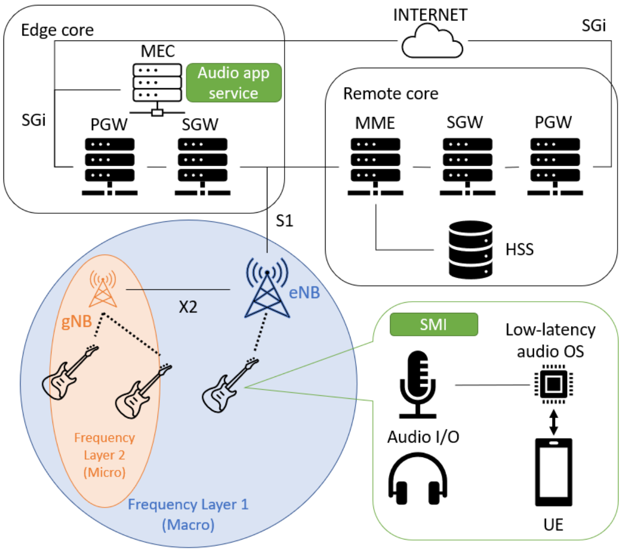


FIGURE 8. Representation of the connect-compute architecture leveraging EN-DC and MEC to implement LL-NMP use cases under a hybrid 4G/5G deployment.

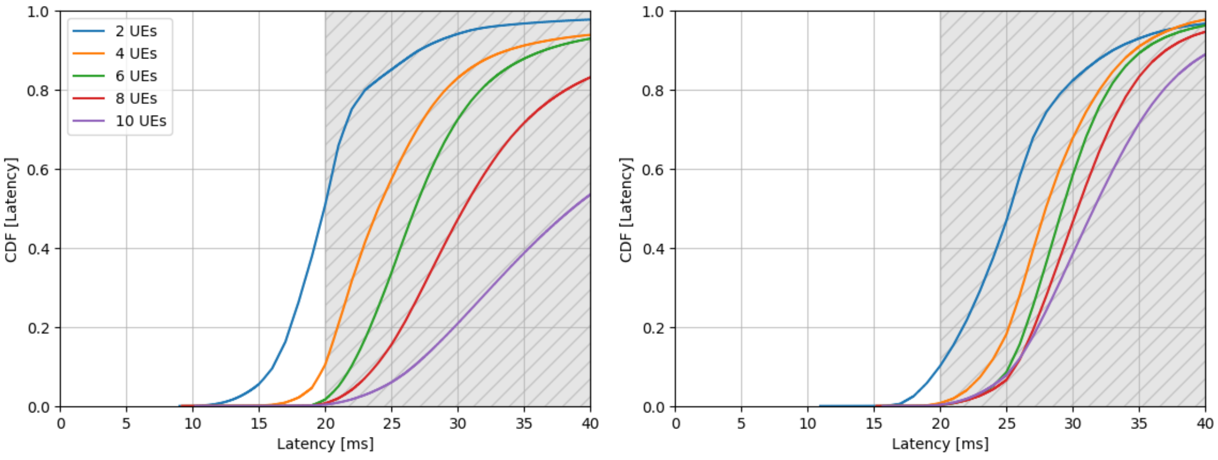
Moreover, we assume that the elements of the EPC or 5G Core (5GC) are implemented as VNFs, and that the latency required to traverse one VNF is 100 μ s [78]. In the 4G and EN-DC scenarios, we assume that the PGW and SGW are co-located with the eNB and that the MEC host is deployed just after them, hence the one-way latency introduced by these VNFs is 200 μs. In the standalone 5G scenario, instead, the MEC host is deployed after one user plane function (UPF) element of the 5GC and the one-way latency is 10 μs. For the X2 interface in the EN-DC deployment, we assume a 1 ms one-way latency [79].

For each scenario, we run 60 independent replicas of a 20-second simulation.

B. SIMU5G RESULTS – SINGLE-TECHNOLOGY SCENARIO

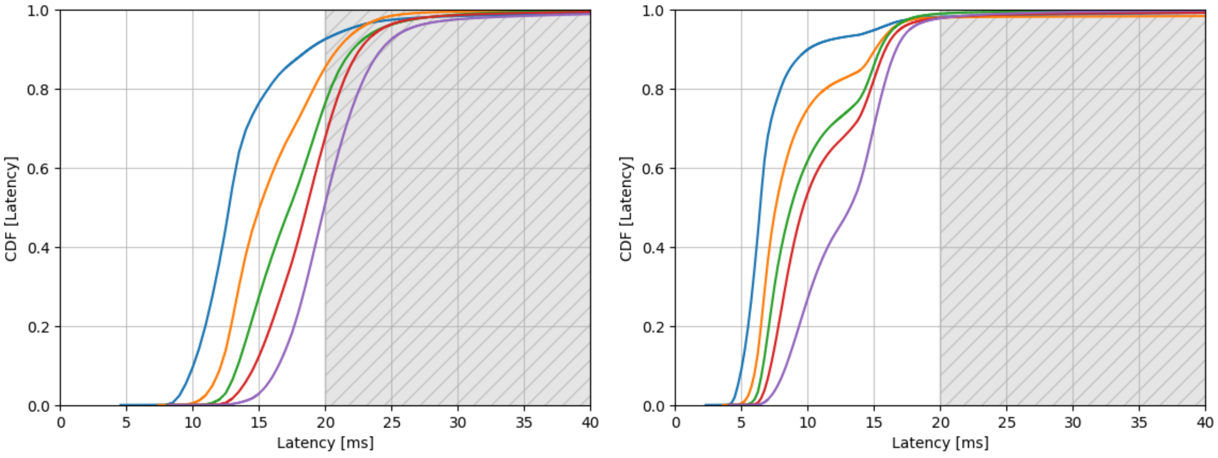
We start our evaluation with Fig. 9, which shows the CDF of the end-to-end packet delivery latency. In this first scenario, all IoMusT devices are connected to the same air interface, hence we align to Fig. 4 and present results separately for 4G and for the three numerologies $\mu = 0, 1, 2$ of 5G NR. We recall that a total round-trip time of 20 ms is the maximum the system can afford to ensure a fully smooth musical performance.

The first observation is that Simu5G simulations offer the same trends as 5G-LENA simulations, with predictably worse performance for increasing number of users. This is observed from the CDF curves moving towards the right of each graph



(a) 4G

(b) 5G, $\mu = 0$



(c) 5G, $\mu = 1$

(d) 5G, $\mu = 2$

FIGURE 9. **Scenario 1: packet delivery latency for 4G and 5G cellular architectures using OMNeT++/Simu5G. (a) 4G; (b) 5G numerology 0, (c) 5G numerology 1, and (d) 5G numerology 2. The trends confirm that end-to-end latency increases for increasing number of users and that higher 5G numerologies successfully reduce the latency, thereby providing support for de-facto real time interactions for a larger number of users.**

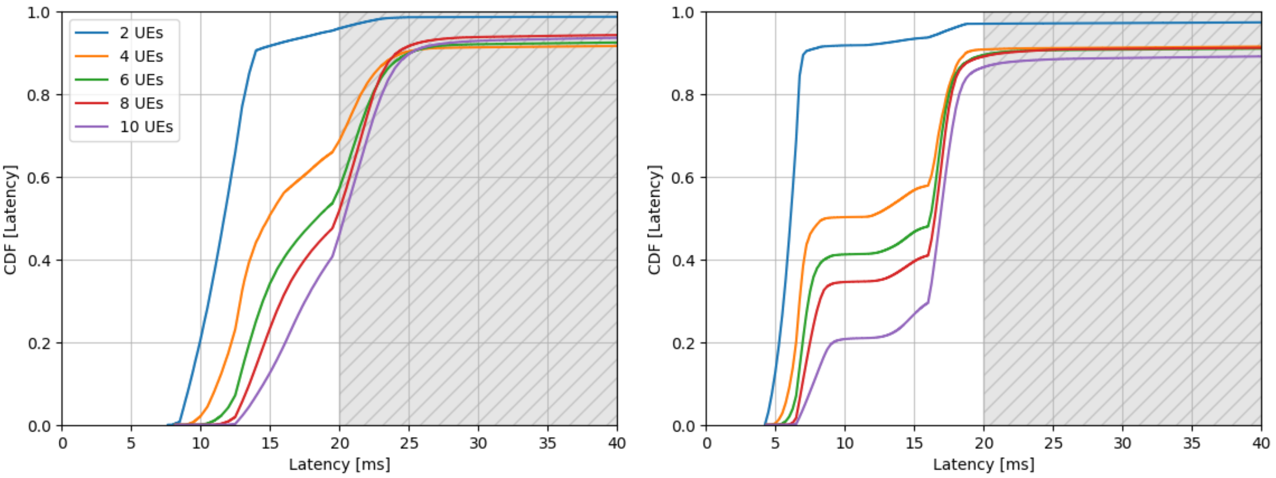
as the number of IoMusT users increases. The proportion of delays not exceeding 40 ms (i.e., the value of each curve at the rightmost part of each graph) also decreases for increasing IoMusT users.

At all numerologies, 5G offers better performance to IoMusT data flows, leading to generally lower delays, except for a slightly higher minimum delay for the case of 2 users under 5G for $\mu = 0$ with respect to 4G. Additionally 5G also leads to steeper CDF curves, which imply a generally more predictable delay. These results are also in line with the outcomes of 5G-LENA simulations.

We also observe two differences between 5G-LENA and Simu5G simulations: minimum delay values are higher in Simu5G than in 5G-LENA, whereas the proportion of packets whose delay exceeds 40 ms is larger in 5G-LENA The former is due to implementation-specific timers related to the exchange of standard-compliant messages, such as the so-called buffer-status reports that UEs send to advertise the amount of data available in their uplink buffers. The reason behind the latter, instead is a larger amount of memory dedicated to UE buffers in 5G-LENA, which causes longer queues even before a packet can finally be uplinked to the 4G or 5G base station. In any event, Simu5G results confirm that higher numerologies in 5G networks offer acceptable delay performances thanks to more efficient modulation schemes, error control, and network scheduling management. Therefore, when available, 5G coverage enables de-facto real-time IoMusT interactions among 6 to 10 users. We remark that we are not testing a single cell in isolation, but a realistic cell deployment scenario, with a tier of six cells surrounding (and interfering with) the cell under observation.

C. RESULTS – EN-DC DEPLOYMENT

Having established a substantial correspondence between the trends observed in 5G-LENA and Simu5G, we now turn to



(a) $\mu = 1$ for the 5G cell

(b) $\mu = 2$ for the 5G cell

FIGURE 10. EN-DC scenario: packet delivery latency using OMNeT++/Simu5G. (a) 5G numerology 1, (b) 5G numerology 2. 5G yields a considerable advantage over a purely 4G cell, both by decreasing the traffic of the 4G cell and by providing higher performance to 5G users.

the EN-DC scenario where a spot 5G cell co-exists within the coverage area of an incumbent 4G cell. Fig. 10 shows the CDF of the end-to-end latency for different numerology values and different numbers of users. Here, we focus on numerologies $\mu = 1, 2$ for the 5G cell, as our results suggest that $\mu = 0$ does not offer any substantial advantage in EN-DC scenarios.

Conversely, for $\mu = 1, 2$ EN-DC greatly helps improve the otherwise poor performance of a purely 4G deployment, because offloading traffic to the gNB increases the availability of 4G resources for the remaining users. In more detail, EN-DC complies with the 20 ms delay constraint at least 50% of the times for $\mu = 1$ (Fig. 10a) and at least 85% of the times for $\mu = 2$, depending on the number of IoMusT users. In all cases, the distribution of the latency exhibits a long tail, and a region of lower slope, typically around 10 ms. The latter is emphasized for $\mu = 2$. The long tail is due to 4G users, which do not experience a performance as critically low as in Fig. 9a, but still access a comparatively slower network that manages traffic less efficiently. Instead, the lower-slope region of the CDFs in Fig. 10b is due to the same effect observed in Fig. 6b. In the specific case of EN-DC, if the first packet of a group to be mixed is served by a 5G user, this packet is likely to arrive early on, and trigger the timeout timer at the mixing server in the MEC. If other packets in the same group come from 4G users, there is a higher chance that these packets will arrive late and trigger a mixing timeout. This effect is less likely if only 2 IoMusT users are present in the cell. Notably, 5G offers a comparatively lower advantage in the presence of 10 users, as there exists a higher chance that some of them will be covered only by a 4G cell.

Finally, Fig. 11 shows the fraction of late packets for different numerologies applied to the 5G cell. This figure confirms two typical trends: that increasing the number of users increasing the chance that packets arrive late for mixing or back to the user; and that increasing the numerology of the 5G cell yields generally better performance, mostly as a consequence of better service delivered to 5G users with respect to 4G ones.

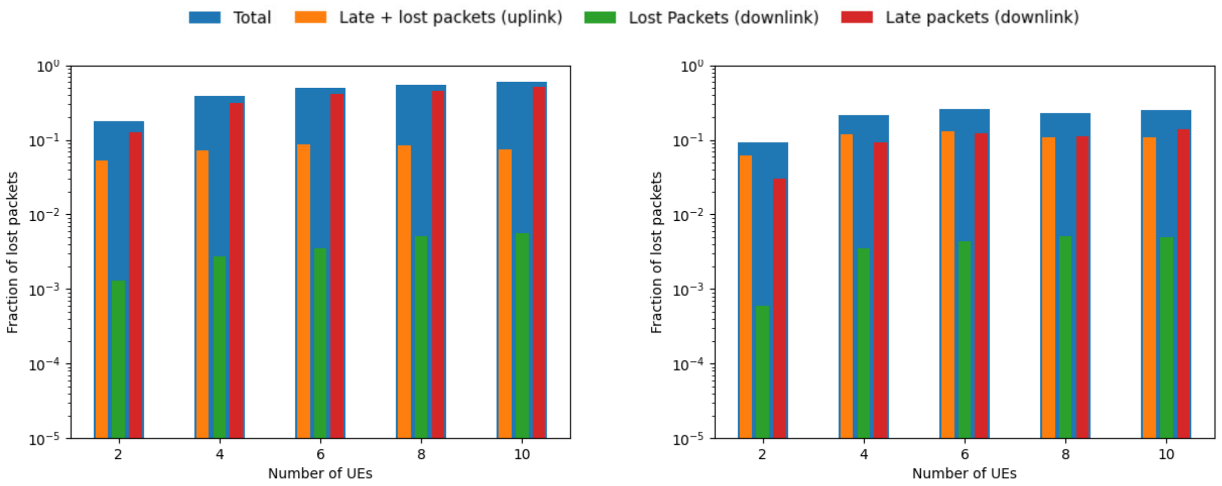
From the above results, we can conclude that EN-DC scenarios offering spot 5G coverage to IoMusT users constitute a good intermediate solution for 5G service deployed in a narrower geographical area than currently widespread 5G networks. While users limited to 4G service are still likely to experience unsatisfactory performance even with EN-DC, the load on the 4G cell is reduced and the delay experienced by networked audio transfers is significantly better than in a loaded 4G network. These characteristics make EN-DC a valid alternative to manage IoMusT traffic in the transition between 4G and a fully 5G coverage.

VI. USER STUDY

We complete our evaluation by conducting a user study. Our objective is to validate our simulation results from the quality of experience standpoint. Using the reliability metrics resulting from the 4G and 5G simulations, we created a set of audio files simulating networked music performance sessions, including packet delays, losses, and the effects of jitter buffer deadlines, possibly leading to discarding packets that were received correctly, but too late. Then, we asked musicians to play over such audio files and to assess their experience.

A. APPARATUS

The apparatus aimed at replicating the conditions of a real NMP session. The experiments were carried out in an acoustically insulated laboratory at the University of Trento, which was equipped like a recording studio. The apparatus comprised a laptop delivering the sound stimuli (via the webMUSHRA listening test framework [80]), which was connected to a soundcard (RME Fireface UFX II).



(a) $\mu = 1$ for the 5G cell (b) $\mu = 2$ for the 5G cell

FIGURE 11. EN-DC scenario: fraction of late packets, considering different numerologies for the 5G cell and using OMNeT++/Simu5G. A higher numerology and a lower number of users lead to better performance, but the presence of an incumbent 4G deployment still leads to several late packets.

Participants used their own electric guitar as well as equipment for sound processing. The sounds resulting from the mix of the electric guitarists' and simulated connected musicians were delivered via headphones (Beyerdynamic DT-770 pro).

B. STIMULI

The following 3 kinds of NMP sessions were simulated, each leveraging the data from the 4G and 5G results reported above:

1) 1 musician playing an arpeggio-based accompaniment on the acoustic guitar;
2) 3 musicians playing a blues piece on piano, drums, and electric bass;
3) 3 musicians playing a funk piece on keyboards, drums, and electric bass.

Each simulated NMP session was repeated twice (for a total of 12 trials) and the order of presentation was randomized for each participant.

The realizations of lost packets from the NS-3 simulations were used to directly simulate the packet loss on these files. To account for the use of packet loss concealment methods, we applied the BurgPLC algorithm, which currently represents the state of the art. We used the PLC Testbench reported in [81] configured with BurgPLC.

C. PARTICIPANTS

Ten electric guitar players (all males, mean age: 30.7 years, standard deviation: ~5.0 years) took part in the experiment. They were all expert musicians, with 12 years of musical expertise on average). Each musician took on average 30 minutes to complete the experiment. The electric guitar was selected because it was an instrument not involved in the audio files simulating the network-connected musicians.

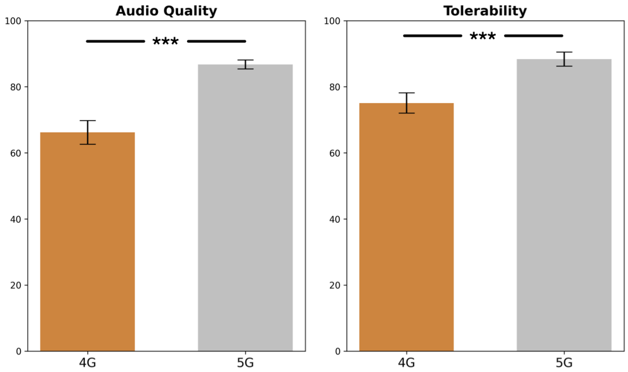


FIGURE 12. Mean and standard error of participants' evaluations in each experimental condition across the audio quality and tolerability metrics.

D. PROCEDURE

The procedure, approved by the local ethics committee, was in accordance with the relevant ethical standards of the Declaration of Helsinki (1964, revised in Fortaleza in 2013), and compliant with the EU GDPR. Participants were given an information sheet and asked to sign a consent form.

Before starting the experiments, participants were instructed to adjust the volume of the sound related to the simulated musicians and that of the electric guitar. Subsequently, musicians were asked to interact with the webMUSHRA applications to start the simulated NMP sessions and to assess their experience across two metrics: audio quality and tolerability (i.e., how they deemed it tolerable to play along with the simulated musicians). They were asked to play over the audio files as soon as the playback started, and to provide their assessment immediately after the playback finished. Specifically, the rankings were conducted over two continuous faders ranging between 0 and 100.

E. RESULTS

An ANOVA was performed on two different linear mixed effect models, one for each metric utilized. Specifically, each model had the metric (audio quality and tolerability) and condition (4G and 5G) as fixed factors, and the playing subject as a random factor. For each model, the assumption on the normality of the residuals was verified.

Results are illustrated in Fig. 12. A significant main effect for factor condition ($p < 0.001$) was found for both audio quality and tolerability, showing that 5G led to significantly higher ratings than 4G.

VII. CONCLUSION

In this paper, we characterized a 5G-based IoMusT framework, showing that the technologies belonging to the mobile network ecosystem can potentially support new use cases such as LL-NMP, allowing ensemble playing even when the musicians are not co-located. We identified the KPIs to be considered for the evaluation of LL-NMP implementations and proposed a general communication architecture. We evaluated the performance of 4G and 5G standards in satisfying the identified metrics via a system-level simulation campaign with ns-3. The simulation results advocate 5G systems as a crucial enabler of the envisioned IoMusT paradigm, overcoming the limitations of the 4G technology. Notably, the latency and reliability results achieved with the performed 5G simulations are in line with the recent results of the performance analysis of the 5G-based IoMusT deployments reported in [33] and [34].

Considering that 4G and 5G technologies will co-exist for a significant amount of time before 5G becomes truly ubiquitous, we demonstrated that a careful setup of a hybrid 4G/5G network can manage traffic much better with respect to a pure 4G system, thanks to the EN-DC architecture. To evaluate the latter, we resorted to the Simu5G simulator, based on the popular OMNeT++ framework. We discussed that relieving 4G networks by routing traffic through 5G gNBs does help reduce the latency incurred by 4G-served IoMusT devices. However, as the latency of 4G traffic remains significant with respect to quasi real-timeness constraints, some mixing timeouts may still occur.

We validated our findings via a user study involving 10 musicians. Each person was asked to play on top of musical traces created by reproducing the sequence of delivery delay values and transmission errors that would occur over a 4G and 5G systems, along with state-of-the-art packet loss concealment methods. The result of this study validates our conclusions about the need for a 5G framework to fully support the IoMusT.

As a further contribution, using two prominent community-maintained frameworks for our simulation campaigns has unveiled that their results are compatible in that they show the same performance trends. Yet, some differences necessarily emerge in each simulator's implementation choices that sometimes lead to different communication timings.

Future work on this topic includes investigating the benefits of network slicing and private networks for widespread adoption of a 5G-enabled IoMusT.

REFERENCES

[1] 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.

[2] L. Turchet, "Smart musical instruments: Vision, design principles, and future directions," *IEEE Access*, vol. 7, pp. 8944–8963, 2019.

[3] B. Matuszewski, "A web-based framework for distributed music system research and creation," *J. Audio Eng. Soc.*, vol. 68, no. 10, pp. 717–726, Dec. 2020.

[4] K. Nishida, A. Yuguchi, J. Kazuhiro, P. Modler, and M. Noisternig, "Border: A live performance based on web AR and a gesture-controlled virtual instrument," in *Proc. Int. Conf. New Interface Musical Expression*, 2019, pp. 43–46.

[5] L. Turchet, T. West, and M. M. Wanderley, "Touching the audience: Musical haptic wearables for augmented and participatory live music performances," *Pers. Ubiquitous Comput.*, vol. 25, no. 4, pp. 749–769, Aug. 2021.

[6] F. Font, "SOURCE: A freesound community music sampler," in *Proc. Audio Mostly*, Sep. 2021, pp. 182–187.

[7] D. Keller, C. Gomes, and L. Aliel, "The Handy Metaphor: Bimanual, touchless interaction for the Internet of Musical Things," in *Proc. 8th Workshop Ubiquitous Music*, 2018, pp. 180–188.

[8] L. Turchet, J. Pauwels, C. Fischione, and G. Fazekas, "Cloud-smart musical instrument interactions: Querying a large music collection with a smart guitar," *ACM Trans. Internet Things*, vol. 1, no. 3, pp. 1–29, Aug. 2020.

[9] A. T. Marasco, E. Berdahl, and J. T. Allison, "Bendit_I/O: A system for networked performance of circuit-bent devices," in *Proc. NIME*, 2019, pp. 331–334.

[10] S. Thorn, "Telematic wearable music: Remote ensembles and inclusive embodied education," in *Proc. Audio Mostly*, Sep. 2021, pp. 188–195.

[11] A. Fraietta, O. Bown, S. Ferguson, S. Gillespie, and L. Bray, "Rapid composition for networked devices: HappyBrackets," *Comput. Music J.*, vol. 43, nos. 2–3, pp. 89–108, Jun. 2020.

[12] A. Fraietta, O. Bown, and S. Ferguson, "Transparent communication within multiplicities," in *Proc. 27th Conf. Open Innov. Assoc. (FRUCT)*, Sep. 2020, pp. 61–72.

[13] S. Roy, D. Sarkar, and D. De, "DewMusic: Crowdsourcing-based Internet of Music Things in dew computing paradigm," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 2, pp. 2103–2119, Feb. 2021.

[14] O. Bown, S. Ferguson, A. D. P. Dos Santos, and K. Mikolajczyk, "Supporting creative practice in wireless distributed sound installations given technical constraints," *J. Audio Eng. Soc.*, vol. 69, no. 10, pp. 757–767, Oct. 2021.

[15] R. Vieira and F. Schiavoni, "Sunflower: An environment for standardized communication of IoMusT," in *Proc. Audio Mostly*, Sep. 2021, pp. 175–181.

[16] L. Turchet and F. Antoniazzi, "Semantic Web of Musical Things: Achieving interoperability in the Internet of Musical Things," *J. Web Semantics*, vol. 75, Jan. 2023, Art. no. 100758.

[17] P. Verma, A. Ilic 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 Innov. Assoc. (FRUCT)*, Sep. 2020, pp. 268–275.

[18] 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.

[19] R. Hupke, M. Nophut, S. Preihs, and J. Peissig, "Toward professional distributed performances: Effects of a global metronome on networked musical ensemble interactions," *J. Audio Eng. Soc.*, vol. 69, no. 10, pp. 720–736, Oct. 2021.

[20] R. Battello, L. Comanducci, F. Antonacci, G. Cospito, and A. Sarti, "Experimenting with adaptive metronomes in networked music performances!" *J. Audio Eng. Soc.*, vol. 69, no. 10, pp. 737–747, Oct. 2021.

[21] M. Bosi, A. Servetti, C. Chafe, and C. Rottondi, "Experiencing remote classical music performance over long distance: A JackTrip concert between two continents during the pandemic," *J. Audio Eng. Soc.*, vol. 69, no. 12, pp. 934–945, Dec. 2021.

[22] 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 *Proc. 152nd Audio Eng. Soc. Conv.* New York, NY, USA: Audio Engineering Society, 2022.

[23] D. Keller, F. Schiavoni, and V. Lazzarini, "Ubiquitous music: Perspectives and challenges," *J. New Music Res.*, vol. 48, no. 4, pp. 309–315, Aug. 2019.

[24] J. Timoney, A. Yaseen, and D. Mcevoy, "The potential role of Internet of Musical Things in therapeutic applications," in *Proc. 10th Workshop Ubiquitous Music*, 2020, pp. 61994–62017.

[25] J.-P. 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, Sep. 2010.

[26] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An overview on networked music performance technologies," *IEEE Access*, vol. 4, pp. 8823–8843, 2016.

[27] S. J. Barro, T. M. Fernández-Caramés, and C. J. Escudero, "Enabling collaborative musical activities through wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 8, no. 3, Mar. 2012, Art. no. 314078.

[28] J. P. M. Avila, C. Greenhalgh, A. Hazzard, S. Benford, and A. Chamberlain, "Encumbered interaction: A study of musicians preparing to perform," in *Proc. CHI Conf. Human Factors Comput. Syst.*, May 2019, pp. 1–13.

[29] X. Jiang, H. Shokri-Ghadikolaei, G. Fodor, E. Modiano, Z. Pang, M. Zorzi, and C. Fischione, "Low-latency networking: Where latency lurks and how to tame it," *Proc. IEEE*, vol. 107, no. 2, pp. 280–306, Feb. 2019.

[30] L. Gabrielli and S. Squartini, *Wireless Networked Music Performance*. Singapore: Springer, 2016, pp. 53–92.

[31] *Remote Music Performance by Ericsson and MIND Music Labs at Mobile World Congress 2019*. Accessed: Apr. 2019. [Online]. Available: https://elk.audio/remote-music-performance-by-ericsson-and-mind-music-labs-at-mwc19/

[32] 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. SMC*, Jun. 2020, pp. 407–412.

[33] J. Dürre, N. Werner, S. Hämäläinen, O. Lindfors, J. Koistinen, M. Saarenmaa, and R. Hupke, "In-depth latency and reliability analysis of a networked music performance over public 5G infrastructure," in *Proc. 153rd Audio Eng. Soc. Conv.* New York, NY, USA: Audio Engineering Society, 2022, pp. 1–10.

[34] 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. 2023.

[35] *System Architecture for the 5G System; Stage 2 (Release 16)*, document 23.501, Rev. 16.1.0, 3GPP TSG-SA, Jun. 2019.

[36] *NR; NR and NG-RAN Overall Description; Stage 2 (Release 15)*, document 38.300, Rev. 15.5.0, 3GPP TSG-RAN, Mar. 2019.

[37] N. Bhushan, T. Ji, O. Koymen, J. Smee, J. Soriaga, S. Subramanian, and Y. Wei, "5G air interface system design principles," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 6–8, Oct. 2017.

[38] J. Vihriälä, A. A. Zaidi, V. Venkatasubramanian, N. He, E. Tiirola, J. Medbo, E. Lähetkangas, K. Werner, K. Pajukoski, A. Cedergren, and R. Baldemair, "Numerology and frame structure for 5G radio access," in *Proc. IEEE 27th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2016, pp. 1–5.

[39] N. H. Mahmood, R. Abreu, R. Böhnke, M. Schubert, G. Berardinelli, and T. H. Jacobsen, "Uplink grant-free access solutions for URLLC services in 5G new radio," in *Proc. 16th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Oulu, Finland, Aug. 2019, pp. 607–612.

[40] M. Centenaro, D. Laselva, J. Steiner, K. Pedersen, and P. Mogensen, "System-level study of data duplication enhancements for 5G downlink URLLC," *IEEE Access*, vol. 8, pp. 565–578, 2020.

[41] A. Baratè, G. Haus, and L. A. Ludovico, "Advanced experience of music through 5G technologies," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 364, Jun. 2018, Art. no. 012021.

[42] A. Baratè, G. Haus, L. A. Ludovico, E. Pagani, and N. Scarabottolo, "5G technology and its applications to music education," in *Proc. MCCSIS*, 2019, pp. 65–72.

[43] N. Hassan, K. A. Yau, and C. Wu, "Edge computing in 5G: A review," *IEEE Access*, vol. 7, pp. 127276–127289, 2019.

[44] S. Kekki, "MEC in 5G networks," ETSI ISG MEC, Sophia Antipolis, France, White Paper 28, 2018.

[45] M. Centenaro, S. Tomasin, N. Benvenuto, and S. Yang, "Predictive voice-over-Internet protocol fallback over vehicular channels: Employing artificial intelligence at the edge of 5G networks," *IEEE Veh. Technol. Mag.*, vol. 15, no. 2, pp. 72–78, Jun. 2020.

[46] Q.-V. Pham, F. Fang, V. N. Ha, Md. J. Piran, M. Le, L. B. Le, W.-J. Hwang, and Z. Ding, "A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art," *IEEE Access*, vol. 8, pp. 116974–117017, 2020.

[47] P. Corcoran and S. K. Datta, "Mobile-edge computing and the Internet of Things for consumers: Extending cloud computing and services to the edge of the network," *IEEE Consum. Electron. Mag.*, vol. 5, no. 4, pp. 73–74, Oct. 2016.

[48] 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.

[49] Cisco Visual Networking Index. *Global Mobile Data Traffic Forecast Update*. Accessed: Jun. 2021. [Online]. Available: http://media.mediapost.com/uploads/CiscoForecast.pdf

[50] Ericsson. (Nov. 2022). *Ericsson Mobility Report*. Accessed: Jan. 2023. [Online]. Available: https://www.ericsson.com/4ae28d/assets/local/reports-papers/mobility-report/documents/2022/ericsson-mobility-report-november-2022.pdf

[51] 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.

[52] *Ns3 Network Simulator*. Accessed: Apr. 2021. [Online]. Available: https://www.nsnam.org

[53] *OMNeT++ Discrete Event Simulator*. Accessed: Jun. 2022. [Online]. Available: https://omnetpp.org/

[54] N. Patriciello, S. Lagen, B. Bojovic, and L. Giupponi, "An E2E simulator for 5G NR networks," *Simul. Model. Pract. Theory*, vol. 96, Nov. 2019, Art. no. 101933.

[55] G. Nardini, D. Sabella, G. Stea, P. Thakkar, and A. Virdis, "Simu5G—An OMNeT++ library for end-to-end performance evaluation of 5G networks," *IEEE Access*, vol. 8, pp. 181176–181191, 2020.

[56] M. Centenaro, P. Casari, and L. Turchet, "Towards a 5G communication architecture for the Internet of Musical Things," in *Proc. 27th Conf. Open Innov. Assoc. (FRUCT)*, Trento, Italy, Sep. 2020, pp. 38–45.

[57] *The LOLA Website*. Accessed: Mar. 2021. [Online]. Available: https://lola.conts.it/

[58] 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.

[59] E. Berdahl, S. Salazar, and M. Borins, "Embedded networking and hardware-accelerated graphics with satellite ccrma," in *Proc. Conf. New Interface Musical Expression*, 2013, pp. 325–330.

[60] E. Meneses, J. Wang, S. Freire, and M. M. Wanderley, "A comparison of open-source Linux frameworks for an augmented musical instrument implementation," in *Proc. Conf. New Interface Musical Expression*, 2019, pp. 222–227.

[61] A. McPherson and V. Zappi, "An environment for submillisecond-latency audio and sensor processing on BeagleBone black," in *Proc. 138th Audio Eng. Soc. Conv.* New York, NY, USA: Audio Engineering Society, 2015, pp. 1–7.

[62] 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, May 2021.

[63] *Study on Audio-visual Service Production; Stage 1 (Release 17)*, document 22.827, Rev. 17.1.0, 3GPP TSG-SA, Dec. 2019.

[64] *AR Framework Architecture*, document 003, Rev. 1.1.1, ETSI ISG ARF Group, Mar. 2020.

[65] L. Turchet, M. Benincaso, and C. Fischione, "Examples of use cases with smart instruments," in *Proc. Audio Mostly Conf.*, 2017, pp. 47:1–47:5.

[66] M. K. Müller, F. Ademaj, T. Dittrich, A. Fastenbauer, B. R. Elbal, A. Nabavi, L. Nagel, S. Schwarz, and M. Rupp, "Flexible multi-node simulation of cellular mobile communications: The Vienna 5G system level simulator," *EURASIP J. Wireless Commun. Netw.*, vol. 2018, no. 1, pp. 1–17, Dec. 2018.

[67] S. Martiradonna, A. Grassi, G. Piro, and G. Boggia, "5G-air-simulator: An open-source tool modeling the 5G air interface," *Comput. Netw.*, vol. 173, May 2020, Art. no. 107151.

[68] Z. Ali, S. Lagén, L. Giupponi, and R. Rouil, "3GPP NR V2X mode 2: Overview, models and system-level evaluation," *IEEE Access*, vol. 9, pp. 89554–89579, 2021.

[69] N. Patriciello, S. Lagén, B. Bojovic, and L. Giupponi, "NR-U and IEEE 802.11 technologies coexistence in unlicensed mmWave spectrum: Models and evaluation," *IEEE Access*, vol. 8, pp. 71254–71271, 2020.

[70] 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.

[71] *Service Requirements for the 5G System; Stage 1 (Release 17)*, document 22.261, Rev. 17.1.0, 3GPP TSG-SA, Dec. 2019.

[72] *Elk Audio OS*. Accessed: Mar. 2021. [Online]. Available: https://elk.audio/audio-os/

[73] R. Fanou, F. Valera, P. Francois, and A. Dhamdhere, "Reshaping the African Internet: From scattered islands to a connected continent," *Comput. Commun.*, vol. 113, pp. 25–42, Nov. 2017.

[74] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5G-enabled tactile Internet," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 460–473, Mar. 2016.

[75] N. Baldo, M. Miozzo, M. Requena-Esteso, and J. Nin-Guerrero, "An open source product-oriented LTE network simulator based on ns-3," in *Proc. 14th ACM Int. Conf. Modelling, Anal. Simulation Wireless Mobile Syst.*, Miami, FL, USA, Oct. 2011, pp. 293–298.

[76] *Evolved Universal Terrestrial Radio Access (E-UTRA) and NR; Multi-connectivity; Stage 2 (Release 16)*, document 37.340, Rev. 16.5.0, 3GPP TSG-RAN, Mar. 2021.

[77] F. Giust, "MEC deployments in 4G and evolution towards 5G," ETSI ISG MEC, Sophia Antipolis, France, White Paper 24, Feb. 2018.

[78] A. Virdis, G. Nardini, G. Stea, and D. Sabella, "End-to-end performance evaluation of MEC deployments in 5G scenarios," *J. Sensor Actuator Netw.*, vol. 9, no. 4, p. 57, Dec. 2020. [Online]. Available: https://www.mdpi.com/2224-2708/9/4/57

[79] M. Polese, M. Mezzavilla, and M. Zorzi, "Performance comparison of dual connectivity and hard handover for LTE-5G tight integration," in *Proc. EAI SIMUTOOLS*, Prague, Czech Republic, 2016, pp. 118–123.

[80] M. Schoeffler, S. Bartoschek, F.-R. Stöter, M. Roess, S. Westphal, B. Edler, and J. Herre, "WebMUSHRA—A comprehensive framework for web-based listening tests," *J. Open Res. Softw.*, vol. 6, no. 1, p. 8, Feb. 2018.

[81] L. Vignati, S. Dallona, and L. Turchet, "PLC testbench: A modular tool for the study and comparison of audio packet loss concealment algorithms," in *Proc. 4th Int. Symp. Internet Sounds*, Oct. 2023, pp. 1–8.

LUCA VIGNATI received the B.Sc. degree in computer and electronic engineering from the University of Pavia, Italy, and the M.Sc. degree in computer science and engineering from Politecnico di Milano, Italy. He is currently pursuing the Ph.D. degree with the Department of Information Engineering and Computer Science, University of Trento. His main research interests include the Internet of Musical Things, audio over 5G networks, packet loss concealment methods, and embedded systems specific for audio applications.

GIOVANNI NARDINI (Member, IEEE) received the M.Sc. degree in computer engineering and the Ph.D. degree in information engineering from the University of Pisa, in 2013 and 2017, respectively. He is currently an Assistant Professor with the University of Pisa. His research interests include resource allocation algorithms in 4G/5G/B5G networks, multi-access edge computing, simulation, and performance evaluation of computer networks. In these fields, he has coauthored six patents and more than 50 articles. He has been involved in EU-funded and industrial research projects. He is serving in the editorial board for the *Computer Communications* journal (Elsevier) as a member of the technical committee and the *Wireless Networks* journal (Springer) as an Editor. He has served as the TPC Chair for IEEE SmartSys, in 2023.

MARCO CENTENARO (Member, IEEE) received the B.S. degree in information engineering and the M.S. and Ph.D. degrees in telecommunication engineering from the University of Padova, Italy, in 2012, 2014, and 2018, respectively. He was on leave as a Visiting Researcher with the Nokia Bell Laboratories, Stuttgart, Germany, and Yokohama National University, Japan, in 2016 and 2017, respectively. Between 2018 and 2019, he was a Postdoctoral Research Fellow with the University of Padova, doing research on projects funded by Huawei Technologies and Aalborg University, Denmark, working in tight collaboration with the Nokia Bell Laboratories, Aalborg. Then, he was an Expert Researcher with Fondazione Bruno Kessler (FBK), Trento, Italy, and at the time of writing, he is currently a System Engineer with Athonet S.r.l., Bolzano Vicentino, Italy. His research interests include telecommunications standards and the Internet of Things enabling technologies. He was a recipient of the Ph.D. Award from Gruppo Telecomunicazioni e Tecnologie dell'Informazione (GTTI—Italian Association of Telecommunications and Information Technologies), in 2018, in recognition of the best Ph.D. theses defended at an Italian University in the areas of communications technologies, and the Best Paper Award at IEEE PIMRC 2018.

PAOLO CASARI (Senior Member, IEEE) received the Ph.D. degree in information engineering from the University of Padova, Italy, in 2008. He is currently an Associate Professor with the Department of Information Engineering and Computer Science, University of Trento, Italy. He was on leave with the Massachusetts Institute of Technology, in 2007, working on underwater communications and networks. He is currently a 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. Between 2015 and 2019, he was with the IMDEA Networks Institute, Madrid, Spain, where he led the Ubiquitous Wireless Networks Group. He received two best paper awards. He regularly collaborates with the organizing committee of several international conferences. He serves on the Editorial Board Member for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, served as an Associate Editor for IEEE TRANSACTIONS ON MOBILE COMPUTING (2018–2023), is currently a Co-Guest Editor of two special issues for *Computer Communications* (Elsevier) and *Computer Networks* (Elsevier) journals, and has been a Co-Guest Editor of a special issue of IEEE ACCESS on underwater acoustic communications and networking. His research interests include many aspects of networked communications, such as channel modeling, network protocol design, localization, simulation, and experimental evaluations.

SANDRA LAGÉN (Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees from Universitat Politècnica de Catalunya (UPC), Spain, in 2011, 2013, and 2016, respectively. Her dissertation was awarded by COIT the best national Ph.D. thesis on high-speed broadband mobile communications (2017) and received a Special Doctoral Award by UPC (2019). In 2017, she joined the CTTC Research Center, Spain, where she is currently a Senior Researcher and the Head of the Open Simulations (OpenSim) Research Unit. She has lead and participated in outstanding collaborations with mayor telecom industry players and coordinates the design and developing team of ns-3 5G-LENA. She has supervised multiple research secondments, ns-3 GSoC projects, and M.Sc./Ph.D. thesis. Her research interests include wireless communications, signal processing, and cross-layer system optimization. She was a recipient of IEEE WCNC 2018 and WNS3 2020 Best Paper Awards and coauthor of two non-provisional and three provisional patent applications. Since 2021, she has been a member of the Executive Board of ns-3 Consortium.

BILJANA BOJOVIC received the M.Sc. degree in electrical and computer engineering from the Faculty of Technical Sciences, Novi Sad, Serbia, in 2008, and the Ph.D. degree in networking engineering from the Polytechnic University of Catalonia, Barcelona, Spain, in 2022. In 2009, she joined the CTTC Research Center, Spain, where she is currently a Senior Researcher and a Promoter of the Open Simulations (OpenSim) Research Unit. She is also a Developer and a Maintainer of the LTE, NR, and NR-U modules of the ns-3 network simulator and the principal author of the LAA and LTE-U modules. She held LTE and NR module tutorials at the ns-3 workshops, in 2016 and 2022, and at CONFTELE, in 2021. In addition, she was a Mentor of ns-3 GSoC on several occasions. In 2020, she received ACM SIGCOMM Networking System Award. She worked on many research projects for industrial clients, such as Wi-Fi Alliance, SpiderCloud, Interdigital, U.S. Department of Defense, NIST, and Facebook. She is the coauthor of one patent application (US20200314906A1). Her research interests include XR traffic enhancements for 5G-advanced, MIMO simulation models for ns-3, and unlicensed/shared spectrum.

STEFANO ZAMBON received the master's (summa cum laude) and Ph.D. degrees in computer science from Verona University. He is currently a Co-Founder and a Chief Technical Officer of Elk. His experience combines academic-level expertise in signal processing with 11 years of industry experience working with large codebases, always for musical applications. He was the Core Developer of the physical-model algorithms behind Physis Piano, an advanced digital piano released, in 2013, by Viscount International S.r.l. He has a strong background in digital signal processing and acoustics, high-performance numerical computation, and embedded software development. He holds three patents and has authored three top-tier journal articles and nine publications in international conferences.

LUCA TURCHET (Senior Member, IEEE) received the master's degree (summa cum laude) in computer science from the University of Verona, in 2006, the degrees in classical guitar and in composition from the Music Conservatory of Verona, in 2007 and 2009, respectively, the Ph.D. degree in media technology from Aalborg University Copenhagen, in 2013, and the degree in electronic music from the Royal College of Music of Stockholm, in 2015. He is currently an Associate Professor with the Department of Information Engineering and Computer Science, University of 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 also a Co-Founder of the Music-Tech Company Elk. He is also 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 a Guest Editor for *IEEE Communications Magazine*, the *Personal and Ubiquitous Computing* journal, the *Journal of the Audio Engineering Society*, *Frontiers in VR*, and *Digital Creativity*.

• • •

Open Access funding provided by 'Università degli Studi di Trento' within the CRUI CARE Agreement