5G key technology enablers for emerging media content production services

Deliverable D2.1
Use cases, requirements and KPIs

Version v2.0

Date: 2022/01/21

Document properties:

Grant Number:	957102
Document Number:	D2.1
Document Title:	*Uses cases, requirements and KPIs*
Editor(s):	Cristina Avellán, Adrián Rodrigo, Irene Alepuz, Salvador García (UPV); Esther Sánchez, Rocío Ortiz (TID)
Authors:	David Gomez-Barquero, Cristina Avellán, Adrián Rodrigo, Irene Alepuz, Salvador García (UPV); Jaime Ruiz, Pablo Pérez, Ignacio Benito (NOK); Darko Ratkaj, Paola Sunna (EBU); Ian Wagdin, Simo Eley (BBC); Jan Dürre, Norbert Werner (SEN); Josue Igual, Emilio Abraham Garrido, Esther Sánchez, Rocio Ortiz (TID); Baruch Altman (LiveU); Francisco Morán, Daniel Berjón (UPM); Davide Milanesio, Federico Pandolfi, Davide Desirello (RAI); Manuel Fuentes, David Martín-Sacristán (5CMM); Antonio (ACC).
Reviewers	Roland Beutler (SWR); De Mi (UNIS); Juul Moen (NEP)
Contractual Date of Delivery:	2021/05/31
Dissemination level:	PU[1]
Status:	Final version
Version:	2.0
File Name:	5G-RECORDS_D2.1_v2.0

Revision History

Revision	**Date**	**Issued by**	**Description**
1.0	2021/05/31	TID	EC Final version
2.0	2022/01/21	TID	Update requested in the Mid Term Review

[1] CO = Confidential, only members of the consortium (including the Commission Services)

PU = Public

Abstract

5G-RECORDS is a European H2020 project that aims to explore the opportunities of new 5G technologies bringing to the professional audio-visual (AV) content production sector.

To this effect, the main goals are to offer improved performance in terms of bandwidth, latency, support for accurate timing, assurance of quality of service and scenario flexibility. Furthermore, it is expected that standardized 5G-based solutions can bring down production costs, reduce environmental impact and increase the operational efficiency and flexibility of production workflows in news gathering, remote production and coverage of live events.

5G-RECORDS has considered three use cases to embrace some of the most challenging scenarios in the professional content production: **live audio production**, a **multiple camera wireless studio** and **live immersive media production**.

To ensure the successful demonstration of these use cases, the project gets together a set of experienced partners whose expertise covers both 5G and content production value chains. Together, they have designed the main components of the architecture for each use case, also identifying the benefits for users, requirements, and technical enablers. Furthermore, the best way to understand each use case is through defined scenarios and workflows, for what is relevant to specify KPIs to measure the performance of the proposed solutions.

Keywords

5G, Audio-visual sector, Campus networks, Core technologies, Component development and integration, End-to-end infrastructures, KPIs, Non-public networks, Professional media content production, Use case requirements, Live events.

Disclaimer

This 5G-RECORDS D2.1 deliverable has been requested for revision after the Mid-Term Review that took place in November 2021. Once the monitoring process involving experts come to an end, i.e., once the deliverable is approved by the European Commission, the final approval decision will be communicated,

Executive Summary

5G-RECORDS aims to explore the opportunities which new 5G technology components bring for professional media content production. To do this, we base ourselves on the needs of the industry that currently has, and we look for the opportunities that new technologies offer us. Besides, it will take a business-to-business (B2B) perspective, but without losing sight of the end user benefits.

Therefore, the project will provide three end-to-end 5G infrastructure: core network (5GC), radio access network (RAN) and end devices. Besides, it aims to use of non-public networks (NPNs).

To be able to tackle all this, 5G-RECORDS has considered three use cases to embrace some of the most challenging scenarios in the framework of professional content production: live audio production, a multiple camera wireless studio and live immersive media production. To meet project's objectives, the first approach is to identify and design the main components of the 5G architecture for each use case. Furthermore, KPIs and requirements to measure the performance of the proposed solutions will be used to benchmark the existing state of the art as well as the 5G components developed within the project. Requirements coming from the content production industry as well as from other 5G PPP projects will be used as an input.

Use Case 1: Live audio production

The activity of a live audio production use case is carried out in places like theatres, production studios or music festivals, where one or several artists are performing live. Artists on stage are equipped with Programme Making and Special Events (PMSE) equipment, i.e., 5G wireless microphones and in-ear monitoring system (IEM). Thus, this use case (UC) aims to deploy a local 5G wireless high-quality ultra-low latency network dedicated for audio production, by designing 5G New Radio Reduced Capacity (NR-Redcap) audio devices prototype based on 3GPP Release 17. The demo will take place in EURECOM's infrastructure, a fully functional network demonstrator and testbed in **Sophia Antipolis, France,** integrating an open source 5G modem, a virtualised and disaggregated RAN, 5G-enabled microphones and IEM systems, spectrum sharing management technologies and a compact 5GC.

Use Case 2: Multiple camera wireless studio

Live or pre-recorded media content production usually requires deploying large amounts of equipment and crew on the event location or studio, all connected to the production facilities. The focus of this UC is to utilise 5G and develop other new technologies to create a wireless, holistic, interconnected content production system that reduces logistic effort and provides all the functionalities needed for broadcasters. This use case will deploy a studio with wireless 5G-enabled cameras wireless studio for operation under remote production scenarios in **Aachen, Germany**. It will also take into consideration operational requirements for working alongside wired studio equipment as well as explore cloud-based workflows.

Use Case 3: Live immersive media services

This use case will allow real-time immersive media experience of sport events. This means that each user can access a specific angle live, offering a unique Quality of Experience (QoE). It will also be possible to reproduce content both live and on demand. The content can then be distributed not only to people attending the event, but also to third parties. Nokia's E2E infrastructure based in **Segovia, Spain**, will incorporate every 5G component needed for the use case deployment. As integration comes to a mature state, end-to-end tests will be performed in terms of network performance and QoE.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

3GPP	3rd Generation Partnership Project
4G	4th Generation of mobile communications systems
5G	5th Generation of mobile communications systems
5G-PPP	5G Public-Private Partnership
AR	Augmented reality
AV	Audio-Visual
BG	Background
CAGR	Compound Annual Growth Rate
CEPT	European Conference of Postal and Telecommunications Administrations
COFDM	Coded Orthogonal Frequency Division Multiplexing
CS	Capture servers
CU	Centralized Unit
CU-CP	CU-Control Plane
CU-UP	CU-User Plane
DÉCOR	Dedicated Core Network
DNN	Data Network Name
DOF	Degrees of Freedom
DTT	Digital terrestrial television
DVB	Digital Video Broadcasting
E2E	End to end
EC	European Commission
ENG	Electronic News Gathering
ES	Edge server
FG	Foreground
FTTH	Fibre To The Home
FVV	Free Viewpoint Video
FWA	Fixed Wireless Access
HD	High Definition
HDM	Head-mounted displays
HDR	High Dynamic Range
HEVC	High Efficiency Video Coding
gNB	gNodeB
GoP	Group of Picture
ICT	Information and Communications Technologies
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IEM	In-Ear monitoring
IPTV	Internet Protocol Television
ITU	International Telecommunications Union

KPI	Key Performance Indicator
LTE	Long Term Evolution
MCR	Master Control Rooms
MEC	Multi-access Edge Computing
mMTC	Massive Machine-Type Communications
mmW	Millimetre Wave
MNO	Mobile Network Operator
MOCG	Media Operational Control Gateway
MPEG	Moving Picture Experts Group (formally, ISO/IEC JTC 1/SC 29/WG 11)
MTC	Machine Type Communications
MVD	Multiview plus depth
NMOS	Network Media Open Specifications
NPN	Non-Public Network
NR	New Radio
NRA	National Regulatory Authority
NSA	Non-standalone
NVF	Network Function Virtualisation
OB	Outside Broadcasting
O-RAN	Open Radio Access Network
PLMN	Public Land Mobile Network
PMSE	Programme Making and Special Events
PNI-NPN	Public Network Integrated Non-Public Network
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Networks
RF	Radio frequency
SA	Standalone
SAB	Services Ancillary to Broadcasting
SAP	Services Ancillary to Programme making
SDN	Software Defined Networking
SDL	Supplementary Downlink-only
SMPTE	Society of Motion Picture and TV Engineers
SNPN	Standalone Non-Public Networks
SUL	Supplementary Uplink-only
TR	Technical Report
TRL	Technology readiness levels
UC	Use Case
UHD	Ultra High Definition
UPF	User Plane Function
VIAPA	Video, Imaging and Audio for Professional Applications
VNF	Virtual Network Functions

VR	Virtual reality
vRAN	Virtualised RAN

Glossary of Terms

Some of these terms are extracted, mostly unchanged from [1], [2], [3], [4] and are reproduced here for convenience.

Table 1. Glossary of terms

Term	Definition
AV production	The process by which audio and video content are combined to produce media content. This could be for live events, media production, conferences, or other professional applications.
Clock synchronisation service	The service to align otherwise independent user equipment (UE) clocks.
Clock synchronicity	The maximum allowed time offset within a synchronisation domain between the master clock and any individual UE clock.
Compressed video	A means of making video file or stream sizes smaller to meet various applications. Some compression methodologies are: - **Mezzanine compression**: low latency and non-complex compression applied to a video signal to keep the maximum amount of information whilst reducing the stream size. - **Visually lossless compression**: the maximum amount of compression that can be applied to a video signal before visible compression artefacts appear. - **Variable bitrate compression:** the bitrate varies according to the source material. This is achieved by only encoding complex parts of the picture and predicting or repeating fewer complex parts often using long Group of Picture (GoP) technologies. - **Constant bitrate compression:** the bitrate is fixed. Typically, this is preferred for live work as it can provide the highest quality video at any given frame without the use of predictive procession. - **Adaptive bitrate compression** varies depending on the transport resources available. - **Highly compressed**: use of compression to distribute content over very low bandwidth connections where the content is more important than the quality of the image.
Communication service availability	Percentage value of the time an end-to-end communication service is delivered according to an agreed Quality of Service (QoS), divided by the time the system is expected

	to deliver the end-to-end service according to the specification in a specific area[2] [4].
Communication service reliability	Ability of the communication service to perform as required for a given time interval, under given conditions[3] [1].
Cue / talkback	Audio messages sent from the broadcast centre to a specific location usually to instruct a presenter when to speak. This is not audible in the broadcast audio. It can consist of multiple switched paths for different function such as camera operations or presenters.
Distributed production	A form of decentralised production where individual functions are enabled at different locations. Typically, utilised software and cloud-based resources as well as low latency connectivity.
End-to-end latency	The time that it takes to transfer a given piece of information from a source to a destination, measured at the communication interface, from the moment it is transmitted by the source to the moment it is successfully received at the destination [4]. Within 5G-RECORDS these terms would also be referred as glass-to-glass, mouth-to-ear, and motion-to-photon latency.
Isochronous data transmission	The time of an event or signal that is recurring at known periodic time intervals. Isochronous data transmission is a form of synchronous data transmission where similar (logically or in size) data frames are sent linked to a periodic clock pulse. Isochronous data transmission ensures that data between the source and the sink of the AV application flows continuously and at a steady rate.
Glass-to-glass/end to end latency	Time to generate an image from a source, to apply a certain amount of processing, to transfer it to a destination, and to render the resulting image on a suitable display device, as measured from the moment a specific event happens to the moment that very same event is displayed on a screen.
Mouth-to-ear	End-to-end maximum latency between the analogue input at the audio source (e.g., wireless microphone) and the analogue output at the audio sink (e.g., IEM). It includes audio application, application interfacing and the time delay introduced by the wireless transmission path.

2 The end point in "end-to-end" is assumed to be the communication service interface. The communication service is considered unavailable if it does not meet the pertinent QoS requirements. If availability is one of these requirements, the following rule applies: the system is considered unavailable in case an expected message is not received within a specified time, which, at minimum, is.

3 Given conditions would include aspects that affect reliability, such as: mode of operation, stress levels, and environmental conditions. Reliability may be quantified using appropriate measures such as meantime to failure, or the probability of no failure within a specified period.

Motion-to-photon latency	Delay between the movement/action of the user and the change in the display showing the appropriate content for that particular motion/action.
In-Ear-Monitoring (IEM) system	A specialist type of earphone usually worn by a performer in which an audio signal is fed to a wireless receive device and attached earphone.
Media clock	Media clocks are used to control the flow (timing and period) of audio and video data acquisition, processing, and playback. Typically, media clocks are generated locally in every mobile or stationary device with a master clock generated by an externally sourced grand master clock (currently GPS).
Non-public network	A network that is intended for non-public use [4].
Outside broadcast	A production where content is being acquired away from the broadcast centre and controlled from the location. Generates output for broadcast, which may be sent back to the broadcast centre for inclusion into a programme or for onward distribution. This may also be known as remote contribution and is mostly unidirectional.
Remote production	A production where the main control elements of a production are based in a different location to the production itself. This may be typified by having cameras and audio sources on site but the production operation being done from a production centre.
Survival time	The time that an application consuming a communication service may continue without an anticipated message [4].
Uncompressed video	Uncompressed video is digital video that either has never been compressed or was generated by decompressing previously compressed digital video (see also decoded video).
Video, imaging, and audio	The means of digital capture, transmission, and storage of still and moving pictures and sound for professional use.

1 Introduction and Objective

5G mobile networks are profoundly changing the way the media industry is creating efficient and scalable solutions for society. They offer improved performance in terms of bandwidth, reduced latency, support for accurate timing, assurance of quality of service and scenario flexibility. Furthermore, it is expected that standardized 5G-based solutions would bring down production costs, reduce environmental impact and increase the operational efficiency and flexibility of production workflows in news gathering, remote production and coverage of live events.

5G-RECORDS targets the development, integration, validation, and demonstration of 5G components for professional media content production. These components will be evaluated in specific end-to-end 5G infrastructures thanks to the participation of all partners (see Figure 1).

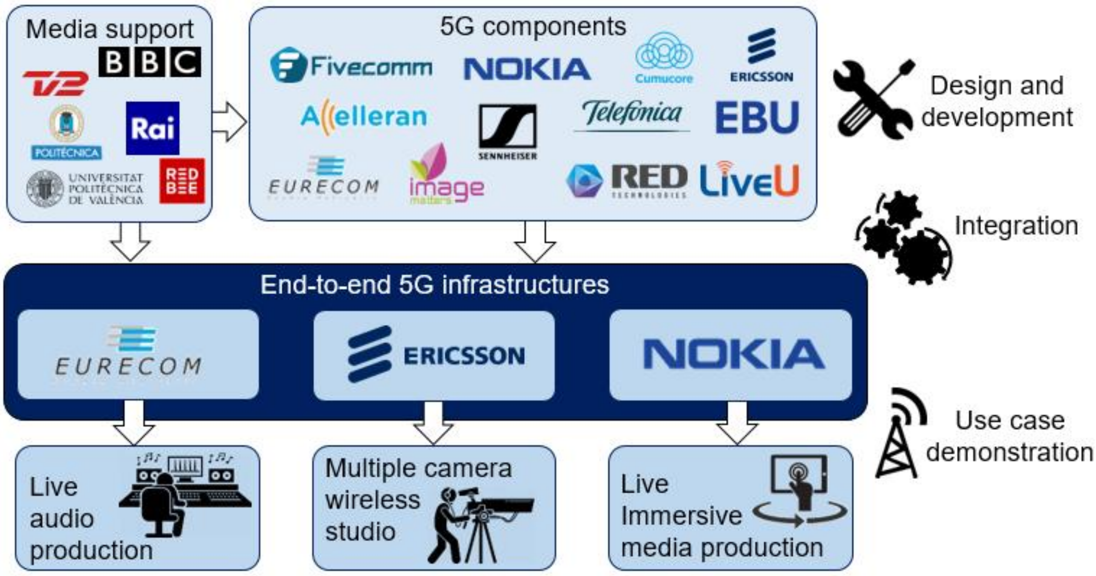

Figure 1. 5G RECORDS overview

This project will build on 5G components developed within previous 5G-PPP projects and earlier R&D investments and will further develop them for applications in content production, thus accelerating the industry rollout of 5G. 5G-RECORDS project focuses on three use cases: live audio production, multiple-camera wireless studio and live immersive media production. Each one of them has been designed to explore different aspects of the media industry, with the aim of exploring the opportunities that 5G brings to the audio-visual sector. It is important to consider throughout the project how 5G satisfies the high-level requirements of each use case and whether the minimum functionalities to be developed are feasible with this technology or not.

Key Technology Enablers

5G-RECORDS aims to explore the opportunities that new 5G technology bring to the professional audio-visual (AV) content production sector, taking advantage of 5G key features.

The project will put special emphasis in the deployment of 5G Non-Public Networks (NPN), given the opportunities they provide in terms of performance, privacy, data security and compliance. A Non-Public Network is a closed network like an enterprise network or like the media production network. The media producer controls, which devices get access to this network and can connect to the private media production infrastructure services. Conversely, NPN are tailored to satisfy the needs of a specific industry, such as media organizations, allowing them to deploy fixed and nomadic networks, with a coverage ranging from small areas like studios to entire premises like a so-called campus network.

The 5G System supports two forms of Non-Public Networks, namely a Standalone Non-Public Networks (SNPN) and a Public Network Integrated Non-Public Network (PNI-NPN).

An SNPN can be realized as completely isolated 5G System, leveraging dedicated spectrum, separated radio & core network hardware and localized computing for media processing. Media production related SNPN typically cover small and localized areas.

A PNI-NPN leverages to some extend resources from a public network. In the simplest form, the private network is realized using a Network Slice of a public network or just using a separate Data Network Name (DNN).

The deployment of these PNI-NPNs may go hand in hand with the use of network slicing for guaranteed Quality of Service (QoS). A paramount requirement for a network slice is to ensure that service level agreements (SLAs) for quality of service are met for all services running in a slice, taking into consideration all available network resources and capabilities. In this project, slices will be allocated to the different users of the networks.

More advanced forms of PNI-NPNs also leverage local breakout into a multi-access edge computing (MEC) environment.

The different forms of NPNs are described by 5G-ACIA in detail for industrial scenarios [5].

A key benefit of an SNPN is the full control and the clear isolation of the media production system. Networking resources and capacity are fully controlled by the media production organization. A key benefit of PNI-NPNs is the leveraging of existing network resources. This can increase the speed of setting up the network coverage and allows cost sharing with other use-cases.

For more details on NPN you can refer to deliverable “D2.3. Business analysis”.

Multi-access Edge Computing (MEC) is a great asset to seize new opportunities with applications for stadiums or localized events. The implementation of new Virtual Network Functions (VNF) instantiated very near to the radio provides very low latency and performance video capabilities. Edge computing will provide the slicing physical resources in terms of processing, storage, and CPU capacity to the required slices.

None of the previous advantages nor technologies would be useful without proper synchronization and timing. Professional content production requires specific QoS and timing parameters, with very high data rates for video sources, low latency for audio signals, very high reliability, and stringent content security measures dictated by the high value of the content. Live media content need precision when sending information (packets, in the case of IP world) to enable production workflows. As transmission speed is considerably high, synchronization is crucial and should be tolerant in terms of very small amounts of time (in terms of nanoseconds). Giving that several sources of media content are considered in a professional environment, it is a complex issue to tackle. Considering its importance, best fitting protocols will be considered and applied.

Another key aspect of 5G-RECORDS is the implementation of solutions based on open-source software running on general purpose processors. It provides developers and users with a great set of tools to simplify network access, reduce cost, increase flexibility, and accelerate the introduction of new services into the market. In this context, the use of up-to-date Software Defined Networking (SDN) is a cornerstone to open the proprietary interfaces to control the RAN hardware/software.

Use cases:

- **Live audio production**

As anticipated in the executive summary, typical locations for this use case range from theatres to festivals, where various artists are performing live.

The use of 5G in this use case brings inherent advantages to the beneficiaries, i.e., producers, artist, and the audience. Therefore, the system used must reach the musicians and artists' expectations on stage. Performers are equipped with PMSE equipment (5G wireless microphones and IEM systems). Depending on the level of expertise users can cope with different values of mouth-to-ear latency while depending on the scenario different audio qualities / data rates are acceptable.

The main objective of this UC focuses on the implementation of a local 5G wireless high-quality ultra-low latency audio production network, by designing 5G New Radio Reduced Capacity (NR-Redcap) audio devices prototype based on 3GPP. Given the low-latency requirement of this use case, and the trend towards on-demand/ad-hoc deployments, short set-up time, and process & workflow automation in live audio production scenarios, the use of NPN, e.g., SNPNs, is crucial. They provide a high level of flexibility in the network configuration as well as the security and privacy this use case requires.

- **Multiple camera wireless studio**

This use case is based around multi-camera audio and video production in a professional environment. This type of multimedia content productions requires high deployment and personnel costs, as well as high coordination of the teams located in the studio and at the event venue.

The use of 5G in this UC is based on the creation of a wireless and interconnected content production system with all the necessary functionalities for quality production and broadcasting. This use case will deploy a studio with 5G-enabled cameras for operation under remote production scenarios as well as exploring the interoperation with existing studio capabilities and logistics for distributed working.

Additionally, one of the use case scenarios will also take into account the operational requirements for working alongside wired studio equipment, as well as explore cloud-based workflows and how they can all be combined to produce high quality content suitable for broadcast.

Regarding the use of Non-Public Networks, the use case focuses on different NPN realizations: the use of Standalone NPNs is primarily in considered for localized production scenarios, where KPIs like a sustainable quality (bitrate) and a low latency are of key focus. The use of Public Network Integrated NPNs is considered for scenarios, where an existing network deployment should be re-used. It is assumed that the performance is guaranteed through a Service Level Agreement (SLA) between the Media Producer and the network provider.

- **Live immersive media services**

This use case considers a real-time, end-to-end FVV system including capturing, 5G contribution, virtual view synthesis on an edge server, delivery to 5G and FTTH users, and visualization on the users' terminals. The use case will allow real-time immersive capture of sport events or live performances. This means that each user can access a specific angle live, offering a unique QoE. It will also be possible to reproduce content both live and on demand. Besides, the content can be distributed not only to people attending the event, also to third parties.

The incorporation of 5G into the FVV pipeline will allow the distribution of the computational load by moving part of the processing to a MEC system (edge cloud), thus paving the way for possible future service virtualization. The extensive use of 5G for subsystem interconnection will allow the synchronization of all involved elements and the remote control and operation of the live immersive service.

The analysis of the use case will cover several scenarios, from which some of them will be selected for deployment and trial. The scenarios will be broken down from three perspectives: content production, content delivery, and system deployment. Thus, the E2E infrastructure will incorporate every 5G component needed for the use case deployment, and later, end-to-end tests will be performed in terms of network performance and QoE.

Besides, use case 3 proposes using a NPN which can be deployed on a need-basis in the location where the live event is going to happen: e.g., a music festival or a sports competition. This NPN with mmWave RAN will provide the required bandwidth to support wireless connectivity of the immersive media production cameras, as well as provide service to end users attending the event. The NPN will be supported by and connected to the network of a MNO (Telefónica), thus ensuring end-to-end QoS management.

1.1 Objective of the document

The goal of D2.1 is to define the use cases, KPIs and requirements addressed by the project. The issues in addressing these use cases with existing networks and approaches will be identified and how they will be enabled by the developments within 5G-RECORDS will be provided. The identified KPIs and requirements will be used to benchmark the existing state of the art as well as the 5G components developed within the project. Requirements coming from the content production industry as well as from other 5G PPP projects will be used as an input.

This project has envisaged three use cases that cannot be enabled by earlier mobile technologies but may be solved by using 5G networks. The deliverable will present the three use cases and their context, by describing the state of the art, scenarios, workflows, benefits, and main goals.

1.2 Structure of the document

The document is structured as follows. Firstly, Section 2 presents the "State of the art" with focus on media industry, users, roles, and workflow of the industry. After that, the core part of this document will consist of a detailed description of each use case: live audio production (UC1) in Section 3, multiple camera wireless studio (UC2) in Section 4 and live immersive media production (UC3) in Section 5.

Every use case section is split into four parts: (*i*) a description containing the user benefits and goals as well as the scenarios and workflows; (*ii*) pre-conditions; (*iii*) UC requirements; and (*iv*) KPIs.

2 State of the art

According to [6], the transformative impact of 5G will go well beyond just enhanced mobile media. It will disrupt the industry on many levels, including media and entertainment. 5G is expected to help bring a new, tactile dimension to entertainment. Nowadays, there is also great interest from the media industry on 5G networks as a key enabler technology for media content workflows, i.e., contribution, production, and distribution.

Media content is especially sensitive to the network performance, not only in the scenarios in which it is produced, but also when it is distributed to final clients where QoE is key. Media content usually requires high bandwidth, low latency, and high network availability. This is hard to achieve, especially in mobile networks such as LTE, with clear limitations. A 5G service-based architecture requires the introduction of new technologies that enable a wide variety of applications (e.g., NPNs, network slicing for guaranteed QoS, multi-access edge computing, virtual network functions, software defined network solutions, and accurate synchronization and timing techniques).

In terms of standardization, a set of professional content production use cases have been defined within the 3GPP Release-17 (Rel-17) study AVPROD. The Technical Report (TR) 22.827 [3] describes relevant use cases and proposes respective potential service requirements for 5G systems to support production of AV content and services. On the other hand, the Technical Specification (TS) 22.263 [2] describes the service and performance requirements for the operation of Video, Imaging and Audio for Professional Applications (VIAPA) via 5G systems.

Several efforts have been made (or are being made) to explore the new capabilities and possibilities that 5G brings to the table. To provide context, Annex A provides some 5GPPP projects related in one way or another to media applications. These are, for instance, 5G-MEDIA, 5G-Xcast, 5G Virtuosa and so forth.

In this context, the 5G-RECORDS project has envisaged 3 innovative use cases with different necessities and unique goals. Each one of them has been designed to explore specific aspects of the media industry. The state of the art of these use cases is described next.

2.1 Live audio production

Wireless audio devices are commonly used in live audio production scenarios since they offer a beneficial degree of flexibility for artists, see Figure 2. Meeting the latency and reliability requirements is a major challenge for such audio equipment. The stringent low latency requirement is because an artist combines audio source (microphone) and audio sink (IEM). The self-created audio signal is received via two separated paths, bone-conduction and IEM. Artists experience irritation and discomfort that can reduce the musical performance if the time difference between both audio paths is too large. Thus, a round-trip-latency (microphone to IEM) of 4 ms should not be exceeded in professional wireless audio systems [7] [8].

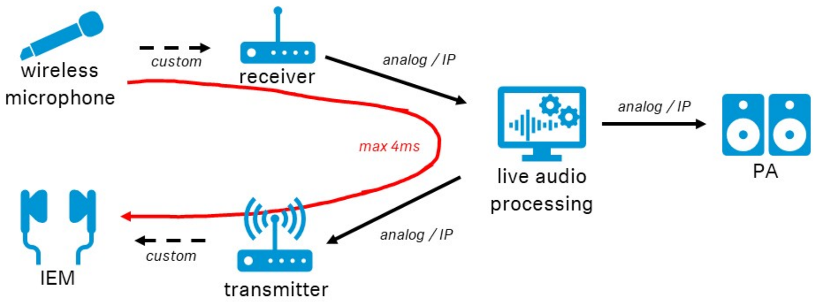

Figure 2. Today's live audio production

Today's wireless audio solutions are based on highly specialized link-based RF-technologies, since currently no standardized RF-technology can meet the strict requirements for latency and reliability. The ongoing transition from analogue to digital RF-technology still poses a major challenge to meet the latency requirement, even with specialized technology [9].

Another drawback is the fact that most wireless audio systems do not offer a method to synchronize media clocks in handheld devices. This implies the wide use of sample rate conversions, resulting in reduced audio quality and increased latency. Moreover, although wired audio production networks are increasingly using IP-based transport, wireless transmissions are typically not IP-based due to the necessary degree of optimization. Thus, current stationary audio transmitters and receivers are acting as gateway devices to IP networks.

On the other hand, planning and deployment of spectrum-use for wireless audio devices is typically done manually by RF-experts. This is increasingly supported by frequency planning software but remains a major effort in using wireless audio systems especially in more complex setups. In large live scenarios, such as festivals, spectrum is often a limiting factor for the number of simultaneously active devices, leading to excessive planning effort for such an event.

2.2 Multiple camera wireless studio

Live video production is a complex activity that typically is served by evolving specialized technologies, networks and radio solutions. The high bandwidth and low latency required to produce real time high-definition video requires dedicated point to point connections that have evolved from analogue, via digital, to IP based solutions.

Wireless connectivity plays a major part in live production where there is a need to have mobility, flexibility, and reliability.

Wireless cameras

There are two main scenarios where wireless cameras are usually deployed:

- Scenario 1: Wireless cameras within a production

Currently this scenario involves the use of RF cameras based on COFDM [10] technologies (Figure 3) which is based on the same technology used in the distribution of terrestrial television (DTT) broadcasting.

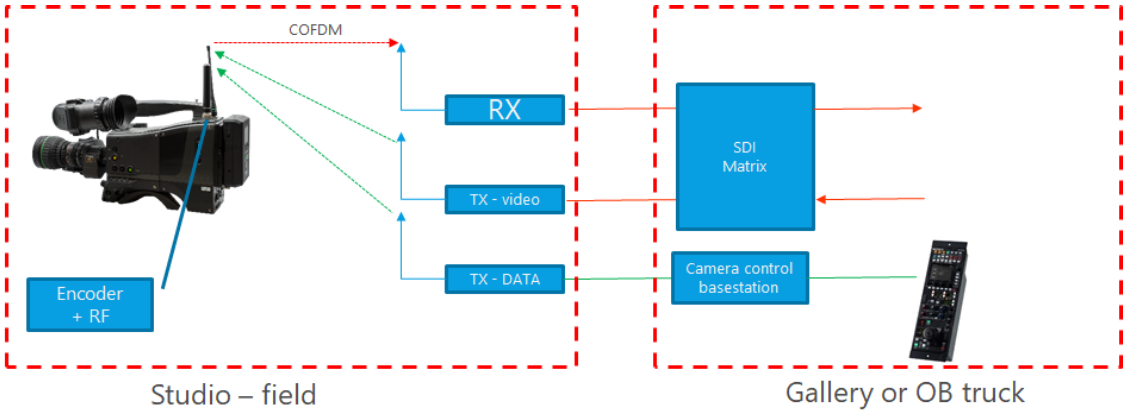

Figure 3. A typical wireless camera solution for studio or OB use.

These solutions have the advantage of being extremely low latency with relatively high bandwidths and with high quality of service availability. Dedicated systems are built around events and spectrum is available for production teams to deploy several cameras at a given location. Typical examples may include a camera on the touchline of a soccer match, on stage at a large music event or on a motorbike to cover a cycle race or marathon.

Different types of networks may be deployed depending on how the camera is used. For a single point-to-point (PTP) link, a dedicated peer to peer solution can be achieved with a simple transmitter and receiver set up. These may use either omni directional or directional antennas. For more complex set ups such as a studio or sporting event then a mesh network with multiple receivers may be set up. This allows the cameras to move freely within the coverage area while maintaining quality of service. Finally, for large area events then aerial relays may be deployed to cover a moving camera on the ground.

While these solutions are extremely robust, they do require specialist skills and knowledge to set them up. They are more common in territories that have DTT solutions, such as Europe, than those that do not, such as the USA. Spectrum is managed by the production technology teams.

As the consortium sees a migration away from digital solutions to IP based workflows and this spectrum becomes repurposed for other uses, there is a desire to explore how 5G may be used to replicate the functionality of these types of wireless cameras. Typically, they must operate alongside wired solutions and so latency and picture quality are of primarily importance and alongside these devices we also see requirements for wireless management, control and communication of camera, sound equipment and monitors.

These workflows are principally used for studio, sport, or entertainment production. The core aim for this scenario is to explore the substitution of COFDM technologies with 5G as well as enabling remote production via wider area connectivity or the cloud.

These scenarios require high level of movement and detail as well as low camera to vision mixer (switcher) latency. It should be noted that often the time taken to encode and decode highly compressed video can add considerably to the overall connection latency and therefore solutions need to use less complex compression which in turn will require hire throughput.

- Scenario 2: Outside remote contribution with distributed production elements

Over the past few years, broadcasters have been using mobile networks for some workflows, specifically using 4G networks to send a live video stream to the production centre. This type of communication has helped revolutionise the way news and events are produced, as reporters and teams could work from anywhere, anytime, if an acceptable coverage is available. To do this, a backpack or camera-mounted device is used to encode and broadcast video without the need for mobile units (vans) and/or satellite or microwave links.

However, the use of 4G networks can bring several disadvantages. For example, due to the bandwidth required, mobile solutions require multiple connections and therefore multiple SIM cards to provide adequate service; this method of connection aggregation is known as bonding. Additionally, when these devices are outside the mobile network provider coverage area, other SIM cards are required to use an alternate network. The video must also be highly compressed due to network bandwidth restrictions. To achieve this, it is common to deploy proprietary solutions that require a paired transmitter and receiver from the same manufacturer. These technologies tend provide a single video link and so if more than one camera is required it needs multiple units that are often timed differently. There is also no differentiation between the networks to which these devices connect and public networks, so in large events 4G connections become unreliable as they struggle for connectivity and bandwidth with other users.

It can be expected that 5G solutions will evolve to meet these workflows with little or no interventions but there is also a demand for a technology that allows you to connect and synchronize multiple audios and video sources as well as better interoperability with existing workflows and open solutions that allow more flexibility in deployments.

The scenarios for contribution are often focused on news and lower budget production. In these scenarios content may be more static with less temporal change or fixed backgrounds so more intense compression may be applied.

Considerations on cloud-based production

Productions typically require long preparation times with large audio and video equipment that is physically moved to external event sites, as well as configured and adjusted for a specific production activity. 5G networks themselves, despite the advantages they introduced, do not solve this problem. Some solutions such as cloud-based production are being investigated, which together with 5G networks may significantly change production workflows, as it will reduce the requirement to move all production equipment to the event site. This may lead to cost reductions or allow more coverage of complex events. For example, multimedia sources such as cameras or microphones would be deployed at the event site, but much of the equipment may be in production centres and be connected over the network to the remote site, examples include audio and video mixers, switching matrixes, storage devices and multi-viewers.

Some functions are coordinated in master control rooms (MCRs). These MCRs pull together multiple internal and outside sources and organise them for presentation to operational galleries. Large broadcast centres have large matrixes that allow multiple audio and video signals to be organised and packaged for both incoming and outgoing feeds.

2.3 Live immersive media production

New media consumption patterns and advances in production technologies have brought new immersive media services into the media ecosystem, especially in the production and distribution of live events, mostly sports. Premium sport events such as the Super Bowl or ATP Tour Masters 1000 have been already produced using 360 or 180 immersive video cameras, delivered to end users wearing head-mounted displays (HMDs). Immersive services for the audience in the stadium have also been experimented: video replay, AR statistics of the match, contextual gaming, etc.

The most significant factor driving the growth of this market is the increasing demand of 3D and 360 content in entertainment and AR/VR applications paired with the proliferation of content delivery devices such as AR/VR HMDs.

Some of the main players in this filed have made several developments in their products, specific to sports and entertainment applications. Microsoft, in partnership with Hammerhead, has developed a volumetric studio named Dimension [11]. Intel has developed a studio for filmmaking and has partnered with Turner Sports for providing its volumetric capture technology [12]. In February 2018, MR Museum in Kyoto, in partnership with Hakuhodo-VRAR, created a volumetric video for visitors about the beautiful and sacred artwork of the Zen temple in Japan narrated by a Zen monk [13]. Also, in February 2019, Intel collaborated with the National Football League (NFL) to offer immersive highlights via Intel's True View to the Super Bowl LIII in Atlanta's Mercedes-Benz Stadium [14]. These developments are expected to provide an impetus to the sports, events, and entertainment applications, resulting in it dominating the global volumetric video market by 2025, growing at the fastest CAGR.

A particularly relevant immersive service is Free Viewpoint Video (FVV), a technology that allows the user to freely move around the scene, navigating along an arbitrary trajectory as if there were a virtual camera that could be positioned anywhere within the scene. This functionality can improve the user experience in event broadcasting, such as sports [15] [16] or performances, and in interactive video communications such as immersive videoconferences [17] or interactive courses.

The development of such systems presents several challenges regarding video quality, real-time operation, and cost, which are often antagonistic. High-quality volumetric video benefits from the quality and number of cameras in the video acquisition stage, which in turn increases the deployment costs of the acquisition setup and requires more resources in terms of processing capabilities (compression and rendering) and network bandwidth. Additionally, high-quality volumetric video typically requires synthesis algorithms of such high computational complexity that prevent its real-time operation despite the use of high-end computation resources.

Examples of widely used commercial systems are the 4DReplay [18] and True View technologies [19]. 4DReplay provides video replays for sports, events, and film scenes in which the camera virtually travels along a predefined path. Their system uses more than 100 professional-grade cameras with hardware synchronization and inter-camera distance between 5 and 10 cm. True View technology uses an array of 30-50 high-end 5K cameras to capture volumetric data (voxels) of all the action. Those voxels allow the technology to render replays in multi-perspective 3D. Both systems can be framed in the high-quality end of FVV systems, requiring a dense array of professional cameras to produce high-quality video replays. However, the FVV functionality is limited, due to the use of predefined virtual paths and non-real-time operation. Moreover, virtual paths are typically designed to improve perceptual quality, rapidly swiping trough virtual views and stopping in viewpoints of physical cameras, which masks rendering artifacts. In the case of 4DReplay, the density of the camera arrays allows generating the viewpoint path from consecutive frames from adjacent physical cameras, without requiring virtual view synthesis.

To overcome these limitations, FVV Live [20] proposes a novel end-to-end FVV system conceived from the outset for real-time operation and high virtual video quality using off-the-shelf hardware, thus enabling low-cost and easy deployment. The key elements of this system are the following:

- An acquisition block comprised of a sparse array of consumer electronics stereo cameras, managed by a set of capture servers (CSs), which yields a multiview plus depth (MVD) format. The acquisition block includes a depth post-processing module which makes use of multi-camera calibration to correct stereo-pair calibration errors, improving the accuracy of the depth estimation.
- A compression and transmission block based on standard video coding schemes and transmission protocols. The design of the compression block focuses on preserving depth data to improve synthesis quality. To that end, it uses a bit re-arrangement step which enables the lossless encoding and transmission of 12-bit depth maps in an 8-bit format. In addition, to limit the overall bitrate, the transmission block adaptively enables/disables the data stream from each real camera depending on the position of the virtual one.
- A view synthesis module providing high-quality video with real-time constraints. This module runs at a single edge server (ES) and uses a layered approach, merging background (BG) and foreground (FG) layers projected from several reference cameras in the virtual view. This strategy affects the whole end-to-end processing and transmission chains, with the adaptation of acquisition (BG depth removal) and the view system modules (which uses FG and BG depth and colour data to efficiently synthesize high quality virtual views).

All these systems require to manage high bit rates for the production, and therefore are typically deployed over wired networks, as 4G did not provide enough network bandwidth in the uplink to support the required data rates.

2.4 Benefits of 5G with respect to 4G

This section outlines the main benefits provided by 5G with respect to 4G in the context of the three use cases described above.

The ITU defines the requirements for IMT-2020 system in ITU-R M.2083-0 [21]. 3GPP studies have shown, that 5G System fulfil the new IMT-2020 requirements. The spider diagram (Figure 4).

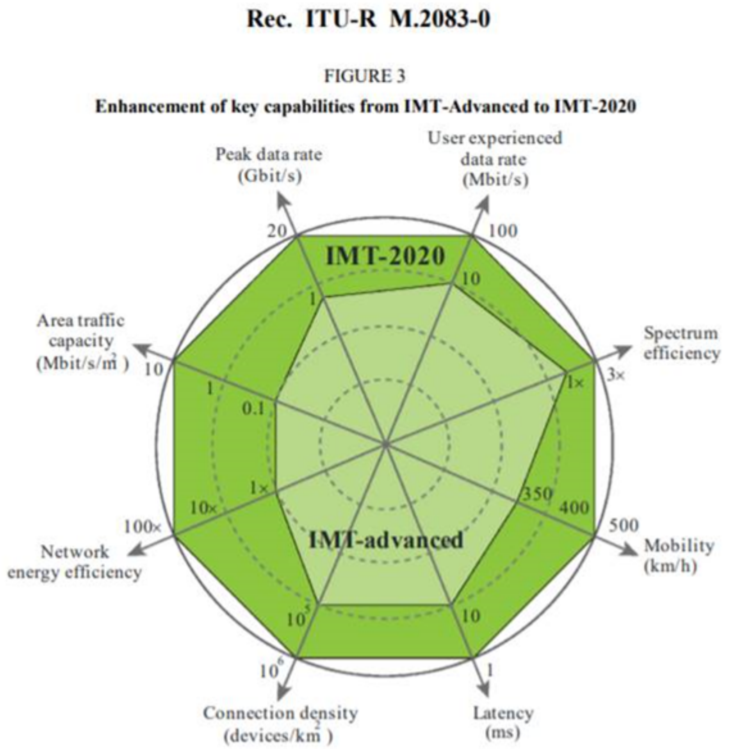

Figure 4. ITU-R's Enhancement of key capabilities from IMT-Advanced to IMT-2020

More specifically, the 5G Radio and Core Network components are improved compared to 4G Radio in the following areas:

- Support for lower latency: 5G supports faster radio response times than 4G, allowing for higher throughputs.
- Higher spectrum utilization: 5G supports better spectrum utilization than 4G, e.g., due to lower overhead.
- Increased energy efficiency: 5G radio supports higher energy efficiency, increasing the battery lifetimes and reducing the power-consumption of infrastructure components.
- Support for massive MIMO: improved MIMO improves other KPIs like coverage, capacity, data rate and reliability.
- Support of Full duplex radio links that reduce latency and increase available bandwidth by allowing simultaneous transmission and reception of data packets.

The capabilities of the 5G components have also improved compared to 4G components.

- Support for much larger bandwidth, allowing for higher bitrates. The new bands in Mid and High Frequency Range support those larger bandwidth
- Support for cloud native deployments, allowing for more flexible, agile and distributed deployments.
- Support for native Ethernet transport and TSN
- Support for IP multicast transmissions
- Support for time-synchronization over 5G (Ethernet and IP based IEEE 1588 profiles with Rel-17)
- Network slicing support in 5G is more flexible than 4G eDECOR.
- Support for Non-public networks, allowing to serve industry verticals.

3 UC1: Live audio production

3.1 Use case description

In a typical live audio production, such as a concert, musical, theatre or studio performance, one or several artists are performing (i.e., acting, dancing, singing, or playing music instruments) live either to create content that can be used later or to entertain an interested audience that can be live on site or follow the content live via stream. Figure 5 depicts this scenario as well as the key enabler technologies that allows its deployment within 5G-RECORDS.

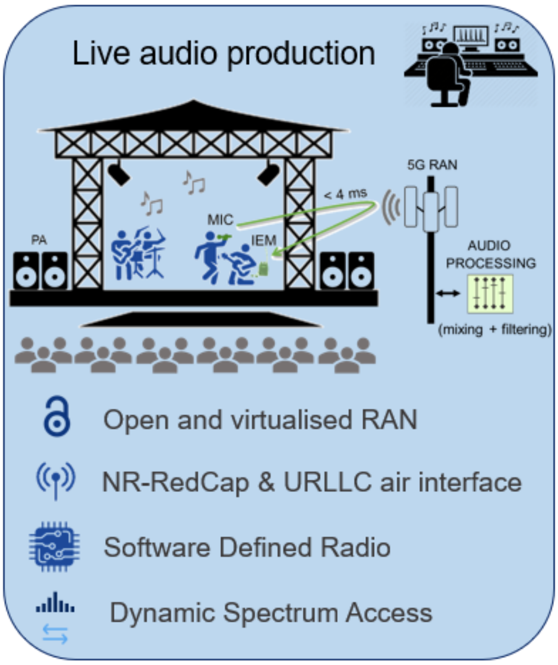

Figure 5. UC1 key enablers overview.

This scenario has considered four roles.

- Role A: main user of any on-stage equipment.
- Role B: the producer that is responsible for any further creative processing of captured audio data to create a final consumable product.
- Role C: Distributor that delivers the product for example via local PA system, broadcasting, or internet stream by a distributor.
- Role D: the consumer.

Capturing and producing of a live event for subsequent use of the cultural and creative content, involves many wireless audio devices. For instance, artists on stage use wireless microphones to capture their voices or instruments' sound while hearing themselves via a wireless IEM system.

Wireless microphones are used for capturing the voice and music signals of the artists within the stage. Consequently, content capture is expected to take place at the highest quality possible, with producers taking steps to ensure the integrity and robustness of content capture and delivery. For these reasons, the quality and reliability of the radio links are fundamental to wireless audio device users. For live audio productions especially, the commercial pressures on operators are significant as there is no opportunity for recovery. It is not possible to ask for a repetition during a live concert, so the tolerance for QoS is extremely low.

As shown in Figure 6, all latency critical processing such as individual IEM mixes is done locally, whereas non-critical audio processing can be done outside the local high-quality network. Today, wireless audio devices are designed in a link-based approach, which means that every audio link consists of one transmitter (wireless microphone) or receiver (IEM) and its corresponding device. In Figure 6 this is simplified by showing only one base station for that purpose.

Many artists rely on receiving a personalized audio mix of the event streamed back to their IEM device. In this context, personalized means that each artist can receive a different audio mix (i.e., unicast downlink transmission) fully adapted to his or her needs and preferences.

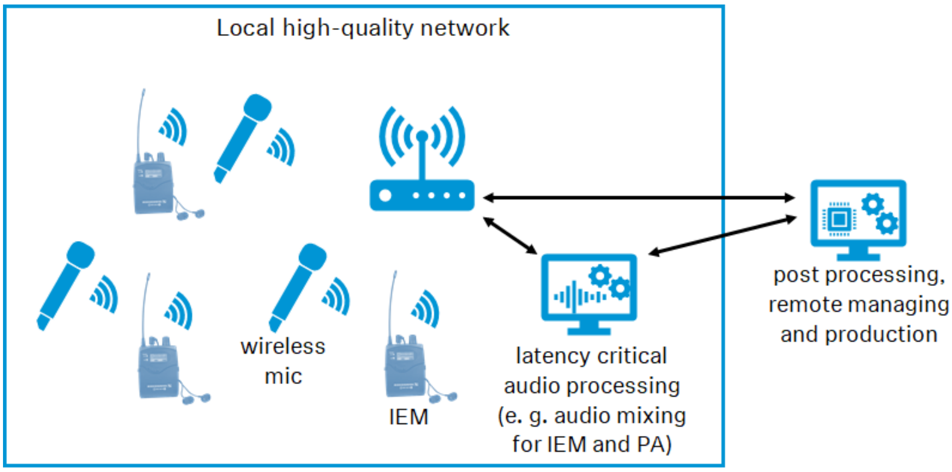

Figure 6. Live audio production network

Live events take place typically in theatres, concert halls, stadiums, and production studios. The stage can be located indoors or outdoors. Typical operation has a duration known in advance. All wireless audio equipment required for the capturing and production of an event is always available at the location of the event. It either belongs to the staging supplier or a rental company has been engaged to deploy it for the event. The wireless communication service is limited to the event area and all audio processing such as audio mixing is done in real time during operation.

Wireless audio devices are most likely to be used at different events, which may have different application set-ups in terms of number of devices and their capabilities. Therefore, before a professional live audio production starts, audio devices may need to be quickly reconfigured. Considering the number of wireless devices that may be involved in a live production, it is desired that those devices can be automatically and quickly provisioned through the network.

Spectrum Sharing and licensing

Access to frequencies used by the PMSE devices in this use case is not based on exclusive licences. Instead, the spectrum used by the PMSE devices can be dynamically shared over time, space, and frequencies. This enables to optimize spectrum utilization through coordination. To ensure a guaranteed quality of spectrum, a given user shall not be interfered more than a pre-defined threshold by other nearby users.

However, prior to accessing shared spectrum, PMSE users will need to obtain authorisation with a licence from the NRA or another spectrum owner (e.g., MNO). There are two types of licences that will be explored in this use case:

1) Long term fixed site licence

A long-term fixed site licence will enable continuous access to spectrum for 5G systems by a local licensee for regular, ongoing production events (e.g., a daily or weekly live production in a studio). In this use case, the licensee is the production studio owner or facilities supplier.

2) Short term temporary licence

A single short term temporary authorisation to access spectrum will be enabled by licensing spectrum for 5G systems to either a production event organizer (e.g., a summer jazz festival in the South of France) or the facilities supplier for a short period at a specific location. The spectrum owner could be an MNO or a local licensee who has available spectrum at the time and location of the event.

3.1.1 User benefits and goals

The use of 5G in this use case brings inherent advantages to the beneficiaries:

- **Producers**
  5G as a low-latency high-bandwidth IP network and gateway to other wide area IP networks will support the general technology and protocol conversion of today's media workflows. This is also related to the increasing relevance of remote production workflows, which are often based on IP networks. In some scenarios 5G allows to use the same network technology for content creation and distribution which reduces the complexity of the production network and simplifies workflows.

  Deploying 5G as the same wireless connectivity technology as a system-based approach for different types of media devices (e.g., microphones and cameras) will result in faster, easier, and more optimized setup and operation of wireless networks, opposed to today's operation of various wireless technologies for different sections of a media production.

  5G as a system-based approach will potentially also significantly reduce today's effort for planning and handling of spectrum for wireless devices before and during an event.

- **Artists**
  5G will transform the way artists make music and, consequently, the way they enjoy it. 5G will bring about some improvements like clock synchronization, which will result in better sound quality for the streamed audio and better production tools.

- **Audience**
  In case this UC is deployed in scenarios like festivals, musicals, or conference, where there is an audience attending a live performance, they can also benefit from this innovative 5G-based setup. The content that is captured by the wireless microphones on stage will also be streamed so that the audience can follow the event. They will be immersed in an improved live experience that will let them enjoy a high-quality performance without interruptions and potential failures.

3.1.2 Scenarios and workflow

For UC1 a couple of different scenarios exist. All using wireless audio transmission for microphone and IEM links and for all scenarios a low mouth-to-ear latency and a high reliability is important. Each scenario could either happen in an outdoor or an indoor environment. Another common part is that the sound captured by the microphones is transmitted to a central audio processing/mixing device where individual audio mixes for the IEM devices are generated. These IEM-mixes are transmitted to the IEM devices. The typical scenarios within this use case are:

- **Festival**

For a festival it must be possible for multiple artists to perform on multiple stages on the festival location. The actors of this scenario are not only the performers, but there might be also a wide audience attending the event.

- **Musical**

For a musical scenario, the service area could be much smaller than for a festival, but e.g., the speed of the artists could be significantly higher.

- **Semi-professional**

In a semi-professional scenario (e.g., small newcomer bands) the number of wireless links is rather small, and the audio quality requirement is a bit lower than for the other scenarios.

- **AV production**

In an audio video production, a very high audio quality is important. The number of wireless links will be moderate.

- **Audio studio**

In an audio studio production scenario, the audio quality is the most important requirement since an artefact in the audio signal transmission will be part of the produced recording. Up to now in this scenario only wired technology is used for that reason.

Table 2 shows an overview of the requirements for the different scenarios of UC1.

Table 2. Scenarios for use case 1

Scenarios	Device type	# of active devices	Device Speed [km/h]	Service Area [m²]	Audio Quality
Festival	Microphone	200	10	500 x 500	better
	IEM	100			
Musical	Microphone	30	50	50 x 50	better
	IEM	20			
	Loudspeaker	10	-		
Semi-professional	Microphone	10	5	10 x 10	good
	IEM	10			
	Loudspeaker	2	-		
AV production	Microphone	20	5	30 x 30	best
	IEM	10			
Audio Studio	Microphone	30	-	10 x 10	best
	IEM	10	5		

This use case will focus exemplarily on the AV production scenario to demonstrate the feasibility of 5G for professional audio equipment. A typical production workflow is shown below, consisting of activities taking place both off-site and on-site.

- **Off-site workflow:**

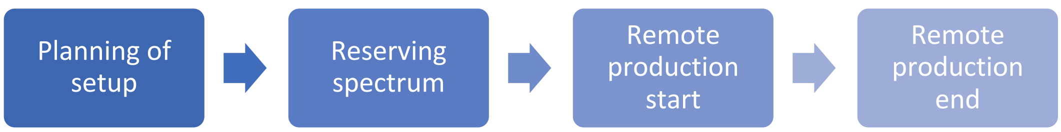

Figure 7. Off-site workflows in UC1.

- **On-site workflow:**

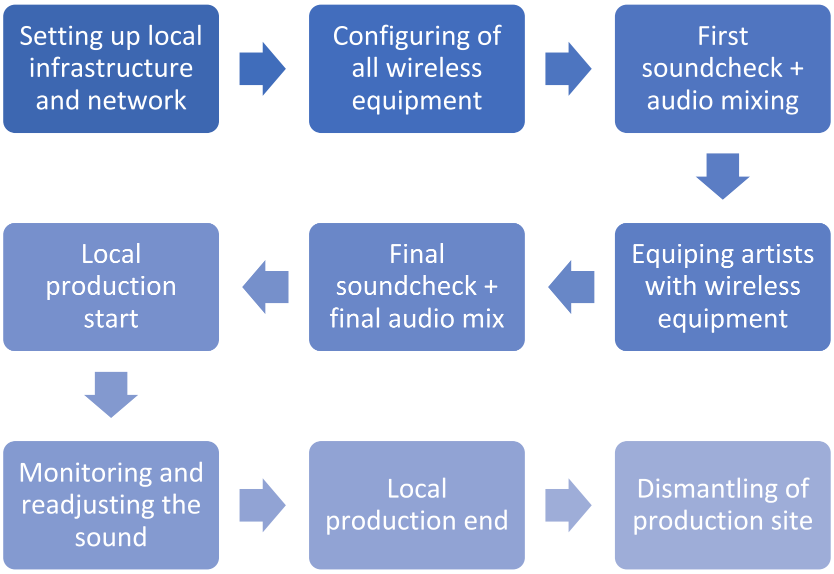

Figure 8. On-site workflows in UC1.

3.2 Pre-Conditions

These following pre-conditions [3] have been set:

- All PMSE equipment required to produce the event is available at the desired location (stage).
- Sufficient bandwidth to satisfy the requirements of the wireless production streams are available at the event location during the whole operation time.
- All wireless microphones and in-ear monitoring devices are provisioned with configuration and credential information for the communication service through the network.
- All wireless microphones and in-ear monitoring devices on stage are switched on and connected to a central audio mixing device through a 5G-RAN.
- All active wireless microphones and in-ear monitoring devices are synchronized at the application level to a reference clock with the required accuracy.
- An audio mixing device is connected to amplifiers and loudspeakers of a public-address system.
- The processing of audio data might be time aligned to the access to the 5G RAN
- An audio mixing device can mix all incoming audio streams into the set of outgoing audio streams for the IEMs.

3.3 Use case requirements

The system has to fulfil the musician’s requirements. Depending on the level of expertise the user is able to cope with different values of mouth-to-ear latency and depending on the acoustic environment different quality of the audio data is acceptable which corresponds to different data rates. In Table 3 the requirements for the audio data flow of UC1 are listed.

Table 3. Requirements for UC1 concerning audio data flow.

Characteristic system parameter		Comment
Mouth-to-ear latency	< 4 ms	Maximum application latency tolerated by a live performer between her analogue audio source (wireless microphone) and her analogue audio output (IEM). It includes two times network latency + audio processing time. Ideal latency would be < 2ms as 2 to 5ms latency can cause change in tone. However, <2ms may not be achievable.
Audio data rate	2.4 Mbit/s	Different user data rates per audio stream need to be supported for different audio demands (e.g., compressed vs. uncompressed audio).
Device must be remote-controllable	≤ 50 kbit/s (Control data rate)	Data rate per control link (UL/DL).
Number of audio streams	20 microphones 10 IEM	Simultaneous audio links.
Service area	30 m x 30 m	Event area, indoor and outdoor. Typical heights of indoor stages: 5 m to 10 m.
User speed	≤ 5 km/h	Devices/artists can move.
Reliability	99.9999%	A high reliability for the audio transmission is needed to guarantee that no noticeable artefacts will be introduced to the audio signal.
Security/ Integrity	The audio application data is encrypted	In some of the applications.

3.4 KPI

KPIs are measurable values which prove that predefined performance targets of different kind of systems are reached. The Table 4 shows the set of KPIs of UC1.

Network latency is composed of transfer interval (periodicity of the packet transfers) plus E2E latency. E2E latency refers to the duration between the transmission of a data packet from the application layer and successful reception at the application layer of the destination. System jitter is not specified because jitter is part of the network latency. The network latency must be constant during the whole operation time.

Packet error ratio is related to a packet size of (transfer interval × data rate). Packets that do not conform with the network latency are also accounted as errors. The given requirement for a packet error rate assumes a uniform error distribution. The requirement for packet error rate is stricter if packet errors occur in bursts.

To fulfil the requirement of a low audio latency and a high audio quality all audio devices in the network have to use synchronized clocks for sampling/processing. This includes the application on UE side and wired audio devices connected to the 5G core.

Table 4. KPIs of UC1.

KPI	Limits	Description
Network latency	< 1 ms	Consists of E2E latency and transfer interval. It is assumed that 2ms are left for audio processing within the mixing console. This is the latency from the application layer on UE side to the application layer on a device connected via UPF to the 5G core (or the other direction).
Synchronicity	<500 ns	The absolute difference between any synchronized clock in the network and the time master has to be below the limit.
Packet error ratio	$< 10^{-6}$	The packet error ratio (PER) of the system shall be below 10^{-6} for a packet size corresponding to 1 ms audio data. Further, a consecutive minimum continuous error-free duration ≥ 100 ms has to be ensured. This is because, to make packet errors inaudible, error concealment is used at application level. Every concealment is capable of handling one specific kind of error distribution.

4 UC2: Multiple camera wireless studio

4.1 Use case description

This use case is based around multi-camera audio and video production in a professional environment. It will enable existing technologies such as COFDM radio cameras to be replicated in terms of performance and capabilities using 5G technology.

The use case will require radio links to integrate into existing studio facilities and perform with or exceed similar characteristics of existing RF solutions. The cameras will operate in a similar way to existing studio setups and require services such as tally lights, control signals and audio.

This use case will also explore multi-location scenarios with production facilities local to an event as well as remote and distributed production models. In some additional scenarios, it is expected to integrate 5G based contribution solutions into production centres using different types of network configuration to provide contribution links.

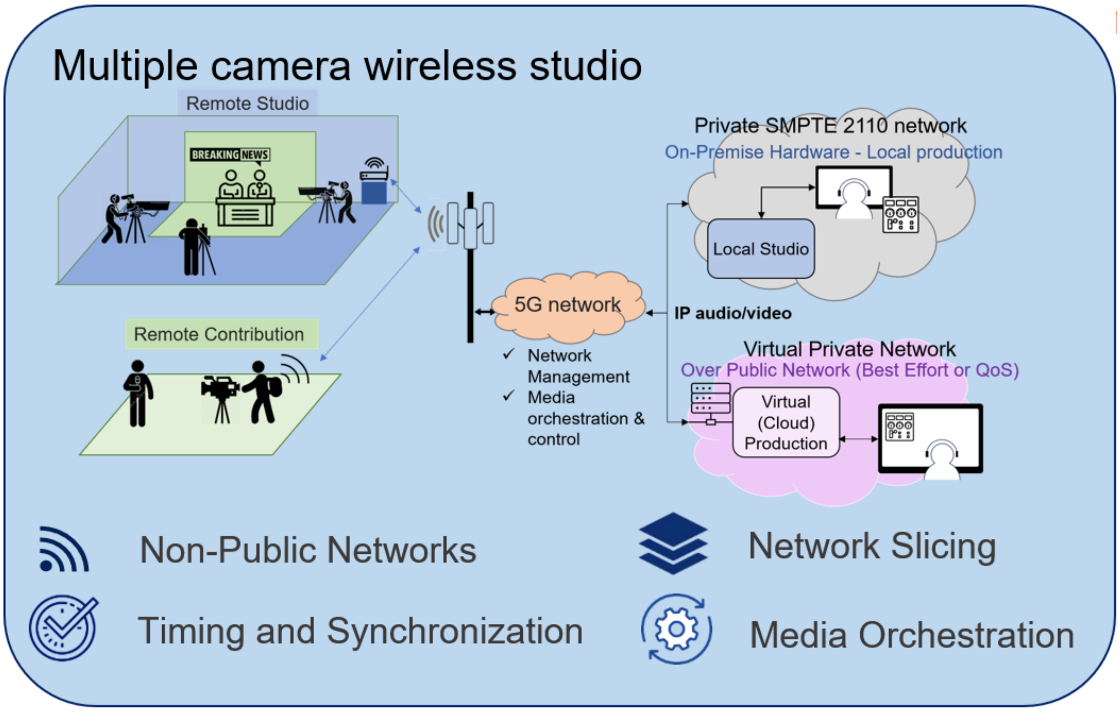

Figure 9. UC 2 key technologies.

In the UC2, it aims to develop a system that can support different scenarios and configurations for media content production, where the following functions are expected to be provided:

- Discovery, integration, and management of 5G-enabled wireless cameras within an IP studio production network.
- Delivery of streams produced by 5G-enabled wireless cameras to a remote studio whose transport infrastructure is based on SMPTE ST 2110 with minimal latency.
- Support for seamless A/V source processing (e.g., mixing with local sources).
- Time synchronisation of devices in the field and in the studio.
- Feedback (e.g., return video or tally) and controls signals (e.g., RCP) are sent back from the studio to the production location.

Since Media Production is generally moving towards IP based networking solution, 5G might bring additional flexibilities and levels of freedom. Using 5G, also a range of off-the shelf technologies, which were developed for consumers can also be leveraged in media production scenarios.

While cabled cameras, guaranteeing error-free steady performance, can be used in many circumstances, there are scenarios where this is not possible or practical, e.g., when the movement of the camera cannot be constrained, or in some outdoor events in locations where deploying long fibre cables is not possible, or in some sports event, where cables could interfere, etc.

The use case will require multiple 5G enabled cameras to be connected to an NPN and operate simultaneously and allow the seamless transmission between all of the cameras with no timing glitches or visible artefacts. The starting format will be 1080p50 HD video, testing on HDR. Higher raster sizes may be explored if compression and network bandwidth are available.

Latency is critical, and it is a requirement that is it needs to be predictable and constant rather than purely pushed to the lowest possible. QoS is even more critical than latency in the most part so techniques that minimise errors at the expense of latency may be applied.

In large broadcast operations, a MCR may be deployed to monitor and manage incoming and outgoing feeds. In these scenarios there may be multiple studio and gallery configurations as well as contribution sources. Depending on the type of set up each studio may have its own NPN or a shared NPN.

In Section 4.1.2 Scenarios and workflow, different scenarios that can benefit from the development of this UC will be presented.

It should be noted that some of the scenarios are reliant on emerging technologies such as cloud-based production which is out of scope for 5G-RECORDS. IP based production is in its infancy and the scenarios are targets, some of which are deliverable in a shorter timeframe whereas some will rely on future 3GPP releases and on the evolvement of technologies such as cloud-based production.

Architecture and inner workings

In Figure 10, the general architecture of the UC2 is shown:

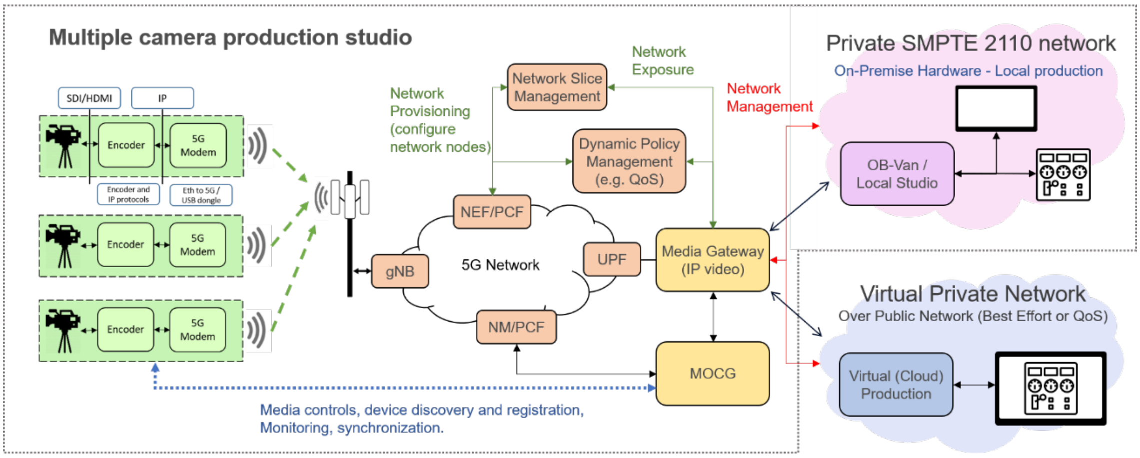

Figure 10. UC2 general architecture.

In this proposed architecture, the main components are shown:

- Media equipment with 5G connectivity: it provides the multimedia streams and need to have connectivity to 5G network and the two gateways. A variety of different media sources can be used in different configurations.
- 5G RAN, core, and functionalities: these provides the base for the use case communications, as well as advanced network functions such as network slicing.
- Media Operational Control Gateways (MOCG): these two components manage the media streams and device authorization, registration, and controls.
- Production networks: provide the infrastructure to work on the media sources and controls.

With this architecture (see Figure 10), it is expected to have a complete production environment; cameras, microphones, etc., deployed for the production in the event location 5G connected and controlled from another location in another location. A more detailed architecture is shown in Figure 11:

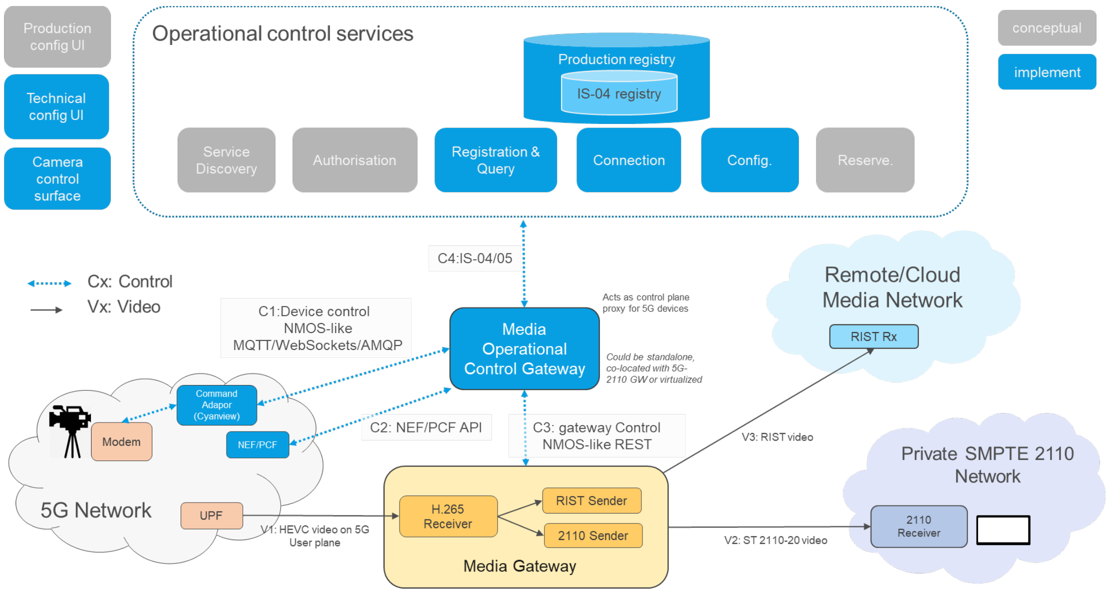

Figure 11. UC2 detailed architecture

As the media and operational control gateways will be a core development focus, following will be a description of these two elements:

Operational control layer & operational control gateway

The operational control layer is a key element in the UC2 5G-RECORDS architecture. Whilst the current Rel. 16/17 standardisation has focused on improving the QoS parameters in terms of latency and bandwidth and timing, the application of professional media production functionalities requires the definition and development of an additional orchestration layer above the 5G infrastructure capabilities in the wireless studio for: (i) automatic discovery and registration of professional media devices, (ii) automatic resource management and provisioning, (iii) control and monitoring functionalities,(iv) timing and genlock capabilities, management and performance management, and ideally, the operational control layer (see Figure 11) will orchestrate the communications flows between the production islands, the 5G core and the gateway.

In an IP based production environment, the management of signals also needs to be more carefully managed than in a traditional broadcast facility where dedicated point to point cables are replaced by IP routing. This means that the addition of a **media control**

layer also needs to be considered. The function of this layer is to provide services that define:

- **Stream transport protocols and formats**

This chooses the optimum transport and codec choices dependant on the resources available and configures all devices in the chain to support this. This also includes raster size and framerate as well as options for HDR and colour gamut.

- **Discovery of resources**

This function identifies resources on the network and their preferred functionality and makes them available as a resource on the production network.

- **Connection of media and data flows**

It manages and prioritises bidirectional traffic.

- **End-to-end Timing and Synchronisation**

It provides timing data to enable synchronisation of multiple sources over the networks.

- **Authentication and Authorisation**

It identifies the device and allows it to connect to the networks.

Other services that may be provisioned are:

- Provisioning of dynamic resources.
- Grouping and labelling of resources.
- Configuration of resources.
- Conversion of media and data flows.
- Operational control.
- Monitoring.
- Logging.
- Process Automation.
- Resource Allocation, Reservation.
- Resilience management and monitoring.

The Media Gateway

This gateway converts the 5G protocols and media formats (e.g., audio and video over RTP) into the ST 2110 compliant protocols and formats and vice-versa so the producers can work with the media streams. This gateway acts as a “protocol translator” that adapts these streams into the format that a specific network or equipment supports.

The interface between the 5G network and the broadcast network (media gateway) will require some definition and translation of signals to and from the radio and studio domains. This gateway will need to maintain high quality at low latency. Latency, while low, should be constant to allow predictable performance that can be measured and compensated for.

4.1.1 User benefits and goals

The main benefits of using 5G network components and remote production by user type are:

- **Producers**
    - Remote production reduces team and equipment travels, increase teams’ safety and cost improving the flexibility, mobility, and reliability of my production operations.

- Usage of 5G Networking equipment reduces the production network setup time (by reducing the cable deployment time) and increases the flexibility.
- When cameras are moving during production (continuously or temporarily), the cabling of cameras is often impractical (e.g., during live events) to impractical (e.g., drone mounted cameras).
- Off-the-shelf 5G End-user equipment like smartphone and tablets can be utilized in some scenarios as professional production equipment, e.g., during news gathering (Tier 3 events) or for editing and control.
- 5G components provide a more sustainable field production while increasing efficiency, providing live coverage of less viewed sports and events.

- **Engineering managers:**
  - Having the ability to provide low cost, high quality outdoor remote production using IP and Cloud technologies together with either 5G NPN Systems or 5G PLMNs.
  - Overcoming the limitations of the RF-based cameras in outdoor, uncontrolled areas.
  - More standardized setup of wireless networks, providing a larger coverage (e.g., entire event site).
  - Simplified equipment management due to more aligned 5G system components. Equipment can be used in different SNPNs (for Tier ½ events) and in PLMNs (for Tier 3 productions)
  - Increased radio performance e.g., using more advanced radio techniques like MIMO, multi-cell systems, etc
  - Usage of on-prem dedicated hardware or public or private clouds for remote productions increases flexibility.
- **Facilities:** supporting multiple instances of mobile remote production.

Potential drawbacks of using 5G Network Components and Remote Production

- **Capacity:** 5G Equipment implementation is driven by downlink-traffic oriented consumer use-cases
- **Latency:** Usage of 5G drives the usage of video and audio compression in order to increase the number of cameras and devices in the 5G System

The main use case goals are:

- To test and trial to demonstrate the capabilities of 5G NPN in a “wireless studio” environment.
- To work together with the different stakeholders (e.g., network providers, manufacturers, service providers, etc) to align the requirements.
- To develop production workflows integrating 5G technology, especially relevant to full IP wireless studio productions.
- To discuss commercial and operational models for the employment of NPNs and network slicing in a wireless studio UC.
- To integrate multiple workflows and deployments to allow a flexible approach to 5G production.

To understand which camera controls, currently proprietary, could be managed through open protocols/APIs.

4.1.2 Scenarios and workflow

In any production there are pre-production as well as production and postproduction workflows. A typical production workflow is shown below:

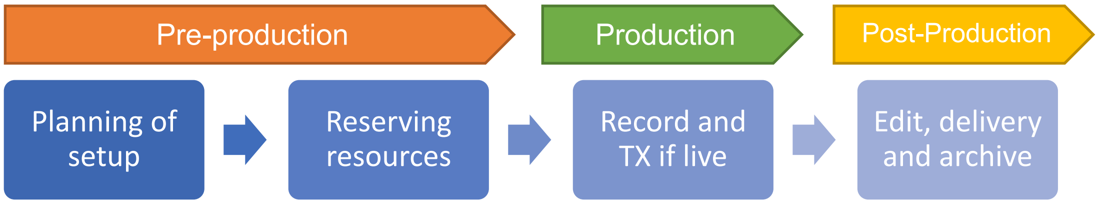

Figure 12. General UC2 workflow

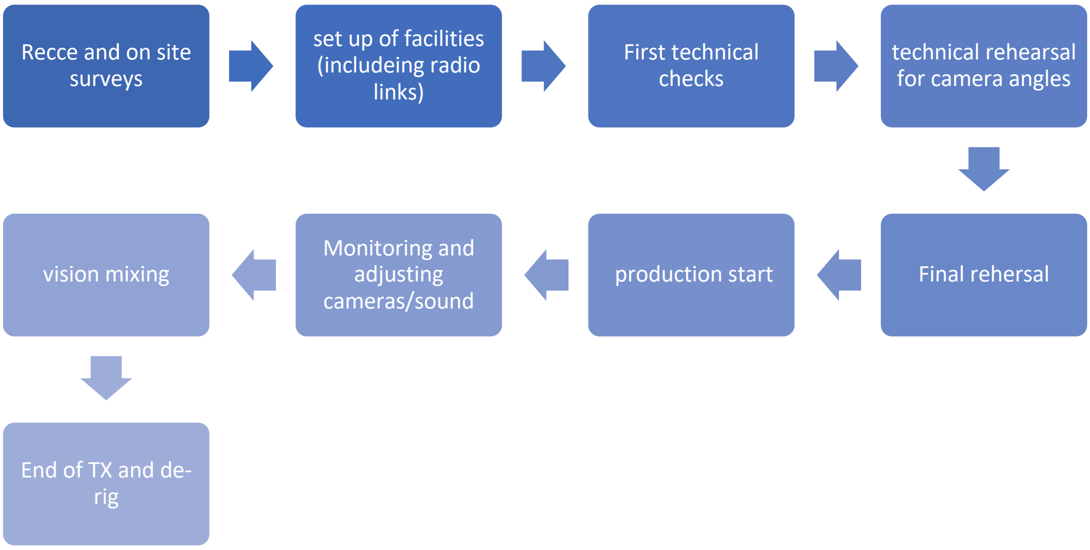

Figure 13. UC2 production workflow.

With respect to the workflow (see Figure 13):

- The operational control layer will orchestrate the communications between the 5G core, the gateway and the production equipment.
- The Network slicing manager (Cumucore) will enable the multiplexing of independent logical networks on the same physical network infrastructure to fulfil diverse requirements.
- The signals (SDI: video plus embedded audio) will be encoded using an external encoder provided by Image Matters or Fivecomm in HEVC over RTP. In scenario 2, A/V feed(s) will be encoded (HEVC) and transmitted via LiveU LU800 multi-cam bonding device. Remote control of cameras will also be provided over the LU800 IP PIPE.
- The RTP streams will feed the modem(s) provided by Fivecomm.
- Through the 5G network, the streams will get to the 5G to ST 2110 gateway (media gateway): this gateway will perform the decoding of the HEVC streams and mapping them to ST 2110 and wrap the RTP stream to RIST for the distribution through the cloud.
- The gateway is also responsible to encode the incoming return video feed and pipe it through the 5G network back to the cameras.

- The video and audio streams are delivered to the “remote studio”/MCR in the cloud.
- The camera controls signals (RCP, tally, and intercom) will be transmitted over IP.
- In scenario 2, the video transmitted by the LU800 will be received, decoded by LiveU LU2000SMPTE server and output in SMPTE 2110 formats (with video redundancy, PTP-synched to local master clock).

In particular, the sequence of the events when connecting a wireless device (camera, microphone, intercom, headset, autocue, lighting, etc) will be:

i. **Connection to the 5G network**: Power on. Initialize modem. Establish IP data link.
ii. **Connect to operational control layer**: Authorise device for use in production. Register device services (e.g., camera) for 5G. Register device for operational control.
iii. **Production configuration**: Get initial setup from production configuration. Enter session-specific details. Set up connections from/to the device.
iv. **Production operation**: Director / vision mixer operator calls shots and makes cuts. Audio mix changes as required. Rack’s operator makes technical adjustments during shoot. Start / stop recording as required. Other activities including logging, monitoring, communications.

For this project we will explore two main scenarios: wireless cameras within a production (scenario 1a); outside broadcast contribution (scenario 2) with the integration of distributed production elements (scenario 1b).

4.1.2.1.1 Scenario 1a: Wireless cameras within a production

This scenario forms the reference production scenario, where the wireless camera(s), are equipped with a compression encoder and a 5G modem to inject the audio/video streams into a 5G network which is then decoded and connected with the SMPTE ST 2110 IP studio (or OB van) via the UPF (User Plane Function) interface. A figure depicting this scenario is shown below:

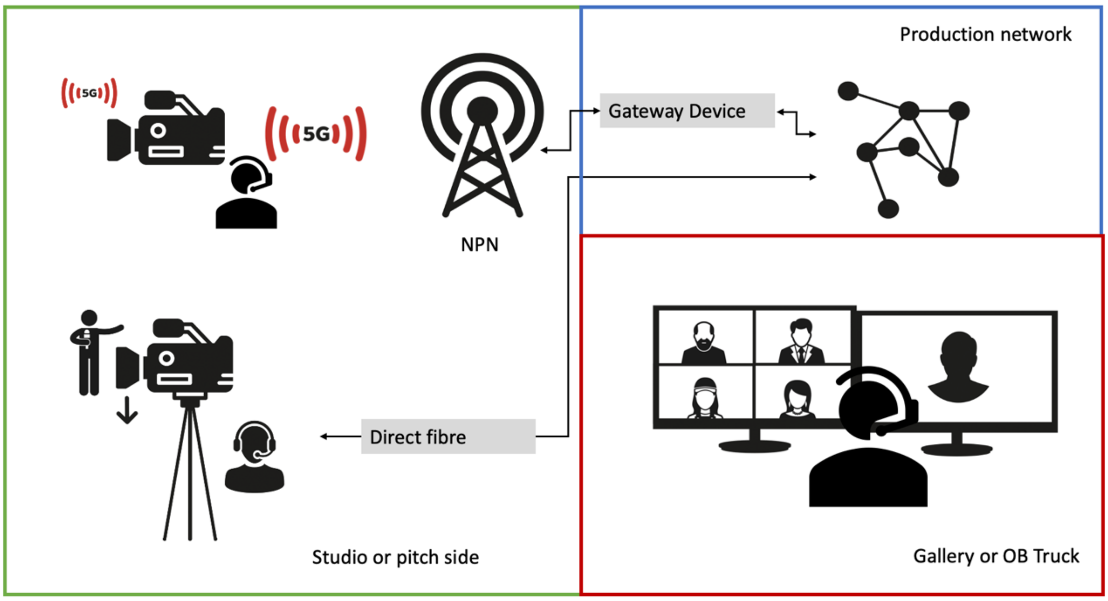

Figure 14. Scenario 1 schematic.

In a wireless studio or OB scenario, a 5G NPN will be used, and all production cameras will be connected over it, replacing the RF/COFDM or cables that are currently used.

In these scenarios there is a requirement to encode complex, fast-moving video in a high quality and with very low latency. This can be achieved by using mezzanine codecs which will require higher throughput bandwidth but can encode and decode video in less time than more complex compression algorithms.

In many scenarios conventional RF/COFDM links are practical enough, they allow only low-range connections (a local receive point or an OB Van is needed). with extended coverage leveraging a larger 5G core network infrastructure, the coverage can be extended to any remote location, provided that network connectivity is fast and reliable enough. With 5G, it is easier to cover a larger area using multiple base station. Transmissions between base stations are coordinated to allow 5G device mobility and to reduce interference. 5G Radio devices have evolved over time, supporting a high range of signal strength improvements or noise reduction techniques such as adaptive Modulation and Coding Schemes (MCS) or different Multiple-Input / Multiple Output (MIMO) schemes.

Further, 5G camera's and other 5G enabled devices can be brought easier between event sites and connect to a different SNPN system.

5G uses extensively IP within the 5G System to connect different 5G Network Components and on 5G User Plane. The same type of equipment can be used to setup the 5G system and the ST 2110 media production network.

Full production with multiple studio cameras will be demonstrated, at a very low latency (ideally < 40ms glass-to-glass), time-synchronization, remote control of the cameras, and control of the other equipment types that are used in a studio such as lighting, monitors, and microphones. The production and equipment will be managed via a new media orchestration platform and a 5G NPN resource allocation layer. The performance will be at least the same as current studio production, with the added benefits provided when using the 5G and the 5G-RECORDS media orchestration layer.

It is expected that the studio infrastructure will operate with uncompressed FHD video at 50 fps (1080p50). If successful, then higher resolutions and framerates will be tested. This is to maintain quality and keep latency down for audio and video in the studio domain. Any radio technology will realistically require compression of video. Any production studio also needs to compare with parallel technologies, so performance is critical. An HD wireless IP solution should have a comparable performance with existing radio cameras that ideally provide an uplink bandwidth of 50 Mb/s with an end-to-end latency of around 20 ms.

This scenario will also explore the addition of remote production to the workflow (presented in Figure 15). Here the gallery operations and camera location are not collocated but connected via an IP link. The gallery will process the incoming feeds as if they are co-located in the studio space. This is known as remote production and differs from an isolated outside broadcast or remote contribution scenario which is self-contained on location. It is an extended scenario where the IP production studio is not co-located to the cameras but in a remote location, potentially in a different country. Here single or multiple sources are sent back to the studio.

There may be a temporary deployment of an ad-hoc NPN at the location or cameras may be individually connected back to the gallery via an MCR. If public IP networks are being used from the NPN to the studio, additional latency and uncontrollable instability might be added, which then needs to be accounted for and overcome.

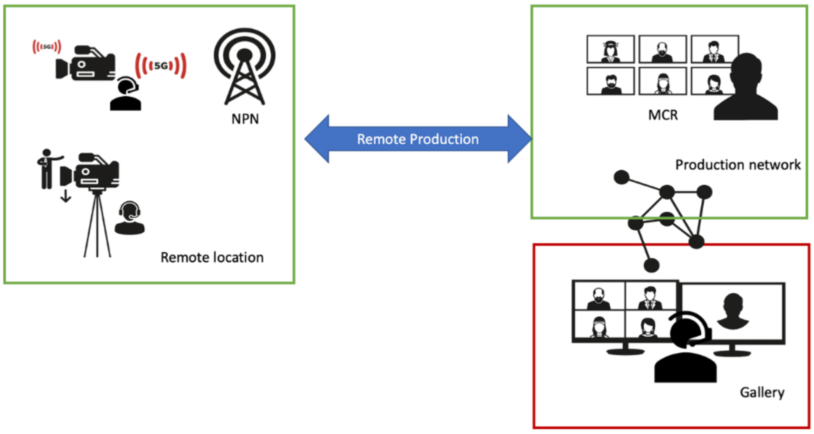

Figure 15. Remote production scenario.

In some cases, the cameras would be relatively close to each other and probably being serviced by the same cell (e.g., a small venue or few cameras), whereas in other cases they could be widely spread and sometimes even mobile, e.g., in a city Marathon. Some of these cameras would be remote controlled whereas others would be locally controlled, or partially remote controlled (e.g., for shading but not for focus, zoom, direction etc).

The communications between the production team in the gallery and the team in the field is important. It should be done within the workflow solution rather than using smartphones etc. Therefore, effective talkback communications with all camera operators and location staff are essential. Low latency is required (< 80 ms) to enable the camera operator to best respond to the gallery instructions. Similarly tally light at low latency is desired. Talk groups should be supported to allow the producer or director to communicate with both single and multiple listeners.

Figure 16 describes an additional scenario, where two 5G networks are connected: the IP studio is directly connected to a “fixed” network, and a second “nomadic” network can be reached via the cloud.

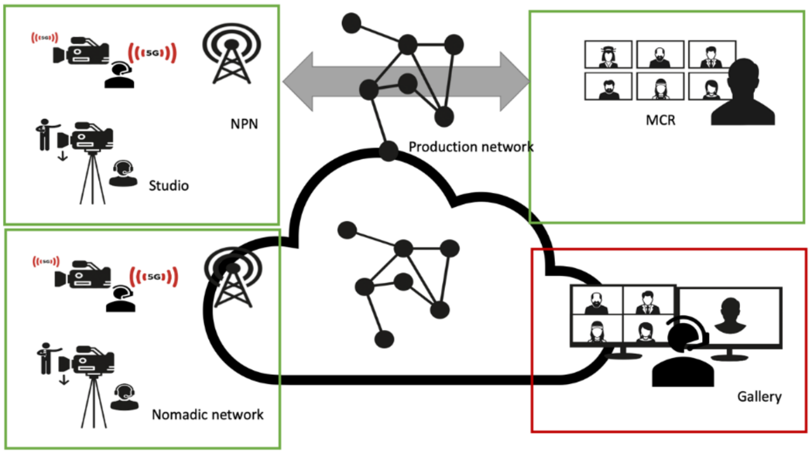

Figure 16. Additional scenario where two 5G networks are connected.

The interconnection with the 5G network and the production network may be done either via a point-to-point link, over a VPN or via the public Internet. This will depend on the requirements of the production in terms of performance or functionality. Connectivity and interoperability between the audio/video generated by the cameras and SMPTE ST 2110 signals should not be done at the border of the 5G network, but at the remote site. In particular:

- A connection based on VPN allows to transparently carry all required signals (audio/video delivery and control signalling) and appear as if carried on a LAN (provided that introduced latency and packet disruption are acceptable). It may require a one-to-one agreement between 5G telecom operator and broadcaster to establish it.
- A connection via the public network may require the use of specific protocols to transport audio/video streams and control signalling. This may be by means of proxies at both ends. It may be that traffic can be prioritised over this network in agreement with the network operator. If the 5G operator and broadcaster IP networks do not have a direct interconnection at an Internet exchange point, a third-party operator may be required to provide an IP link that guarantees adequate throughput and QoS.

Camera units

Typically, a studio set up will consist of multiple signals These may be centred around a studio or OB truck or around a single camera depending on the scale of the production. In addition to the video signal out it can include a return video feed, camera control, tally light, communication etc. as shown on Figure 17 below.

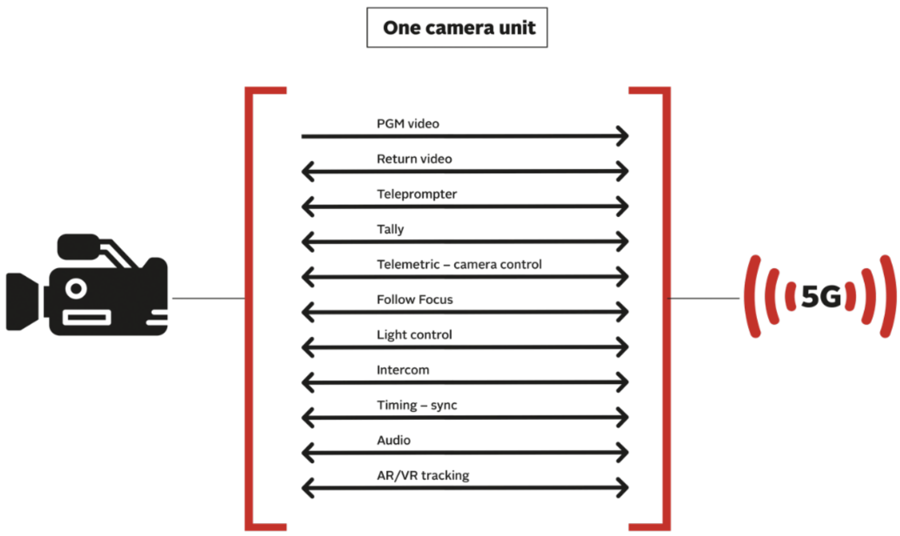

Figure 17. Different signals connected to a camera.

In a conventional setup these signals may be carried on separate radio links and in different frequency bands. It is hoped that they can all be combined in a single 5G connection. Not all production scenarios will use all the services for any given production. Table 5 indicates some likely scenarios but is not exhaustive.

Table 5. Production scenarios service requirements.

Scenario	PGM Video	Return Video	Teleprompter	Tally	Telemetric	Focus	Lighting control	Intercom	Timing	audio	AR/VR
Sport	☑	☑		☑	☑			☑	☑	☑	☑
Studio	☑	☑	☑	☑	☑		☑	☑	☑	☑	☑
Steadicam	☑	☑		☑	☑	☑		☑	☑		
Stage	☑	☑		☑	☑				☑		

Table 6. Camera unit traffic characteristics.

Data	Bandwidth	Direction	Latency	Application
Program Video	High (50Mbs)	uplink	Low (<20ms)	Main video output from camera.
Return Video	Medium (5-10 Mb/s)	downlink	low	Reverse video feed for monitoring.
Teleprompter	Medium (5-10 Mb/s)	downlink	low	Text feed for teleprompter.
Tally	Very low	downlink	low	Red lights on active camera.
Telemetries	low	Bi-directional	low	Remote controls for exposure and colour control, possible PTZ.
Follow Focus	low	downlink	low	Focus control.
Lighting control	low	downlink	low	Control onboard camera lights.
Intercom	medium	Bi-directional	low	Allows director to talk to camera op. May be 2 feeds.
Timing and sync	low	downlink		Provide timing data to camera to allow for video synchronisation and IP time stamps.

4.1.2.1.2 Scenario 1b: Integration of cloud-based distributed production

In large broadcast operations an MCR may be deployed to monitor and manage the incoming and outgoing feeds. In these scenarios there may be multiple studio and gallery configurations as well as contribution sources. Depending on the type of set up each studio may have its own NPN or a shared NPN may operate over the campus.

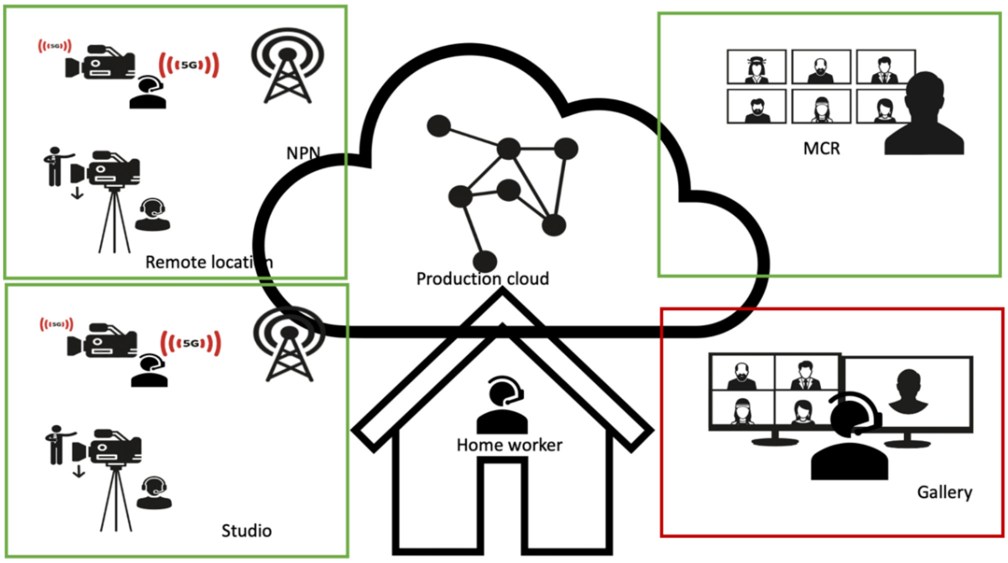

Figure 18. Distributed cloud-based scenario.

Some gallery functions may be performed in different physical locations (Figure 18), for instance the vision mixer (or switcher) may be in a different location to the person controlling the camera racks (or shading) and other functions such as the director, producer, graphics, or sound operations could be performed anywhere with adequate connectivity. In some instances, this will require the generation and distribution of lower quality 'proxy' video signals that will be distributed amongst the production team who then select the source to which the high-quality video stream will be sent. In some instances (such as checking camera focus or shading) it is necessary to see the video in as high quality as possible.

4.1.2.1.3 Scenario 2: Outside remote contribution with distributed production elements

The second scenario is where a self-contained operation is managed at the location and a feed of either a single stream of pre-produced content or individual feeds of multiple sources are returned to the broadcast centre usually via an MCR. One possible media feed is 5G-enabled LU800 device – a multi-link video encoder-transmitter.

This scenario is very common in news and minority sports and provides a flexible low-cost solution for live production. In this scenario remote contribution is provisioned over 5G and the receiver could be either located at a central location or in the cloud. The network will likely be a PLMN but may have traffic prioritisation or slicing provisioned by the network operator. The basic schema of this scenario is presented in Figure 19.In this scenario the feed from the field will be received by an SMPTE2110 video receiver (LU2000-SMPTE), which also communicates with the field device to dynamically coordinate the transmission according to the momentary multi-hop route performances. The feed can then be integrated into the studio SMPTE video processing.

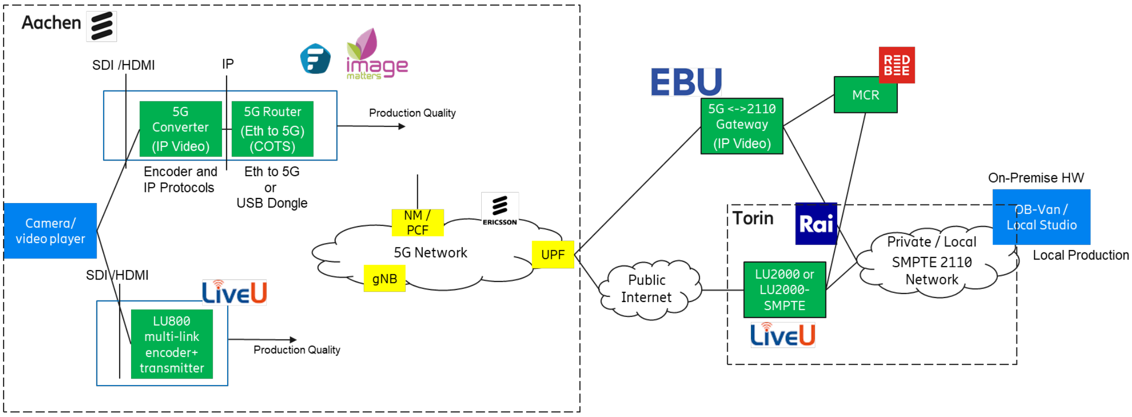

Figure 19. Remote contribution basic scenario

4.1.2.1.4 Combination of different scenarios in a production environment

A broadcast TV programme may require a combination of both scenarios 1 and 2. If we consider a news or magazine programme, we may have a studio based on scenario 1 technologies and workflows and contributions from outside sources that based on scenario 2.

To integrate both to work together to produce a multi-location event and prove interoperability between multiple production environments with varied platforms and functionalities, as shown in Figure 20, all scenarios may require signal routeing and monitoring via MCR functionality which may be either physical or virtual.

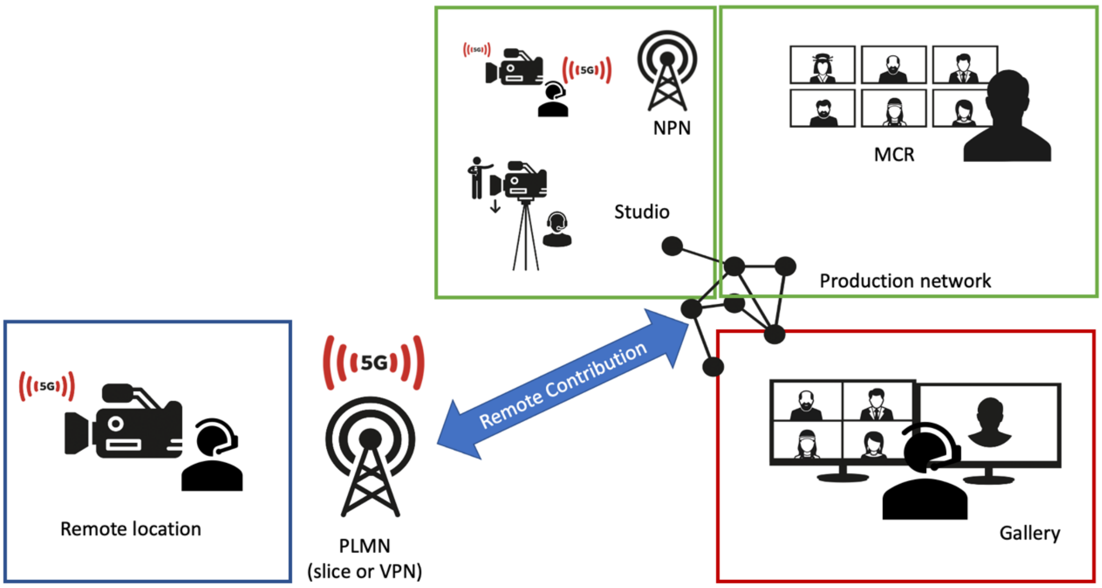

Figure 20. Integrated scenario with remote production

In this model all incoming sources are available on a matrix and those relevant for a specific production are then routed through to the studio gallery. Alongside this there may be requirements to route contributions elsewhere for parallel production for other output channels such as radio or online production.

The MCR will also route reverse signalling to the studio or remote location, this may include reverse audio and video, control signals and other data.

The MCR may be based physically in the same location as the studio and gallery or, more recently, hosted in a cloud or virtual environment. The MCR will also support other functions such as output routing where a final finished programme is sent for transmission.

4.2 Pre-Conditions

- All PMSE equipment required for production (e.g., cameras, microphones) are available at the desired location.
- All PMSE equipment and software required to produce the programme (e.g., monitors, production switcher control, etc) are available on premises and/or remotely.
- Sufficient connectivity to satisfy the requirements of the wireless production streams are available at the event location during the whole operation time.
- All the equipment is synchronized to a reference clock with the required accuracy.

Specific pre-conditions for scenario 1:

- An SA NPN that supports the user requirements.
- Connectivity to the production network.
- Device configuration and management solutions.
- Availability of the MCR service
- Connectivity between the MCR and the gallery/galleries
- Connectivity in the cloud
- Production tools availability in the cloud

Specific pre-conditions for scenario 2:

- A PLMN that available at the contribution location.
- (Optional) a slice of a PLMN at the contribution location.
- A receive server connected to the production network.

4.3 Use case requirements

- The system should support a consistent latency and jitter model.
- The system shall support a standard MTU.
- The interactions, including authentication and authorization, between the equipment on premises, the 5G network and the remote production island shall be orchestrated by the Operational Control Layer.
- A gateway shall perform the conversion of media and protocols coming from the 5G network to provide ST2110 uncompressed streams and RIST streams.
- Communication between the local crew and remote crew shall be enabled.
- All the equipment can be controlled remotely.
- Control of cameras on premises, remote controls shall include racking operation and PTZ.

The following tables (Table 7 and Table 8) describe the use case requirements that are scenario-specific:

Table 7. Requirements for UC2 Scenario 1.

Characteristic system parameter		Comment
Glass to Glass latency	20-300 ms (ideal <40ms)	Latency from an image being captured by a camera to the point it becomes usable in a production gallery (discounting onward distribution.
Video uplink Data Rate	>50 Mb/s (ideal 100Mb/s+)	This is to allow high quality video. different compression algorithms may be deployed depending on the format of the video.
Service area	1000m^2	Typical small studio area.
Mobility	≤10km/h	Support for walking speed or robotic mount.
Number of Streams	Up to 5	

Table 8. Requirements for UC2 Scenario 2.

Characteristic system parameter		Comment
Glass to Glass latency	<1s	Latency from an image being captured by a camera to the point it becomes usable in a production gallery (discounting onward distribution.
Video uplink Data Rate	>20 Mb/s	For this type of contribution, we typically compress more and avoid movement/detail within the frame.
Mobility	≤100km/h	For use in a moving vehicle.

4.4 KPI

KPIs (Key Performance Indicators) are measurable values which prove that predefined performance targets of different kind of systems are reached. Table 9 shows the set of KPIs of scenario 1 and Table 10 the KPIs of scenario 2.

Table 9. Key Performance Indicators (KPIs) of UC2 (Scenario 1).

KPI	Limits	Description
Synchronicity	< 1ms	The absolute difference between any synchronized clock in the network and the time master must be below the limit.
Packet error ratio	$< 10^{-8}$	(Packets that do not conform with the end-to-end latency are also accounted as error (the packet error rate requirement is calculated considering 1500 B packets, and 1 packet error per hour is $10^{-5}/(3*x)$, where x is the data rate in Mbps and then rounded).

Table 10. Key Performance Indicators (KPIs) of UC2 (Scenario 2).

KPI	Limits	Description
Glass to Glass latency	<1s	Latency from an image being captured by a camera to the point it becomes usable in a production gallery (discounting onward distribution.
Uplink Bandwidth	>30 Mbps	Consistent (not peak) over 15 minutes to allow high quality HD, good quality UHD or good quality dual-camera streams
SMPTE ST2110-20 compliance	A/V output with redundancy	Compliance of the LU2000SMPTE Gateway with ST2110-20 and ST2110-30, video redundancy, PTP master clock support; minor issues allowed

5 UC3: Live immersive media services

5.1 Use case description

This use case considers a real-time end-to-end FVV system including capturing, 5G contribution, virtual view synthesis on an edge server, 5G delivery, and visualization on the user's terminal. The system will generate in real-time a synthesized video stream from a free-moving virtual position. An FVV system generates synthetic views of a 3D scene from a virtual viewpoint chosen by the user (yellow camera) by combining the video information from several real (red, green, and blue) reference cameras.

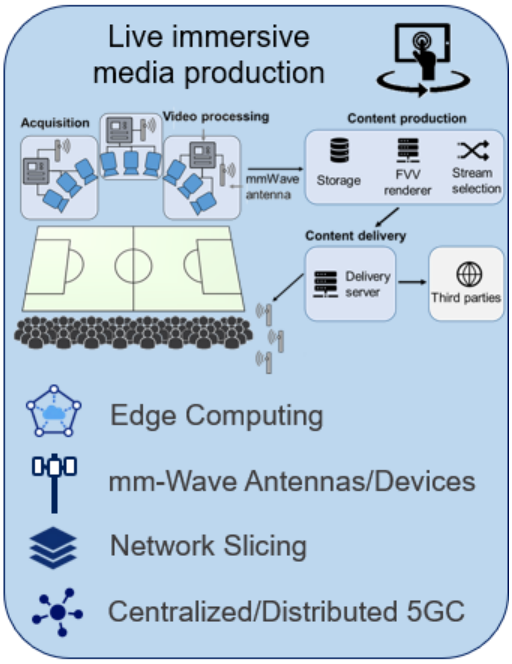

Figure 21. UC3 key enabler overview.

The proposed system builds upon an existing FVV prototype called FVV Live conceived for simple videoconferencing scenarios under controlled lighting. Instead, the envisaged use case targets, among other possibilities, the real-time immersive capture of sport or entertainment events (such as a basketball game, a dance performance, or a concert). It will be possible to reproduce content both live and offline (replay) of free-viewpoint trajectories around one specific area of the event (e.g., one basket of the court, the downstage, etc.). The content can then be distributed not only to people attending the event (local delivery), but also to third parties. Finally, it must be stressed, once more, that the position of the virtual camera can be freely chosen at any time.

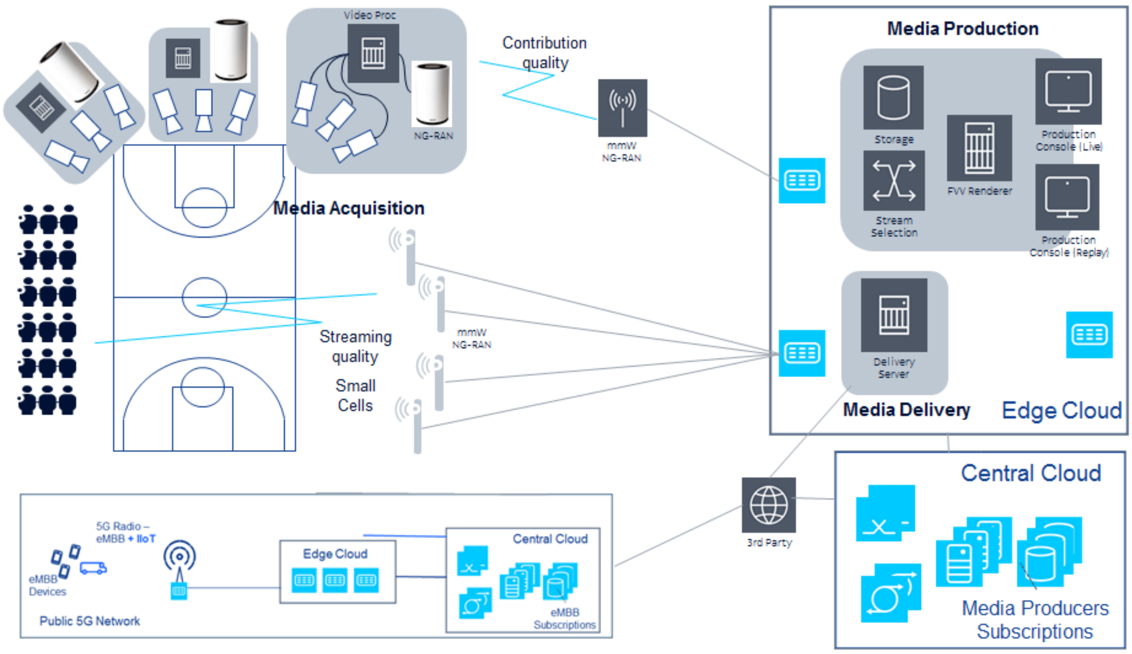

Figure 22. Live Immersive Media Production system concept.

The live immersive media use case is also considered from an end-to-end perspective, which not only includes the acquisition and production elements, but also the delivery to end users both locally and remotely. The use case builds around the idea of *smart venue*, the concept of multi-purpose modern sport stadiums or theatres, which can host a variety of event times.

For instance, the most famous sporting mecca of all, Madison Square Garden in New York, sometimes holds three different events in the same space on a single day: a circus performance in the morning, an ice hockey match in the afternoon and a basketball game in the evening. Visitors to any of these events rarely know another event took place on the same day when the physical infrastructure works as it should. However, they should still be able to access through the technology to new services that enhance the fan experience, such as watching the scene from different viewpoints through their smartphones.

5.1.1 User benefits and goals

5G connectivity will allow a portable FVV system to operate in real time, so that content producers can choose dynamically the trajectory of the synthesized view during a live broadcast. The incorporation of 5G into the free-viewpoint video pipeline will allow the distribution of the computational load, thus paving the way for possible future service virtualization. The extensive use of 5G for subsystem interconnection will allow the synchronization of all involved elements and the remote control and operation of the live immersive service. Thus, the innovations stemming from this use case will help fast location deployment and, therefore, will simplify on-the-spot media acquisition.

The foreseen system will operate with high-resolution cameras, will profit from fast calibration procedures, and will be endowed with enhanced depth-map estimation and pre- and post-processing for optimized approaches to background and foreground modelling, leading to the generation of high-quality synthetic views. All system interfacing and control will be handled by the 5G network.

Additionally, the FVV system will be integrated into an end-to-end workflow, including local attendants to the events, final users in other locations, and local or remote content producers. The QoE of the different users must be maximized, and also weighted with respect to their criticality (e.g., event producers should have higher priority than individual end users). Different scenarios will be evaluated to study the scalability of the use case to several deployments, from small theatres to large stadiums. A representative subset of such scenarios will be selected for trial and evaluation.

Some innovative features of the proposed live immersive media service are:

- Wireless (5G) connectivity between elements.
- Portable: fast setup and calibration of the whole system.
- Live generation of FVV stream (not only for replay, but also non-freezing video while changing viewpoint).
- Scalable and adaptable to different deployment scenarios.

Moreover, the foreseen scenario for the use case operation will include:

- Real-time immersive capture of an area of the scene.
- Live and offline (replay) generation of free-viewpoint trajectories.
- Distribution to people attending the event and to third parties.
- End-to-end QoE guarantee based on network slicing.

All the described advantages and innovations are intended to be used and enjoyed by the beneficiaries listed below:

- **Big events organizers and professional producers**, who will get a never seen before system and will be able to offer a unique service and QoE.
- **Big sports organizations, music bands** or any kind of group interested in covering an event, who will be interested in offering a service in a differentiative way and will be able to choose this kind of solutions.
- **Remote events attendees**, who would be interested in reproducing events from home for example, but without losing any immersivity.
- **On-site events attendees**, who could be interested in replaying a short play lapse from a specific angle while sitting at some stadium.
- **Telco industry companies**, who will have a new tailored service to offer in their portfolio (in terms of cloud computing solutions, assured QoS network slice, 5G infrastructure and other technical features).

5.1.2 Scenarios and workflow

From a high-level perspective, the use case provides the content producer with a virtual video camera, which can merge into the source of several high-resolution physical cameras to produce a virtual view. The video pipeline will be based on three main stages:

- **Capture:** A set of high-resolution cameras will partially surround the object of interest, and synchronization will be provided by the 5G network. Depth information will be extracted in real time and jointly delivered with the acquired videos. A scene-adaptive segmentation will be carried out to help layered synthesis and bandwidth savings. Fast camera calibration procedures will enable rapid deployment of the system.
- **Encoding and transmission:** Real-time encoding of the camera streams (colour and depth) will require specific processing and lossless delivery for depth information as it represents the scene structure. Thus, extended pixel precision and adapted 4:2:0 video structures will be used for this depth information.
- **Synthesis and visualization:** A live FVV stream will be generated with minimum latency by considering only the reference cameras that are closest to the virtual viewpoint at any given moment. This real-time synthesis is a very complex task that operates on unreliable depth data. Specifically designed soft transitions mix contributions from the closest reference cameras. Dense off-line background model and layered synthesis will be employed.

The use case scenario is shown in the following Figure 23. It is split in several locations:

i. **Smart venue**. A stadium, theatre, festival arena or similar location, where a deployment of 5G mmWave RAN makes it possible to provide service to: *i)* a FVV system, *ii)* the content produce, who operates the virtual camera offered by the FVV platform, and *iii)* a number of end users, which can access the live production of the event as part of their entertainment experience.
ii. **Near Edge**. It hosts the VNFs required to process the FVV streams and generate the virtual views which are produced.
iii. **Cloud Edge** and beyond. It hosts the VNFs required to deliver the produced streams to end users and other third parties.
iv. **Remote locations:** where users different from those located in the Smart venue will access the live production of the event either as end-consumers of the content or as remote producers that will include the media flow into an Edited Content Creation.

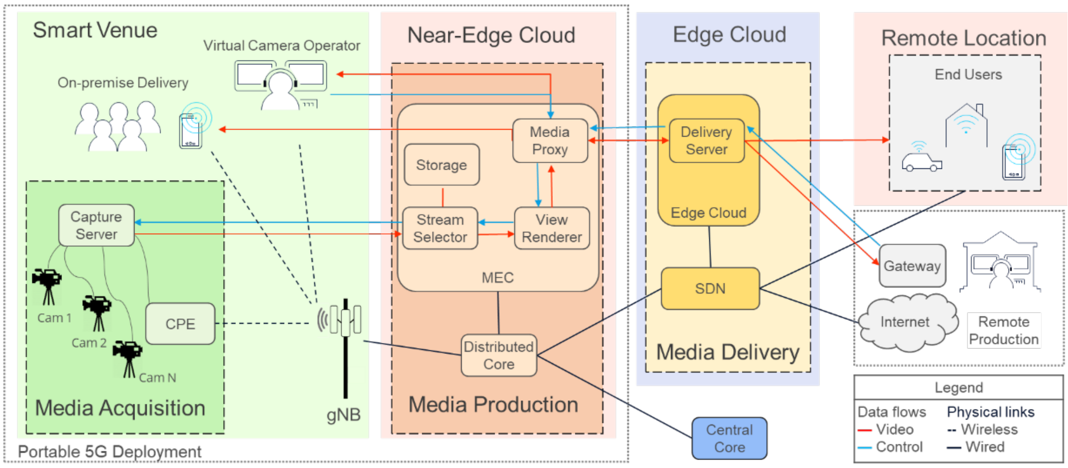

Figure 23. Use case 3 high-level architecture.

The use case deployment considers three main subsystems:

i. **Media acquisition:** A set of media acquisition sub-units each one connected to the 5G networks and endowed with a local video processor to manage the output streams of a plurality of high-resolution cameras. All devices (cameras and video processors) will be remotely controlled through the 5G connectivity and there will be a RGB + Depth stream per camera. This process requires very high UL contributions, transport, storage, and processing.

ii. **Media production:** An FVV renderer and encoder for live or on-demand requests, where the operation will be managed through several Virtual Production Consoles (local or remote). For every generated video stream, either a dynamic selection of the virtual viewpoint or a pre-decided trajectory can be considered and, thus, the required camera streams will be selected. It demands lot of GPU virtualized resources, storage, and processing capabilities.

iii. **Media delivery:** Delivery to third parties (content producers, broadcasters) in contribution quality and delivery to event attendees in streaming quality (transcoding & caching and/or multicast). It requires high processing for the video streaming and lot of Transport and Down Link contribution in the private area and also connectivity or federation to other public 5G networks for external delivery.

Besides, the use case is based in two horizontal elements: (i) **flexible deployment**, and (ii) **end-to-end- QoE guarantee based on network slicing**.

The analysis of the use case will cover several scenarios, from which some of them will be selected for deployment and trial. The scenarios will be broken down from three perspectives: content production, content delivery, and system deployment.

5.1.2.1 Content production scenarios

Two possible scenarios are envisaged for this use case from the media production perspective, that we will refer to as “simple” and “advanced”, with the idea of first being able to demonstrate the simple scenario and progress towards the other as the project goes on.

In a first simple scenario, a small real scene is to be captured with a few cameras that will be placed at relatively short range, with few people being featured in the scene to minimize the number of necessary cameras and associated network requirements. The reference cameras will be physically attached to several dedicated physical computers that will carry out the depth estimation, foreground segmentation and video coding tasks and send the resulting video + depth reference streams via 5G link to another physical dedicated computer, where an operator will be able to select the virtual viewpoint and where the virtual view will be synthesized and coded to deliver to end users, which will not be able to change the viewpoint. The existence of a single virtual viewpoint enables the system to save network resources by only sending the reference streams from the physical cameras closest to the selected virtual viewpoint.

In a more advanced scenario, more network bandwidth is expected to be available, as well as GPU virtualization capabilities from a MEC. This should allow to operate more reference cameras simultaneously, thus covering a bigger real scene and allowing more objects to be present therein or enabling a greater range of movement for the virtual viewpoint. Aside from using more cameras, in an advanced scenario the reference streams would be delivered to the 5G network, which would then route them to one or more virtualized nodes, each tasked with the synthesis of one virtual view, thus supporting multiple independent users, which would select their desired viewpoint remotely and receive the resulting stream. The 5G network could also store the reference views for later use (off-line).

Table 11. UC3 content production scenarios

Scenario		Simple	Advanced
Real scene	Lighting	Controlled (indoor)	
	Complexity	**Low: few** objects and occlusions, **short** depth range	**High: many** objects and occlusions, **wide** depth range
Virtual viewpoint characteristics	Selection	By users, in real time	
	Range (angle)	**Narrow**	**Wide**
	# DOFs (example)	**1** (within arc defined by reference cameras)	**2** (also forward and backward)
Network requirements	Bandwidth	**High**	**Very high**
	Latency	Low	Low

5.1.2.2 Content delivery scenarios

In the Media Delivery subsystem, it is critical to maintain several network requirements that ensure QoE both for the traffic ingestion and for the traffic delivery.

As a strategy to reach those network requirements we will explore end to end network slicing in the context of content delivery for live immersive media productions. The delivery server will be in Madrid, Spain (at Peñuelas Central Office), whereas the users will be distributed among Segovia, Madrid, Barcelona and Valladolid. Those users will have different network requirements based on their profile:

- Local Producers.
- End-consumers attending locally the event.
- Premium remote end-consumers.
- Regular (non-Premium) remote end user-consumers.
- If a Remote Producer is needed a special type of user for remote production tasks will be also considered.

The related KPIs for each of them are presented in Section 5.4 and are subject to further definition during the project execution. To meet the different network demands, we will deploy and test performance on two different network slices:

- **Slice Best-effort**: traffic will not be prioritized on any part of the network.
- **Slice Multimedia-Gold**: traffic will be prioritized in as many parts of the network as possible (access, transport, and core).

The following figure shows the relationship between those network slices and each of the traffic flows:

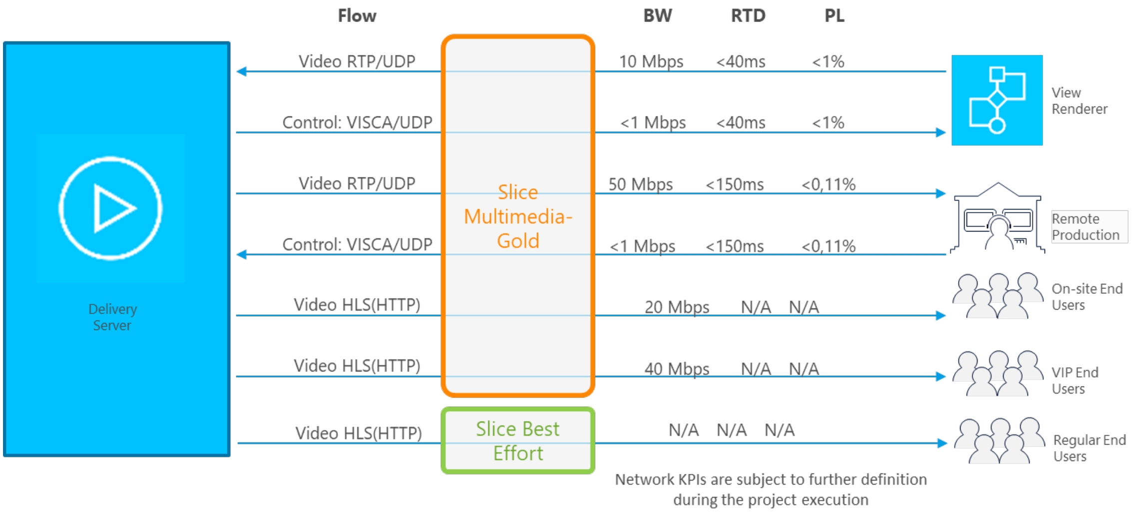

Figure 24. UC3 delivery flows and slices

Finally, some of those users will be connected via a commercial Fiber to the home (FTTH) access (and WiFi-5 Access Points) while other will use the trial 5G network of the project.

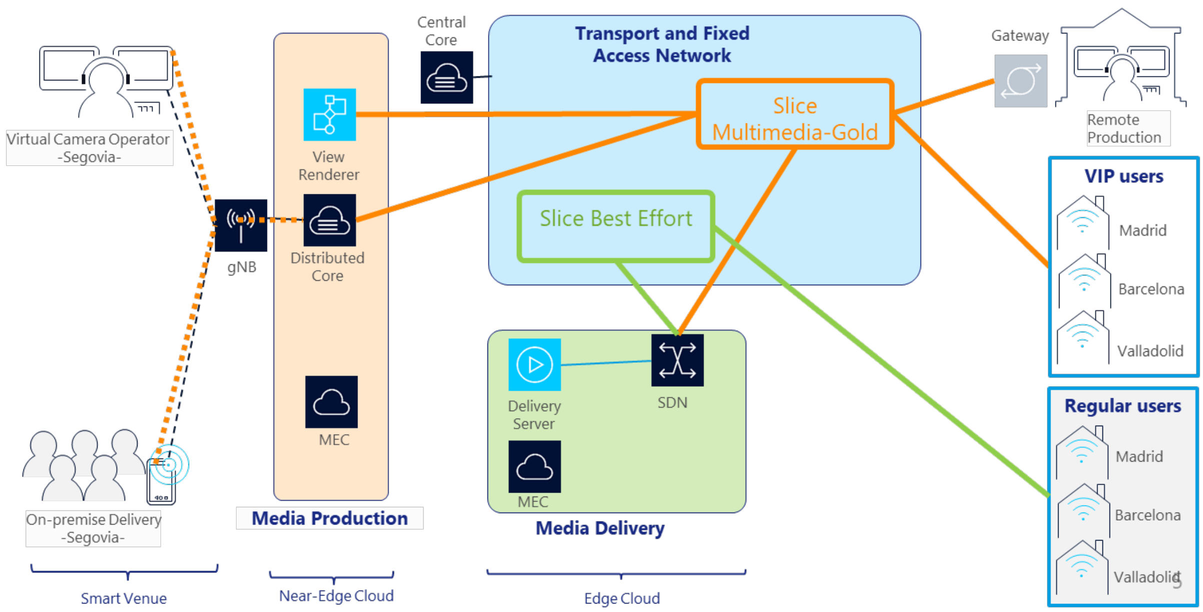

Figure 25. UC3 network slices.

Based on this set-up, we will work on three different scenarios:

Table 12. UC3 content delivery scenarios

Scenario	Slice multimedia-gold E2E for:	Slice Multimedia-Gold partial (near edge to edge cloud) for:
1	View Renderer Premium end-user (FTTH)	Local Producer (5G mmW) On premise end user (5G mmW)
2	View Renderer Premium end-user (FTTH) Local Producer (5G mmW) On premise end user (5G mmW)	Not applicable
3	View Renderer Premium end-user (FTTH) Local Producer (5G mmW) On premise end user (5G mmW) Remote producers (FTTH)*	Remote producers (FTTH)*

* To Be Defined

Regarding delivery protocols. Input is received (from the Media Renderer) H.264 / RTP / UDP, one stream per channel. Each flow goes on a port (and a range of UDP ports should be reserved), filter by source address or, if necessary, by SSRC. At the output, there are two flows. Before the content producers, it would continue to be RTP / UDP. To the end users, HTTP adaptive streaming (HLS or MPEG-DASH) would be delivered. Usually, TCP port can be 80 or 8000. It is also possible to inject an IPTV channel (MPEG2-TS / RTP / UDP / IP multicast).

SDN functionality will be able to establish and manage connections between the Near cloud, Central cloud instances, and external connections to the users. Also, SDN manages connectivity prioritization between the media delivery and the end users. Managing different types of QoS allow producers to evaluate impact in the service.

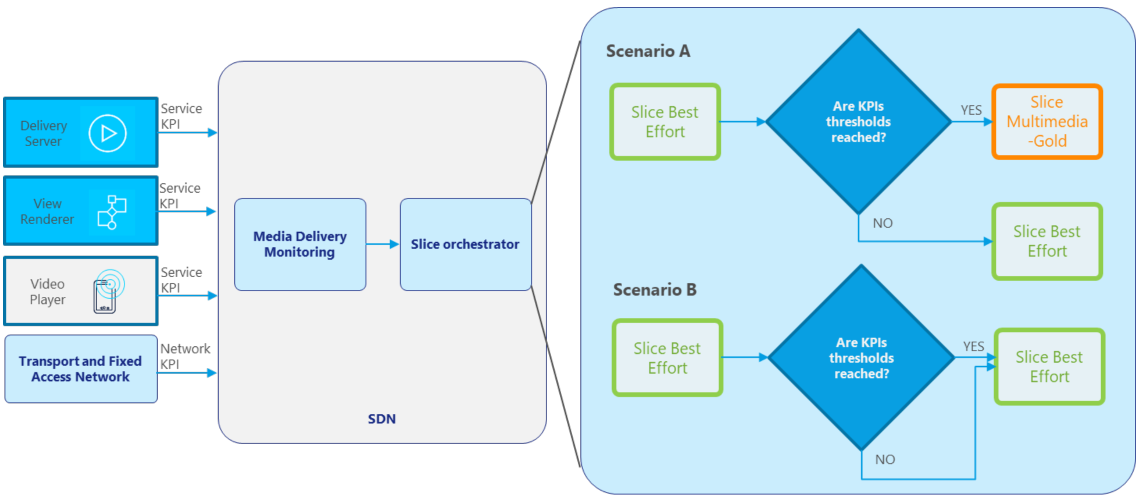

Figure 26. UC3 slice orchestration in the SDN.

5.1.2.3 Deployment scenarios

From the deployment perspective, the target of the use case is creating an end-to-end solution for advanced media production and delivery, which can be applied to both NPN and MNO deployments in any-size Smart Venues, either permanent or ad-hoc, where high capacity is demanded (e. g. concerts, festivals, sport stadiums, etc.). The concept of smart venue implies providing 5G connectivity to sport and entertainment venues, leading into better production capacities (e.g., to support live immersive media production) and better fan experience (e.g., to access free-viewpoint video scenes from the phone during the event). In this context, three deployment sizes are considered: 5G festival, 5G theatre and 5G stadium.

The **5G Festival** describes the deployment of ad-hoc infrastructure to cover a specific event in a location with no previous high-capacity coverage. RAN deployment configuration and network slicing are used to provide the possible service, over SA. An example could be a one-week music festival, or an international sport championship taking place in a secondary city: one-off events in places which would not justify the deployment of permanent 5G mmW infrastructure. Some characteristics of this scenario are:

- Priority to production but support for other services.
- Single or multiple content producer.
- Moderate added-value services to users.

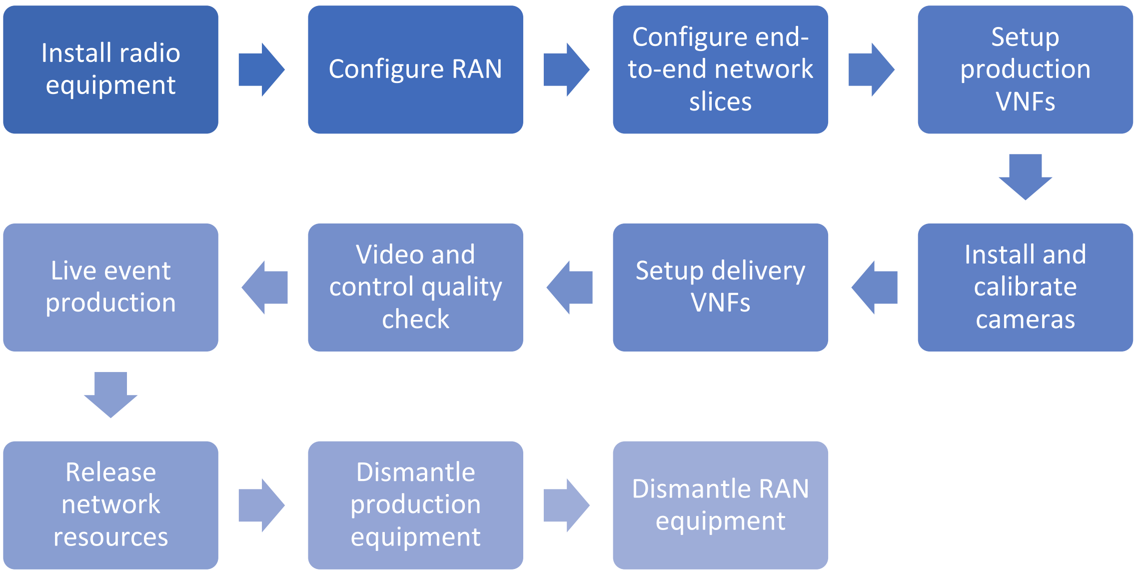

Figure 27. 5G Festival workflow UC3.

The **5G Theatre** models the deployment where an existing 5G network is configured to support the production of a specific event. Network slicing is used to guarantee uplink and MEC capacity over existing infrastructure, which may be either SA or, in a first phase, NSA. An example could be a theatre in an urban area with mmWave coverage, where there is a specific event (concert, theatre play, fashion show) with limited local audience, but global outreach. Some characteristics of this scenario are:

- Focus on prioritizing production service over anything else.
- Single content producer.
- Limited added-value services to users.

The **5G Stadium** models the configuration of an already existing high-capacity 5G network in a tier-1 smart venue (e.g., a football stadium from a team which plays UEFA Champions League). Network slicing is used to separate highly demanding services. This is a use case which only be considered from a theoretical perspective in 5G-RECORDS but, due to its size, its implementation is out of the scope of the project. Some characteristics of this scenario are:

- Multiple slices for different services.
- Multiple content producer.
- Complex added-value services to users.

The following workflow is the same for both 5G Theatre and 5G Stadium scenarios:

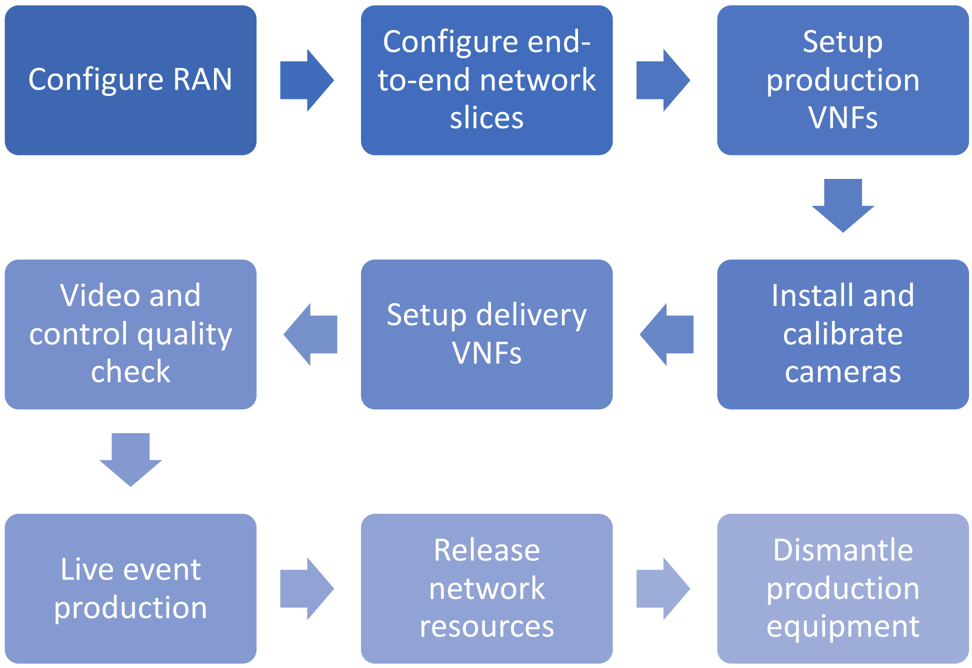

Figure 28. 5G Theatre and 5G Stadium workflow UC3.

The following Table 13 shows a summary of the scenarios.

Table 13. UC3 deployment scenarios.

Scenario	5G	Infra	Slicing	Local user services	Size	Trial target
5G Theatre	NSA	Existing	Yes	Limited	S-M	Phase 1
5G Festival	SA	New	Yes	Moderate	M-L	Phase 2
5G Stadium	SA	Existing	Yes	Demanding	L	Out of scope

5.2 Pre-Conditions

To identify this use case pre-conditions, similarly to [3], the immersive environment characteristics must be considered. To achieve communication in production systems for producing real-time immersive media, it would require different system deployments depending on the type of the event.

When covering sports events, pre-conditions will slightly vary depending on the sport. There is no similarity in the size of the area to be covered, nor in the speed of the players, acceleration, and number of cameras. For that reason, a specific setup will be configured for each specific sport.

This use case is also considering other kind of events such as live music performances, which will conform a different setup compared to a sports event. Camera virtual views will move differently than in a basketball match for example, and in a reduced space which is the stage.

General pre-conditions for a generic event are:

- Professional FVV capture system with cameras, processing units, and UE capabilities: the cameras are registered into the 5G network, connected with the media production system, and perform UE configuration updates.
- Available venue for tests and initial deployment.
- 5G mmW infrastructure deployed in the event venture.
- Near cloud computing system with graphic processing capabilities connected to the on-site equipment to process and deliver the media production.
- Cloud storage system to enable on-demand service.
- Cloud computing system with high networking capabilities for media delivery.

5.3 Use case requirements

This section describes the common requirements for the use case.

Media Acquisition

- A set of media acquisition sub-units, each one connected to the 5G networks and endowed with a local video processor, will be used to manage the output streams of a plurality of high-resolution cameras.
- Each individual camera will be remotely controlled from the VNFs in the MEC.
- The Cameras cluster shall establish a mutually authenticated and secure communication session with the MEC.
- Service continuity shall be guaranteed during the recording service.
- The communication link shall be reliable and potentially encrypted.
- Real-time response and high reliability of the communication session shall be maintained during the recording time.
- The network shall be able to send notifications about expected QoS changes (i.e., QoS prediction) to remote operator, if such notifications are available.
- Bandwidth for Uplink for each UE should be about 150 Mbps.
- Bandwidth for Downlink for each UE should be about 10 Mbps.
- The network must provide a latency between nodes low and constant enough to support clock synchronization with better than 1 ms accuracy.

Media Production

- The media production VNF shall render an encode a FVV stream, working as a virtual camera.
- For every generated video stream, either dynamic selection of the viewpoint or pre-defined trajectories may be used.
- Selection of the viewpoint will be done remotely by a camera control operator using open protocols.
- The camera control shall establish a mutually authenticated and secure communication session with the MEC.
- Recording of the content for off-line view generation shall also be supported.
- Several simultaneous renders shall be supported. The system should control that each required camera stream is only sent once via the radio access network.

Media Delivery

- Each camera render shall be sent to third parties (content producers, broadcasters) in contribution quality.
- Each camera render shall be delivered to end users in streaming quality (transcoding & caching and/or multicast).
- Final users may be in different physical distant places, including the smart venue.

Flexible deployment

- It shall be possible do deploy and remove the whole immersive media production system for one-off events.
- Cameras shall support on-site calibration.
- Different deployment sizes shall be supported. System adaptation to the deployment size should be done via network and software configuration.

End-to-end slicing

- Camera uplink traffic shall be prioritized to guarantee the availability of streams for virtual view generation.
- Slicing of MEC resources shall allow several parallel view renderer VNFs.
- Content delivery will support at least two types of users, Premium user and Non-Premium user.
- Additionally, if remote production is needed, a special type of user will be considered for those tasks.

5.4 KPI

Content Production KPIs

The FVV system proposed in this use case will leverage many resources from the 5G network, which will allow for unprecedented flexibility in the deployment and exploitation of the system, but ultimately the performance and quality of the end application must be measured to ensure meaningfulness of the use case. The next Table 14 summarizes the KPIs that have been considered to evaluate the success of the use case from the perspective of the free viewpoint operation.

Table 14. UC3 content production KPIs

KPI		Scenario	
		Simple	Advanced
Virtual viewpoint	# DOFs	1	2
	# ref. cameras	9	9
	Range (angle)	90º	90º
3D scene reconstruction and virtual view quality	RGB+Depth resolution	720p @ 15 fps	
	Virtual view resolution	720p @ 15 fps	
	Subjective QoE	3.5 in 1-5 MOS/DMOS scale*	
Network requirements	Bitrate (per ref. camera)	100 Mb/s	50 Mb/s
	Motion-to-photon latency	170 ms	
Production functionality	# Virtual cameras	1	2
	Live production	Yes	Yes
	Recorded production	No	Yes

Content Delivery KPIs

Table 15 summarizes the KPIs related to the content delivery scenario. Content Delivery KPIs are focused on guaranteeing enough transport capacity for individual users, even under heavy background load, and ensuring end-to-end QoS management.

Table 15.UC3 content delivery KPIs.

User	Flow	Protocol	Throughput	Jitter	Latency	Packet Loss
Local Producer	Control: Delivery Server -> View Renderer	RTP	10 Mbps	N/A*	<40ms	<1%
	Video: Delivery Server -> View Renderer					
Remote Producer	Control: Delivery Server -> Gateway (or up to end user)	RTP	50 Mbps	N/A*	<150ms	<0.1%
	Video: Delivery Server -> Gateway (or up to end user)					
End-consumers attending locally	Video: Delivery Server -> end user	HTTP	20 Mbps / 10 Mbps**	N/A*	N/A*	N/A*

Premium remote end-consumers	Video: Delivery Server -> end user	HTTP	40 Mbps / 20 Mbps**	N/A*	N/A*	N/A*
Regular (non-Premium) remote end user-consumers	Video: Delivery Server -> end user	HTTP	Best effort	N/A*	N/A*	N/A*

* Not Applicable
** Peak/Mean

Deployment KPIs

Table 16 presents the deployment KPIs within UC3. Uplink/downlink throughputs are achieved through 5G mmWave RAN.

Table 16. UC3 deployment KPIs

KPI		Scenario	
		5G Theater	5G Festival
Infrastructure	*Portable* camera system	Yes	Yes
	Portable RAN	No	Yes
	5G mode	NSA	NSA**
Deployment size	# Capture UEs	3	3
	# Delivery UEs	1	100*
	# MEC units	1	2
Network capacity	UL throughput	180 Mbps	400 Mbps
	DL throughput	350 Mbps	1 Gbps
	RTT (UE to MEC)	40ms	40ms

*Emulated
** Ideally, 5G Festival scenario should be validated using 5G standalone (SA). Due to the foreseen unavailability of standalone deployments with mmWave radio, NSA will be used instead.

A Annex A: media-related 5GPPP projects

Several efforts have been made to explore the new capabilities and possibilities that 5G brings to the table. To provide context, this annex provides some 5GPPP projects related in one way or another to media applications.

- 5GCity [22] presented 3 UCs related to video production, distribution, and immersive services that benefit from a proposed open platform. This platform extends the centralized cloud model to the extreme edge of the network.
- 5G-MEDIA [23] presented 3 UCs related to immersive media and virtual reality, media production and UHD content delivery.
- 5G-MoNArch [24] implemented a testbed to address a set of media and entertainment use cases. It showed the applicability and functioning of E2E network slicing focusing on features including 5G new radio interface, control and user plane separation and orchestration.
- 5GTANGO [25] developed and validated 3 pilots. One of them showcased how 5G networks can enhance the experience of end users regarding media services by providing, high-quality live stream video in 360º environments and additional content, like social media channels to the live video streams.
- 5G-Xcast [26] devised, assessed, and demonstrated a conceptually novel and forward-looking 5G network architecture for large scale immersive media delivery. It enabled broadcast and multicast point to multipoint (PTM) capabilities for 5G.
- Matilda [27] proposed one UC that proposed to combine the functionalities of two of the developed systems, to provide 5G Personal Assistance in Crowded Events (5GPACE), by offering end-users immersive media services combined with machine learning-based personal retail recommendations.
- 5G-SOLUTIONS [28] aims to prove and validate that 5G technologies will provide a wide range of services in contrast to 4G. One of the trials proposed for this project is to demonstrate a media and entertainment UC in Greece and Norway.
- 5G-TOURS [29] is currently deploying full end-to-end infrastructures to bring 5G to real users in 13 representative use cases. One of the use cases will focus on media broadcast and production, located in Turin.
- 5G-VICTORI [30] is conducting large scale trials for advanced vertical use case verification focusing on transportation, energy, media content delivery and factories of the future and cross-vertical use cases.
- FUDGE-5G [31] aims to create a new generation 5G core network for private 5G networks. One of the UCs has the objective of utilizing a localized 5GLAN among many mobile users to facilitate the use of 5G multicast for opportunistic multimedia delivery when transmitting popular content.
- 5GMediaHUB [32] is about testing and validation of innovative 5G-empowered media applications and NetApps from 3rd party experimenters and NetApps developers, through an open, integrated and fully featured experimentation facility.
- 5G Virtuosa [33] targets to explore through real-life examples how 5G wireless communication can be combined with virtualization concepts from the IT industry to enable broadcasters to produce live content (such as sports or music coverage) more efficiently and cost-effectively across locations, to meet growing consumer demand.

B Annex B: technology readiness level (TRL)

5G-RECORDS will significantly improve the 5G technology readiness levels of the three use cases: live audio production, multiple camera wireless studio and live immersive media service. Following the work that is anticipated to be carried out within these use cases and the traction that will be received to this end, the TRL of the various components is expected to increase significantly, which broadens the potential towards the wider and faster deployment of 5G-enabled technologies. The TRLs within each use case are as follows:

Table 17. Influence on the technology readiness levels (TRL) by the 5G RECORDS project within UC1.

Initial TRL	Target TRL	Justification
2	3	Analysis of UC1 requirements and translation to 5G KPIs. Concept development of the 5G architecture.
3	4	Development of first 5G test platforms and corresponding audio platforms to evaluate relevant 5G parameters such as latency, reliability and synchronicity.
		Measurement campaign to verify the achievable performance quality.
4	5	Integration of audio platforms into the 5G infrastructure to provide specific UEs for AV production. Building up a local network consisting of latency optimized RAN and core network, edge computing functionalities and AV devices.
		Measurement campaign to validate the performance in the context of UC1.
		System design and development of an audio production NPN, composed of UEs acting as wireless microphones and IEMs as well as latency-optimized RAN, core network and edge computing.
		Testing of the NPN for audio production in the lab.
5	6	Testing of audio production NPN in a 5G relevant environment.
		Demonstration of the use case using an NPN in a 5G relevant environment.
6	7	Testing and demonstration of audio production NPNs in 5G real scenarios.

Table 18. Influence on the technology readiness levels (TRL) by the 5G RECORDS project within UC2.

Initial TRL	Target TRL	Justification
2	3	Analysis of specific KPIs and requirements.
		First experimental proof of concept in a laboratory environment.
3	4	Development of first 5G test platforms to evaluate relevant 5G parameters such as throughput, latency, and reliability.
		Technology validation in the laboratory.
4	5	Adapting and integrating components to the specific needs of a multiple camera wireless studio and remote production deployments.
		Technology validation in the context of UC2.
5	6	Use case demonstrated in a controlled environment relevant to the media industry. 5G network functions required for network slicing and edge computing, i.e., Network Repository Function (NRF) and Network Slice Selection Function (NSSF).
6	7	Testing and demonstration of UC2 in a 5G real scenario

Table 19. Influence on the technology readiness levels (TRL) by the 5G RECORDS project within UC3.

Initial TRL	Target TRL	Justification
2	3	Incorporation of 5G into the free-viewpoint video pipeline to allow the distribution of the computational load.
3	4	Technology adaptations of the free viewpoint video system to operate within a 5G environment to interconnect subsystems, to synchronize all involved elements, and to allow remote control and operation.
		Technology upgrades of the free viewpoint video system to deal with high-resolution cameras, to achieve fast calibration procedures, and to include optimized approaches to background and foreground modelling,
		Validation of the free viewpoint video system in the laboratory.
4	5	Adaptation of the free viewpoint video subsystems to the requirements and characteristics of a 5G relevant environment: camera deployment and calibration, background modelling, and foreground segmentation.
		Validation of the cross-operation of the different subsystems of the live immersive media service.
5	6	Adaptation of the free viewpoint video system to the requirements and characteristics of a 5G relevant environment.
		Demonstration of the live immersive media service in the relevant environment.
6	7	Operation in the considered sports environment under a 5G NPN and demonstration of the live immersive media service.

References

[1] 3GPP, “TS 22.104 Service requirements for cyber-physical control applications in vertical domains. Stage 1. Release 17.,” 2020.

[2] 3GPP, “TS 22.263 Service requirements for video, imaging and audio for professional applications (VIAPA); Stage 1. Release 17,” 2020.

[3] 3GPPP, “TR 22.827 Study on Audio-Visual Service Production. Stage 1. Release 17.,” 2020.

[4] 3GPP, “TS 22.261 Service requirements for the 5G system, Stage 1, Release 17”.

[5] 5G Alliance for Connected Industries and Automation, “www.5g-acia.org,” July 2019. [Online]. Available: https://5g-acia.org/wp-content/uploads/2021/04/WP_5G_NPN_2019_01.pdf.

[6] “How 5G will transform the business of media & entertainment,” Ovum - Intel, [Online]. Available: https://newsroom.intel.com/wp-content/uploads/sites/11/2018/10/ovum%E2%80%93intel%E2%80%935g%E2%80%93ebook.pdf.

[7] M. a. W. Brunner, ““On the audibility of comb filter distortions”,” in *Audio Engineering Society Convention*, 2007, p. 122.

[8] a. B. Lester, ““The Effects on Latency on Live Sound Monitoring”,” in *Audio Engineering Society Convention* , 2007, p. 123.

[9] H. S. a. S. Pilz, ““Professional Live Audio Production: A Highly Synchronized Use Case for 5G URLLC Systems”, IEEE Network March/April 2018,” 2018.

[10] J. B. R. a. Development, “EBU Technical review,” Winter 1998. [Online]. Available: https://tech.ebu.ch/docs/techreview/trev_278-stott.pdf.

[11] Dimensionstudio, “Hammerhead becomes Dimension,” [Online]. Available: https://www.dimensionstudio.co/news/hammerhead-becomes-dimension.

[12] Intel, “Turner Sports and Intel Announce Multiyear Partnership to Bring Virtual Reality to NBA Fans,” 2017. [Online]. Available: https://newsroom.intel.com/news-releases/turner-sports-intel-announce-multiyear-partnership-bring-virtual-reality-nba-fans/ .

[13] Microsoft Japan Co., Ltd, “Press Release,” 21 February 2018. [Online]. Available: https://news.microsoft.com/uploads/2018/02/Media-Release-hakuhodo-VRAR-and-Kennin-ji-Hold-MR-Museum-in-Kyoto.pdf.

[14] Intel, “Intel newsroom,” 14 September 2018. [Online]. Available: https://newsroom.intel.com/news/nfl-intel-outfit-super-bowl-liii-host-mercedes-benz-stadium-intel-true-view/#gs.shqjzd.

[15] Ryo Suenaga, Kazuyoshi Suzuki, Tomoyuki Tezuka, Mehrdad Panahpour Tehrani, Keita Takahashi, and Toshiaki Fujii, “A practical implementation of free viewpoint video system for soccer games,” in Three-Dimensional Image Process., Measurement (3DIPM), and Appl. 2015 vol. 9393, (SPIE, 2015), pp. 93930G.

[16] Houari Sabirin, Qiang Yao, Keisuke Nonaka, Hiroshi Sankoh, and Sei Naito, “Semi-automatic generation of free viewpoint video contents for sport events: Toward real-time delivery of immersive experience,” IEEE MultiMedia 99 (2018), pp. 2..

[17] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L. Davidso…, “Holoportation: Virtual 3D Teleportation in Real-time,” in Proc. 29th Annu. Symp. on User Interface Softw. and Technol. (ACM, 2016), pp. 741--754.

[18] “4D Replay,” [Online]. Available: www.4dreplay.com.

[19] “Intel True View,” [Online]. Available: http://www.intel.com/content/www/us/en/sports/technology/true-view.html.

[20] D. Berjón, P. Carballeira, J. Cabrera, C. Carmona, D. Corregidor, C. Díaz, F. Morán, N. García, “FVV Live: real-time, low-cost free viewpoint video,” IEEE Int. Conf. on Multimedia and Expo Workshops, ICME 2020, London, United Kingdom, pp. 1-2, 6-10 Jul. 2020..

[21] Radiocommunication Sector of ITU, “ITU - R,” 09 2015. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf.

[22] “5GCity,” 5G PPP, [Online]. Available: https://5g-ppp.eu/5g-city/.

[23] “5G-MEDIA,” 5G PPP, [Online]. Available: https://5g-ppp.eu/5g-media/.

[24] “5G-MoNArch,” 5G PPP, [Online]. Available: https://5g-ppp.eu/5g-monarch/.

[25] “5GTANGO,” 5G PPP, [Online]. Available: https://5g-ppp.eu/5gtango/.

[26] “5G-Xcast,” 5G PPP, [Online]. Available: https://5g-ppp.eu/5g-xcast/.

[27] “Matilda,” 5G PPP, [Online]. Available: https://5g-ppp.eu/matilda/.

[28] “5G-SOLUTIONS,” 5G PPP, [Online]. Available: https://5g-ppp.eu/5g-solutions/.

[29] “5G-TOURS,” 5G PPP, [Online]. Available: https://5g-ppp.eu/5g-tours/.

[30] “5G-VICTORI,” 5G PPP, [Online]. Available: https://5g-ppp.eu/5g-victori/.

[31] “FUDGE-5G,” 5G PPP, [Online]. Available: https://5g-ppp.eu/fudge-5g/.

[32] “5GMediaHUB,” [Online]. Available: https://www.5gmediahub.eu/.

[33] “5G Virtuosa,” [Online]. Available: https://5g-virtuosa.eu/.

[34] “Programme Making and Special Events,” European Comission, [Online]. Available: https://ec.europa.eu/digital-single-market/en/programme-making-and-special-events-pmse.