

## RESEARCH ARTICLE

# Techno-economic analysis of 5G immersive media services in cloud-enabled small cell networks: The neutral host business model

## Providing techno-economic guidelines for the successful provision of 5G innovative services in small cell networks

Pietro Paglierani<sup>1</sup>  | Ioannis Neokosmidis<sup>2</sup> | Theodoros Rokkas<sup>2</sup> | Claudio Meani<sup>1</sup> | Karim M. Nasr<sup>3,4</sup> | Klaus Moessner<sup>4</sup> | Pouria Sayyad Khodashenas<sup>5</sup>

<sup>1</sup>Italtel SpA, Settimo Milanese, Italy

<sup>2</sup>inCITES Consulting SARL, Strassen, Luxembourg

<sup>3</sup>Faculty of Engineering & Science, University of Greenwich, Kent, UK

<sup>4</sup>Institute for Communication Systems, University of Surrey, Surrey, UK

<sup>5</sup>i2CAT Foundation, Barcelona, Spain

### Correspondence

Pietro Paglierani, Italtel SpA, Via Aganippo Brocchi, 20019 Settimo Milanese MI, Italy.  
Email: [pietro.paglierani@gmail.com](mailto:pietro.paglierani@gmail.com)

### Funding information

EC H2020 5G-PPP, Grant/Award Number: 671596

### Abstract

Fifth generation (5G) envisages a “hyperconnected society” with an enormous number of interconnected devices, anywhere and at any time. Edge computing plays a pivotal role in this vision, enabling low latency, large traffic volumes, and improved quality of experience. The advent of 5G and edge computing encourages vertical industries to develop innovative services, which can meet the challenging demands coming from consumers. However, economic feasibility is the ultimate factor that determines the viability of a new service. Hence, effective techniques for the economic assessment of such services are needed. This paper analyzes the provision of immersive media services in crowded events, through a cloud-enabled small cell network owned by a neutral host, and offered in multitenancy to different mobile network operators. We initially develop a planning model to predict the required compute, storage, and radio resources. Taking into account dynamic factors such as service penetration and price evolution, we then provide a number of economic indices, such as net present value, internal rate of return, and expected payback period to assess the viability of a potential investment in a 5G infrastructure for immersive media services. The presented analysis will guide small cell network operators in the provision of 5G innovative media services.

## 1 | INTRODUCTION

Forthcoming fifth-generation (5G) networks promise a ubiquitous solution, featuring extremely high data rates and device densities, multitenancy, on-demand service-oriented resource allocation, and automated management.<sup>1</sup> These aspects call for a fundamental change in mobile network infrastructures, which should shift from conventional data transport means to intelligent information technology (IT) entities, equipped with virtualized compute and storage resources.

The trend analysis of data traffic and consumers' behavior can shed some light on the reasons behind this change of paradigm in telecommunication networks. According to Cisco,<sup>2</sup> Ericsson,<sup>3</sup> and Amdocs,<sup>4</sup> end users originate about 80% of the overall mobile data traffic indoor, ie, during concerts or sporting events; standing at airports, train stations, or

commercial malls; at work; in enterprise premises; or at home. Currently, indoor traffic is increasing at a rate of 20% a year, more rapidly than outdoor traffic, and this trend will continue in the next years.

At the same time, consumers' behavior is showing a clear shift from traditional services, such as voice calls or text messaging, toward CPU-and-data-intensive activities like gaming, social networking, or immersive video services based on 4 k, 360-degree, or 3D media formats.<sup>3</sup>

Emerging video-based applications are driving the traffic growth in mobile networks. Video traffic, which now represents 74% of all mobile data traffic, is expected to increase by around 35% per year through 2024.<sup>2,3</sup>

The provision of immersive media services (IMS) in highly crowded areas, eg, during festivals, music concerts, or key sporting events, has recently attracted the attention of network operators and equipment manufacturers.<sup>5,6</sup> In particular, transforming sporadic crowded events (CEs) in immersive experiences, with the possibility of downloading high quality multimedia contents in seconds even under challenging network conditions, has become one of the most attractive 5G use cases.

Many research projects have dedicated significant efforts to these topics. Among these, SESAME (Small cELLS coordination for multitenancy and edge services)<sup>7</sup> has been a research and innovation project, financed within the 5G Public-Private Partnership (5GPPP,<sup>8</sup>) initiative by the European Commission (Horizon 2020 program), with a recognized role in the development of the small cell (SC) concept.<sup>9</sup> SESAME has focused its activities on three central elements:

1. The placement of network intelligence and applications at the network's edge, through the adoption of the edge computing paradigm<sup>10</sup>;
2. The evolution of the SC concept, already mainstream in 4G networks, but expected to deliver its full potential in high density 5G scenarios;
3. The consolidation of multitenancy in communication infrastructures, allowing multiple Mobile Network Operators (MNOs) to share both access capacity and edge computing capabilities offered by an SC network.

Within 5GPPP, other projects are carrying forward the seminal work of SESAME, which completed its activities in December 2107. In particular, 5GESSENCE<sup>11</sup> is developing the "SC-as-a-Service" paradigm in straight continuity with SESAME, whereas 5GCITY<sup>12</sup> is focusing on the combined use of SC and edge cloud computing in the smart city scenario. Moreover, both projects are contributing to the development of new markets and business models to support the successful emergence of SC network operators (SCNOs) as 5G neutral hosts.

The neutral host is the owner of the radio and IT infrastructure deployed in a venue. It can be a municipality or an enterprise, eventually supported by a collaborating service provider, renting local connectivity and IT resources to different telecom operators through multitenancy.<sup>9,13</sup>

Neutral host solutions based on technologies other than SC are available. Distributed antenna systems (DAS) and Wi-Fi networks can increase radio network capacity and coverage in highly crowded areas. Nonetheless, these technologies can present relevant drawbacks for operators and venue owners, such as high costs (DAS) or low quality of service (Wi-Fi).<sup>9</sup> Small cells can successfully address such problems; hence, they have become the most appealing approach to develop effective neutral host solutions in 5G networks.

Some recent studies have analyzed the use of small cells to achieve 5G network densification. In particular, Oughton et al<sup>14,15</sup> provide a techno-economical assessment of the 5G network rollout in the UK and in the Netherlands, respectively, whereas Neokosmidis et al<sup>16</sup> analyze the main technological and socio-economic issues that influence the market adoption of 5G networks, including the use of SCs for network densification.

Conversely, to the best of the authors' knowledge, the techno-economic analysis of the emerging neutral host business model has so far received minor attention. Nonetheless, factors including capital and operating costs (CAPEX/OPEX), as well as revenue generation potentials, ultimately determine the viability of any solution to be deployed. Thus, clear and careful economic studies to justify proper investments in this context are needed.

The presented analysis aims at translating technical capabilities into projected economic figures, thus helping interested parties to understand better the potential of an investment. To this end, the paper adopts a bottom-up approach, providing an accurate techno-economic assessment of the neutral host business model, which focuses on the offer of IMS during a CE. The presented approach, however, can be straightforwardly extended to any other edge cloud architecture and innovative service, following the guidelines developed in this paper.

We can summarize the main innovative contributions of this paper as follows. Using the cloud-enabled small cell (CESC) model proposed by SESAME, this paper presents a novel planning tool, which can accurately predict the compute, storage, and radio resources needed in a venue to provide IMS during a sporadic CE, as a function of the number of

participants and few, specific quality-of-experience (QoE)–related indices. This tool can also be useful to assign resources to the different MNOs sharing the SC network capabilities.

Based on the output of the planning tool, the paper investigates, over a ten-year study period, the dynamic behavior of a number of economic parameters, such as CAPEX/OPEX, time to reach breakeven, and expected payback period to help verify the economic feasibility of a proposed investment. As a realistic example, we analyze the provision of an IMS capable infrastructure in a large stadium with a typical capacity of 90 000 people and an area of 40 000 m<sup>2</sup>. The presented analysis will guide neutral hosts toward a successful deployment of the needed infrastructure to offer innovative 5G services.

This paper structure is as follows: in Section 2, we introduce the overall scenario, ie, we define in details the adopted CESC model and the services provided to the consumers participating to a CE. In Section 3, we describe the planning tool that can predict the required radio and IT resources to guarantee an adequate service provisioning in the venue of interest. In Section 4, we apply the presented planning tool to a simple but realistic real-life example to clarify its usage. In Section 5, we present the details of the techno-economic analysis carried out in this study. Finally, we discuss the example results in Section 6, whereas in Section 7, we provide an outlook for investors and this paper' conclusions.

## 2 | THE STUDY FRAMEWORK

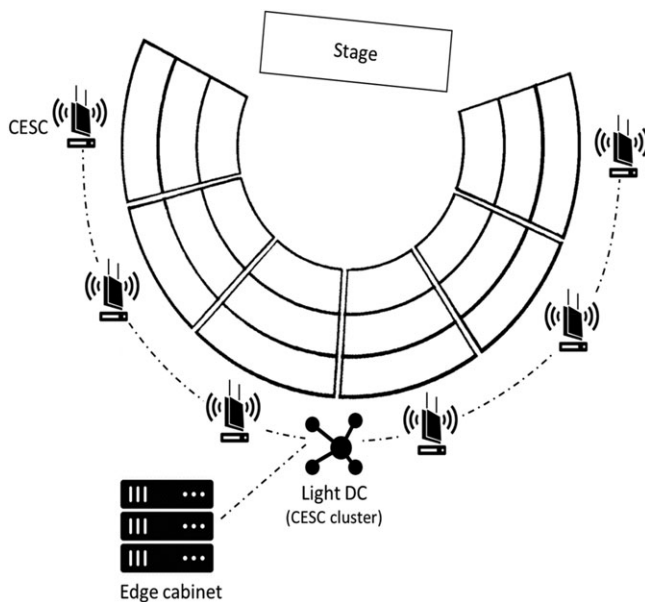
The focus of this techno-economic study is on the IT and radio infrastructure deployed by a neutral host in venues like sport stadiums or concert halls. These venues are mainly characterized by their area and number of attendees. Figure 1 shows a possible venue layout, where a CESC infrastructure has been deployed to offer connectivity and edge cloud computing capabilities. To proceed in the analysis, we present the adopted CESC model, together with an introduction of the services offered to consumers during sporadic CEs.

### 2.1 | The SESAME CESC model

The CESC is an enhanced SC that integrates a virtualised execution platform (microserver) equipped with IT resources (RAM, CPU, storage), as shown in Figure 2.<sup>13</sup> The SC connects to the microserver through a dedicated interface, usually a standard Gigabit Ethernet connection, which links the SC physical network function (PNF) to the SC virtual network functions (VNFs). The combined functionalities offered by the SC PNF and VNFs implement all the protocol stack layers of an evolved Node B.

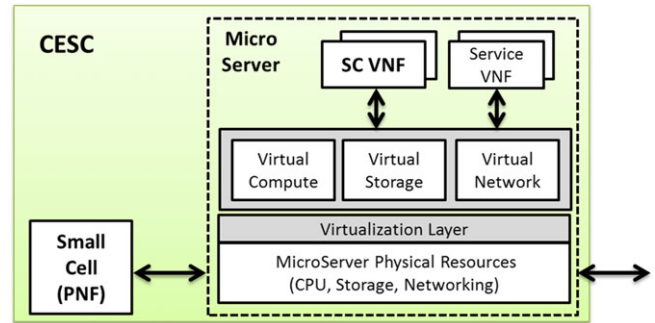
There is currently a debate about the most suitable functional split of the SC protocol stack to define which parts should run as VNFs and which should remain in the PNF.<sup>9</sup> In the SESAME CESC model, the layers above the Packet Data Convergence Protocol (PDCP) run in the microserver as SC VNFs.<sup>17</sup> The remaining layers constitute the SC PNF.

Leveraging network function virtualization,<sup>18,19</sup> the collocated CESC computational and storage capabilities can be used for running additional service VNFs, as shown in Figure 2. However, as the individual CESC resources might

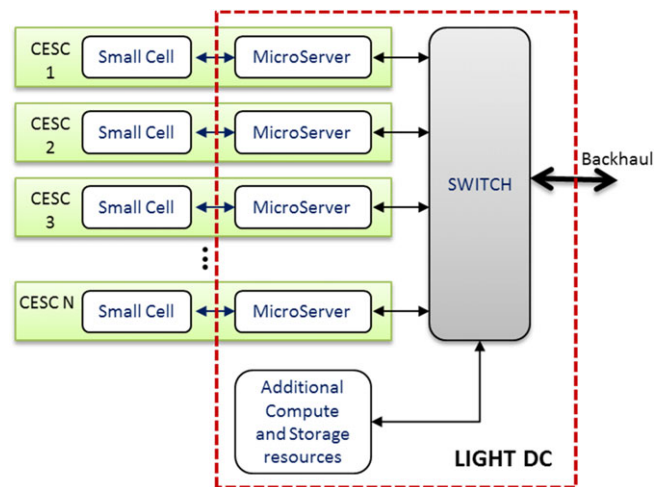


**FIGURE 1** Illustration of the case study framework. CESC, cloud-enabled small cell

**FIGURE 2** Simplified view of the main blocks constituting the cloud-enabled small cell (CESC). SC VNF, small cell virtual network function



**FIGURE 3** Simplified view of the main blocks constituting the Light DC. CESC, cloud-enabled small cell



be insufficient, the SESAME model suggests CESC clustering. This leads to the creation of a distributed virtualization infrastructure, referred to as a Light Data Center (Light DC),<sup>20</sup> shown in Figure 3.

To enable multitenancy in the CESC cluster, two main requirements must be met:

- Adequate radio coverage and capacity at the venue to support blocking free connectivity and achieve the target data-rates required to provide IMS;
- Adequate IT resources at the Light DC to run both the SC VNF, which enables multitenancy and edge computing capabilities<sup>17</sup> and the additional Service VNFs. The SC VNF works as a bridge between the radio and cloud domains; it intercepts an S1 radio bearer per tenant and performs its encapsulation/decapsulation to enable data processing in proximity to end users.<sup>17,21</sup>

To run CPU-and-data-intensive service VNFs, it is essential to enhance the Light DC by coallocating extra IT resources and hardware accelerators, such as graphics processing units (GPUs), which can run the highly demanding workloads originated by IMS. In the example analyzed in this paper, the extra IT resources in the Light DC will provide compute-intensive video processing capabilities during CEs to extend the basic services commonly offered by mobile radio networks.

## 2.2 | The complete service offer

In this study, the basic assumption is that IMS are offered to consumers together on top of basic connectivity, ie, conventional voice services and data connectivity. We will consider three service bundles, with different service level agreements defined as follows:

- The bronze bundle: it includes only basic connectivity.
- The silver bundle: it adds IMS with medium (720p) resolution video contents, in addition to bronze basic connectivity.
- The gold bundle: adds high (1080p) resolution video contents on top of the silver bundle.

The bundle share defines the percentage of users for each bundle, eg, 5% of gold bundle users, 10% of silver bundle users, and 85% of bronze bundle users, and helps modeling the dynamic behavior of end users over time. An evolution in end

users' consumption profile is expected to occur as 5G technology matures. For example, an initial consumption profile of 5% gold, 10% silver, and 85% bronze in a venue can change to 75% gold, 15% silver, and 10% bronze over a duration of 5 years.

### 2.3 | The IMS framework

The IMS included in the silver and gold bundles are provided by a specific service VNF running in the Light DC, namely, the video processing unit (VPU) VNF. The VPU enables the possibility for consumers standing at the venue to share high-resolution video contents, anywhere, at any time and via any device, with (perceived) real-time interaction with the system and among users.<sup>20,21</sup> In addition to conventional content delivery services, in which users only play the role of content consumers, the VPU also offers the possibility to CE participants to create and share video contents in real-time, within a pre-defined group of peers.

Consumers can access IMS from any browser through a web-service interface or through an App running on their personal devices.<sup>20</sup> This way, CE participants can register to IMS and create ad hoc groups of peers with other selected users. Within a group, peers can share live video streams or upload them as prerecorded files. Any other user in the group can receive such live contents, or access them offline later. The VPU saves any video sequence on a local distributed storage area for successive content production and sharing.

Crowded event managers can transmit event related information or advertisements to all users, thus bypassing group constraints. The produced contents must be available as flash event, ie, in the shortest possible time, and adapted to the consumers' device capabilities to optimize the perceived QoE.

To match the different media capabilities of users' devices, the VPU produces copies at different resolutions of each received video file or live stream. The output video files at the desired resolutions are stored in the distributed storage system, ready for immediate streaming on demand. The VPU can also adjust the format of live streams, when users share contents in real-time between devices with different media capabilities. This way, the IMS system can guarantee content accessibility to any registered device in a group, with minimum delay.

Transcoding of video files and live streams is by far the most compute intensive workload for the VPU. Other offered functionalities, such as audio processing or instant messaging, are significantly less compute intensive. Thus, in the following, we will mainly focus on video transcoding.

## 3 | THE RESOURCE PLANNING TOOL

The description of both the SESAME CESC and the offered services allows developing a planning tool to predict the overall IT and radio resources that the neutral host must deploy at the venue. In particular, by suitably modeling the VPU, we can estimate the additional compute and storage resources needed by the Light DC. However, the VPU offline and real-time online transcoding capabilities require ad hoc models, which we will analyze separately.

### 3.1 | Resource planning for video content offline transcoding

Offline transcoding applies to video contents that users produce and send to the VPU as data files. In this case, we consider a subset of  $N$  silver or gold participants to the CE. We model such content originating users as  $N$  independent Poisson processes, with rate  $\lambda_0$  file per second (Fps). Consequently, also the superposition of video-file arrivals at the VPU results in a Poisson process, with rate  $\lambda = N\lambda_0$ .<sup>22</sup> For the sake of simplicity, we will assume that all video sequences are at the same frame rate, with mean duration (expressed in video frames) equal to  $N_{\text{FILE}}$  frames.

Once the VPU receives an input video-data file, it immediately starts a multiple transcoding process, consisting of  $Q$  (parallel or sequential) single-transcoding sessions. The VPU consequently outputs  $Q$  compressed video sequences at the desired resolutions and formats, ready for on demand transmission to the users.

To proceed with a simple but realistic example, we assume that only the ITU-T H.264 standard<sup>22,23</sup> is used to compress the output files, with  $Q = 2$  target video resolutions, namely, 720p and 1080p, which correspond to medium and high definition (HD) qualities, respectively.

A single-transcoding session consists of three stages. At first, the input video file is decoded from its original format, which depends on end users' preferences or device capabilities (this stage is in common among the single-transcoding processes, being performed only once for any received file); then, its resolution is adjusted to the target resolution; finally, it is compressed to the target compression scheme.

The most intensive task in this procedure is by far the final encoding process, whereas the first two stages typically present a much lower computation complexity.<sup>20,24</sup> Therefore, we can characterize each single-transcoding session by one parameter, ie, the transcoding rate, expressed in frames per second (fps), independently of the input sequence format. The transcoding rate assumes a different value for each target resolution. With  $Q = 2$ , we will indicate such values as  $\mu_0$  and  $\mu_1$ , which represent the transcoding rates of the medium- and high- contents, respectively.

If the VPU runs as a VNF on x86-architecture CPUs,  $\mu_0$  and  $\mu_1$  mainly depend on the number of virtual cores assigned to the considered VPU instance. In this case, one can evaluate  $\mu_0$  and  $\mu_1$  by few properly designed experiments, as discussed in the work of Paglierani et al.<sup>20</sup> If video processing relies on HW acceleration devices, such as GPUs,  $\mu_0$  and  $\mu_1$  can still be obtained experimentally; alternatively, the device manufacturer can provide them as a priori performance parameters.<sup>20</sup>

The VPU performance, however, depends on the overall service rate  $\mu$  of the multiple transcoding process, which is a function of  $\mu_0$  and  $\mu_1$ . To evaluate  $\mu$ , we need to elaborate further the VPU statistical model presented so far.

### 3.1.1 | Estimating the number of number of VPU instances

It is convenient to describe the VPU as an M/G/1 queueing system, with Egalitarian Processor Sharing (EPS) discipline.<sup>25</sup> This is a simple and common approach to model computing systems and relies on the reasonable assumption that the VPU equally shares its resources among the active single-transcoding sessions.\*

To obtain  $\mu$ , we observe that, under the EPS discipline, the time needed by two single transcoding sessions running in parallel to process a given input file is equal to the time needed by the same sessions applied in a strictly sequential order. We provide a formal prove of this fact, generalized to the case of  $Q$  single transcoding processes, in Appendix. For  $Q = 2$ , we can easily obtain the overall service rate in fps as

$$\mu = (1/\mu_0 + 1/\mu_1)^{-1} = \mu_0\mu_1/(\mu_0 + \mu_1). \quad (1)$$

Since the VPU processes video data files, it is convenient to express its service rate in Fps, rather than in fps. The overall service rate  $\mu_F$  in Fps is

$$\mu_F = \mu/N_{\text{FILE}}. \quad (2)$$

We indicate with  $D_0 = 1/\mu_F$  the time needed by the VPU to transcode one single video-data file of size  $N_{\text{FILE}}$ , when this is the only file present in the system. Conversely, we let  $D$  represent the time-varying delay that an in-service VPU introduces when processing additional files, randomly sent by the users. We can evaluate the average VPU delay, ie, the average time between the arrival of a video file and the availability of the  $Q = 2$  output files at medium and high resolutions through the following expression, valid for an M/G/1-EPS system

$$E[D] = 1/(\mu_F - \lambda) = D_0/(1 - \rho), \quad (3)$$

where  $E[.]$  is the expectation operator, while  $\rho = \lambda/\mu_F$  is usually referred to as CPU usage.<sup>25</sup> To maintain system stability,  $\rho$  must be strictly lower than unity. Moreover, (3) clearly shows that  $E[D]$  increases as  $\rho$  approaches unity. To guarantee to customers an adequate QoE,  $D$  should be as low as possible. Thus, we set a constraint on its average, ie,  $E[D] \leq D_E$ , where  $D_E$  is the desired average VPU delay. Using the relation  $\lambda = N\lambda_0$  in (3), the maximum number  $N$  of users served by one instance of the VPU so that  $E[D] \leq D_E$  can be evaluated as

$$N(D_E, \mu_F, \lambda_0) \leq (D_E\mu_F - 1)/(D_E\lambda_0), \quad (4)$$

where the expression  $N(D_E, \mu_F, \lambda_0)$  explicits the dependence of  $N$  on  $D_E$ ,  $\mu_F$ , and  $\lambda_0$ . From the total number  $U_{\text{TOTAL}}$  of expected users originating video contents, one can obtain the corresponding total number of needed VPU instances as

$$M_{\text{VPUOFF}} = U_{\text{TOTAL}}/N(D_E, \mu_F, \lambda_0). \quad (5)$$

\*The validity of the EPS model has been investigated through a number of laboratory experiments (some of which were carried out during the SESAME project) summarized and discussed in other works.<sup>20,24,26</sup> The interested reader can specifically refer to the work of Paglierani et al<sup>20</sup>, where the behavior of the VPU has been thoroughly characterized when running on an x86-architecture, with and without Hardware acceleration, and on an ARM processor. In particular, the results in section VII in the work of Paglierani et al<sup>20</sup> (eg, see Figures 7 and 8) show how the aggregated performance of the VPU in terms of processed frame per second remains with good approximation constant, equally sharing the processing resources among the varying number of active transcoding sessions.

Moreover, one can use (5) with different values of  $(D_E, \mu_F, \lambda_0)$ , if the differentiated models are needed for gold and silver users. In this case, the minimum number  $M_{VPUOFF}$  of VPU instances is

$$M_{VPUOFF} = \text{round} \left[ U_G/N_G (D_{EG}, \mu_{FG}, \lambda_{0G}) + U_S/N_S (D_{ES}, \mu_{FS}, \lambda_{0S}) \right], \quad (6)$$

where  $U_S$  and  $U_G$  are the number of expected silver and gold users, respectively, whereas  $N_G (D_{EG}, \mu_{FG}, \lambda_{0G})$  and  $N_S (D_{ES}, \mu_{FS}, \lambda_{0S})$  are given by (4).

Having determined  $M_{VPUOFF}$ , the next step in the development of the planning tool is calculating the amount of RAM that each single VPU instance requires.

### 3.1.2 | Estimating the amount of RAM per VPU instance

From the properties of EPS systems, the steady-state probability that a VPU instance processes  $i$  video files concurrently is<sup>25</sup>

$$\pi_i = \rho^i (1 - \rho). \quad (7)$$

Setting the probability  $\pi_i$  to an arbitrarily low value, eg,  $10e-5$ , we can obtain the corresponding number of concurrently processed files as

$$i = \lceil [\log(\pi_i) - \log(1 - \rho)] / \log(\rho) \rceil. \quad (8)$$

In the experiments carried out in this research, it has been experimentally verified that one instance of the VPU, with no active transcoding sessions, requires approximately 6 GB of RAM, and each multiple transcoding session requires additional 0.5 GB of RAM.<sup>20</sup> Thus, the minimum quantity of RAM required to process  $i$  files is

$$\text{RAM (in GB)} = 0.5 i + 6. \quad (9)$$

In the IMS system, the VPU immediately processes any received video data file. In practice, however, the amount of RAM assigned to the VPU through (9) limits the number of concurrent transcoding processes. If this limit has been reached, eg, for a peak in the number of CE participants, one option is to temporarily buffer the input file in the VPU local storage system until one of the ongoing transcoding processes finishes. However, the CESC architecture allows instantiating a new VPU and balance the load among the active VNFs. In cloud terminology, this process is called scaling.<sup>18</sup> Having more VPU instances will result in more hardware resources and this is exactly how the change in the consumption behavior of end users affects capital costs (CAPEX).

## 3.2 | Resource planning for media content online transcoding

To analyze video streaming, we will model a participant in a CE generating a live stream toward the VPU as a Poisson process, with rate equal to  $\gamma_0$  streams per second, independent from the video-file originating process considered in previous Section 3.1. The average duration of live streams is  $T$  seconds.

The VPU forwards in real-time the received stream to all the users in the group wishing to receive it, through an ad hoc notification mechanism. More specific details on this mechanism and on the forwarding algorithm are outside the scope of this paper (the interested reader can see the work of Paglierani et al<sup>20</sup>). Furthermore, we will assume that the VPU only adjusts the resolution of video contents from high to medium (in the interest of a silver user). This is a reasonable assumption because a change in resolution from medium to high would bring about quality degradation.

For the sake of simplicity, we will consider separately the resource estimation procedure for gold and silver users. In particular, we will assume that one VPU instance can provide  $M_{MAXGOLD}$  online transcoding sessions when the originator is a gold user and  $M_{MAXSILVER}$  online transcoding sessions in the case of a silver user. As in the offline transcoding analysis,  $M_{MAXGOLD}$  and  $M_{MAXSILVER}$  can be evaluated by means of ad hoc performance tests.

An extra parameter, ie, the blocking probability  $E_m$  is needed to evaluate the additional resources required by the online transcoding process.  $E_m$ , which is the probability that a user requiring online video transcoding finds all  $m$  available transcoding resources already occupied, can be obtained using the well-known Erlang-B formula. A simple iterative way to evaluate  $E_m$  is

$$E_m = 1/I_m; I_m = 1 + (m/A)I_{m-1}; m = 1, 2, \dots, M, \quad (10)$$

where  $A = N \gamma_0 T$  is the offered traffic stated in Erlang,  $m$  is the number of identical parallel resources, and  $E_0 = 1$  and  $I_0 = 1/E_0$  are the iteration initial values.<sup>25</sup> Setting the blocking probability  $E_M$  at a desired value, from (10), one can obtain the corresponding number  $M$  of required resources. Indicating with  $M_{MAX}$  the maximum number of concurrent sessions that a single VPU instance can run, the overall number of VPU instances needed to serve all the transcoding requests is

$$M_{VPUON} = M/M_{MAX}. \quad (11)$$

Finally, we can separately evaluate the needed resources for gold and silver users by inserting in (11)  $M_{MAXGOLD}$  and  $M_{MAXSILVER}$ , respectively.

### 3.3 | Estimating the overall storage capacity

To estimate the storage requirements for the IMS system, we assume that the average duration of video sequences is  $T$  seconds, and that the data rates assigned to silver and gold users are equal to  $b_S$  and  $b_G$  bit per second (bps), respectively. The CE duration (in seconds) is  $T_E$ . With the above assumptions, and considering that the transcoding processes originate both medium and high resolution contents for gold users, and only medium resolution contents for silver users, the storage capacity  $C$  (in bits) is:

$$C = T_E * (\lambda_0 + \gamma_0) * (U_G * b_G + U_G * b_S + U_S * b_S) * T \quad (12)$$

### 3.4 | Summary of VPU model parameters

The VPU model presented so far depends on a number of parameters. For the sake of clarity, in Table 1, we summarize all the VPU model parameters of interest for the planning tool.

### 3.5 | Radio resource estimation

To determine the number of SCs needed for a particular venue, two main factors are considered: (1) the maximum data rate that can be supported by an SC and (2) the coverage area of an SC.

The current 4G LTE-A standard allows a maximum downlink data rate of 403.3 Mbps and a maximum uplink rate of 100.8 Mbps, as summarized in Table 2.<sup>27,28</sup> However, as 5G technology matures, aggregated data rates of more than 1 Gbps are expected.<sup>28</sup> For the purpose of the planning study in this paper, we are assuming a downlink scenario and a typical coverage area of 800 m<sup>2</sup> per SC.

**TABLE 1** Main parameters to model the video processing unit (VPU)

Parameter	Description
$\lambda_0$ (Fps)	Rate of generation of s per second (Fps) per single user
$N$	Total number of users served by a specific VPU instance
$\lambda$ (Fps)	Rate of generation of video Fps per $N$ users
$Q$	Number of single-transcoding processes running in parallel
$\mu_0, \mu_1$ (fps)	Transcoding rates for medium and high resolution contents (fps)
$\mu$ (fps)	Overall service rate of the multiple-transcoding process.
$\mu_F$ (Fps)	The overall service rate in Fps
$\gamma_0$	Number of live streams per second toward the VPU, originated by each participant to the CE
$T$ (s)	Average duration of a video stream
$D_0$ (s)	Delay introduced by the VPU when processing the average-size file only

**TABLE 2** Small cell parameters for current LTE/A standard

Parameter	Value
Maximum bandwidth	20 MHz
Number of PRB*	100
Modulation scheme	64 QAM
MIMO scheme	4 × 4
Uplink data rate	100.8 Mbps
Downlink data rate	403.3 Mbps

\*physical resource block



The radio planning strategy relies on calculating the minimum number of needed SC based on the largest figure obtained after considering the aggregate data rate of users in the venue and the coverage area per SC. In other words, it follows a worst-case planning strategy.

Assuming the number of gold, silver, and bronze users to be  $U_G$ ,  $U_S$ , and  $U_B$  and the corresponding data rate per user to be  $D_G$ ,  $D_S$ , and  $D_B$ , then the aggregate data rate  $D_{tot}$  can be written as

$$D_{tot} = (U_G \times D_G) + (U_S \times D_S) + (U_B \times D_B). \quad (13)$$

If the maximum data rate supported by an SC is  $D_{sc}$ , then the number of needed SCs  $N_{SCd}$  using the data rate-based design can be calculated as

$$N_{SCd} = \text{round}(D_{tot}/D_{sc}). \quad (14)$$

Assuming the area of the venue to be  $A$  and the coverage area of an SC to be  $A_{sc}$ , then the number of needed SCs  $N_{SCa}$  based on the coverage area design can be calculated as

$$N_{SCa} = \text{round}(A/A_{sc}). \quad (15)$$

In some cases, a single SC can only support a maximum number of users  $U_{max}$  (eg, 128) irrespective of the individual data rate. This adds an additional constraint to the radio planning design and can eventually increase the number of needed SCs. If this is taken into consideration, then the needed number of SCs  $N_{SCu}$  can be estimated as

$$N_{SCu} = \text{round}(U/U_{max}). \quad (16)$$

Finally, the number of needed SCs,  $N_{sc}$  can be obtained as

$$N_{SC} = \text{maximum}(N_{SCd}, N_{SCa}, N_{SCu}). \quad (17)$$

## 4 | THE PLANNING STUDY

The planning tool provides the means to calculate the needed IT and radio resources to guarantee adequate service provisioning in a venue with an average number of attendees per event. In the following section, we discuss some simple but meaningful examples of its use in real-life study cases.

### 4.1 | Use cases and scenarios

The main target use cases for IMS are conference halls and sports venues of different dimensions and user capacities. Users participating to a CE are classified depending on their selected bundle into different classes that define the corresponding target data rates, as presented in Table 3.

### 4.2 | Resource planning examples

In the CESC architecture, for each  $U_{max} = 128$  CE participants, one instance of the SC VNF should run in the Light DC.<sup>29,30</sup> We will assume that the physical servers deployed in the Light DC are equipped with eight core Intel Xeon CPUs

Bundle	User data rate	Tariff	Applications
Gold	7 Mbps	1 (€/event)	voice, text, images, streaming HD video
Silver	1.7 Mbps	0.5 (€/event)	voice, text, images, streaming SD video
Bronze	0.5 Mbps	0.25 (€/event)	voice, text, images, and low definition video

TABLE 3 Users and target data rates

**TABLE 4** Resource planning example

Parameter	Type (Input/Output)	Value
$\lambda_0$	Input	$5 \times 10^4$ Fps
$\mu_0, \mu_1$	Input	400, 700 fps
$U_{\text{GOLD}}$	Input	500 users
$D_E$	Input	5 s
$N_{\text{FILE}}$	Input	1000 frames
$\pi_i$	Input	$10e-5$
T	Input	40 s
$T_E$	Input	14 400 s
$\mu_F$	Output, from (2)	0.255 Fps
$N(D_E, \mu_F, \lambda_0)$	Output form (4)	110
$i$	Output from (8)	6 channels
RAM	Output from (9)	9 GB
C	Output from (12)	235

at 2.4 GHz. A subset of these, dedicated to provide IMS, includes one (or more) Nvidia M4000 GPU.<sup>20</sup> The required IT resources for a single SC VNF ( $UR_{\text{SC VNF}}$ ) is a set of 2 CPU cores, 4 GB of RAM, and 2 GB of HDD.<sup>26</sup> The total required resources ( $R_{\text{total}}$ ) to provide the basic service bundle (bronze) in a venue with  $U_{\text{TOTAL}}$  number of attendees is

$$R_{\text{total}} = \text{round} (U_{\text{TOTAL}}/U_{\text{max}}) \times UR_{\text{SC VNF}}. \quad (18)$$

The additional resources to provide IMS in the venue can be estimated through the planning tool. To clarify its use in a real-life scenario, in Table 4, we summarize a simple but meaningful application example. The reported parameters refer to a stadium with  $U_{\text{TOTAL}} = 10\,000$  attending users; for the sake of simplicity, the number of silver users has been set to zero. The single transcoding rates  $\mu_0$  and  $\mu_1$  have been achieved with Nvidia M4000 GPUs.<sup>20,24</sup> The value  $N(D_E, \mu_F, \lambda_0) = 110$  has been obtained with  $D_E$  equal to 5 seconds. However, if we increase  $D_E$  to 10 seconds leaving  $\mu_F, \lambda_0$  unchanged,  $N(D_E, \mu_F, \lambda_0)$  becomes 310. Further relaxing  $D_E$  to 15 seconds or 20 seconds provides values of  $N(D_E, \mu_F, \lambda_0)$  equal to 376 and 410, respectively.

The second column in Table 4 indicates if a parameter is an input or on output in the planning tool; in this latter case, we provide the formula used to calculate it.

## 5 | TECHNO-ECONOMIC ANALYSIS

### 5.1 | Methodology and tools

The techno-economic methodology used in this paper for the assessment of IMS provisioning during CEs is based on an excel-based tool developed within the IST-TONIC (TecnO-ecoNomICs of IP optimized networks and services) and CELTIC-ECOSYS (techno-ECONomics of integrated communication SYStems and services) projects, which have already been successfully applied in numerous similar studies.<sup>31-33</sup>

Firstly, the tool requires the selection of a suitable study period. The final output of the analysis will be a set of meaningful economic parameters, calculated on a per-year basis, over the study period.

The core part of the used approach is in a database regularly updated with network components, data collected from the largest European telecommunication companies and vendors, and from benchmarks from the telecom market. Starting from the estimated IT and radio resources, provided by the planning tool described in the previous sections, the database returns the “shopping list,” ie, a detailed list containing all the required elements (equipment, cables, racks, installation, etc) calculated for each year of the study period.<sup>32</sup> To this end, the tool performs demand forecasting using existing methodologies and market data. In the literature, several studies are available to understand the diffusion of new products, services, and technologies.<sup>34,35</sup> In this paper, to forecast both the demand for subscriptions and the penetration of IMS, a four-parameter logistic model is used. This model is recommended for long-term forecasts and for new services and is commonly used for fixed and mobile networks.<sup>36</sup> It relies on the following expression:

$$Y_t = M / (1 + e^{a+bt})^c, \quad (19)$$

where  $Y_t$  is the actual or forecasted demand at time  $t$  as a population percentage;  $M$  is the demand saturation level as a population percentage;  $t$  is the time in years; and  $a$ ,  $b$ , and  $c$  are diffusion parameters that can be estimated by a regression analysis using existing market data.<sup>36</sup>

The price evolution through the study period for all network components is obtained using the extended learning curve model.<sup>34</sup> Finally, CAPEX is calculated by combining the required number of components and their price for each year.<sup>32</sup>

Furthermore, using (19) and available market forecasts, one can calculate the future market penetration of the services and the tariffs associated with them; inserting such information in the tool, one obtains the service revenues for each year of the study period.<sup>31,32</sup>

The assessment presented in the following sections also includes operating costs, ie, maintenance costs and operating expenditures (OPEX).

Maintenance costs consist of two parts: (1) the cost of repair calculated as a fixed percentage of the total investments in network elements and (2) the cost of repair work calculated based on the mean time between failures (MTBF) and the mean time to repair (MTTR).

Operating expenditures (OPEX) are also calculated. For example, energy costs are evaluated based on the power consumption of components and the average cost of 1 kWh.<sup>37</sup> It should be highlighted that, for the investigated CE case, it is assumed that network equipment is used during the event and switched off before and after.

By combining service revenues, investments (CAPEX), operating costs, and other general economic inputs (eg, discount rate, tax rate), the adopted tool can perform a discounted cash flow (DCF) analysis.<sup>31,32</sup> The DCF analysis takes into account the time value of money and the risks of investing in a project. The main advantages of DCF are that it is a simple quantitative method to implement, it is widely accepted and provides clear and consistent metrics such as cash flows, net present value (NPV), internal rate of return (IRR), payback period, and other economic figures of merit.<sup>31</sup>

The interested reader can find all the details of the methodology used by the economic tool in other works.<sup>31-34</sup> The focus of this paper, conversely, is on its application to the provision of IMS during a sporadic event to advance the knowledge about the neutral host business model.

## 5.2 | Study assumptions, parameters, and data

This paper analyzes the investment of an SCNO who would like to deploy an IMS capable SESAME-like SC network in a stadium. A large stadium is assumed, with a capacity of 90 000 and an area of 40 000 m<sup>2</sup>. We will assume a study period of 10 years. The 10-year period starts in 2020 (the expected year of 5G introduction) and ends in 2029. The discount rate is 10% while taxes are equal to 20%. We assume that three MNOs, with market shares of 50%, 30%, and 20%, respectively, have an agreement to access the SC network. The deployment of macrosites or any other infrastructure and equipment of MNOs will not be taken into account.

Extensive research and opinion polls from various market research firms<sup>38</sup> and manufacturers<sup>3</sup> reasonably ensure that the customer base will be equipped with 5G-ready smartphones starting in 2019. Hence, it seems that the entire telecom value chain will push for 5G to become a commercial reality in 2019 in developed Asia and USA and in 2020 in Europe. By combining market insights and 4G historical data along with (19), the demand model for 5G is derived as shown in Figure 4.

The penetration for each service bundle is modeled using (19) and with the following assumptions. During the first years of 5G services introduction, the preferred packet will be the bronze bundle. However, in the following years, silver and gold bundles will prevail, attracting an increasing number of subscribers. Customers will be willing to pay higher subscriptions thanks to innovative services and applications available only to these bundles as illustrated in Figure 5.

The SCNO will also incur in operational expenses to assure proper access to MNOs. These expenses entail the payment of its employees, the rental cost, electricity cost, etc.

Each employee's monthly fee is assumed to be €4000. The total number of personnel is calculated assuming that there is a need for an employee per 2000 subscribers of the MNOs. The building (where IT and networking infrastructure will be installed) rent per month and cost per kilowatt hour are considered to be €2000 and €0.205, respectively.

## 6 | RESULTS AND DISCUSSION

In this section, the results of the techno-economic model are presented and discussed. The results include network investments, operational expenditures, and revenues while the financial outcome is expressed by the financial indices discussed before.

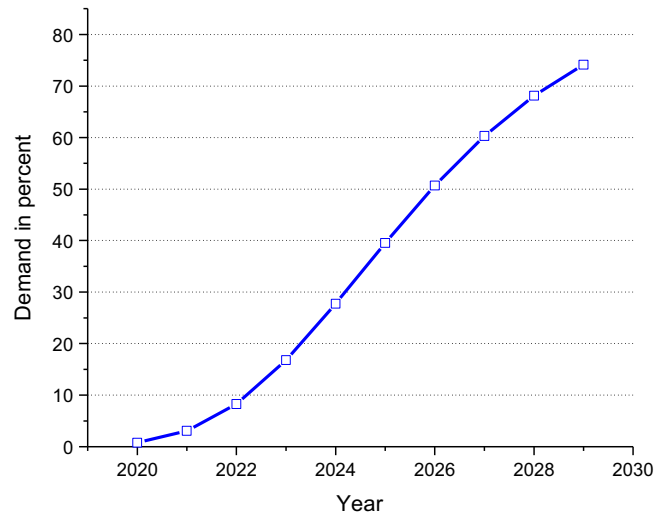


FIGURE 4 Fifth-generation demand model [Source: INCITES]

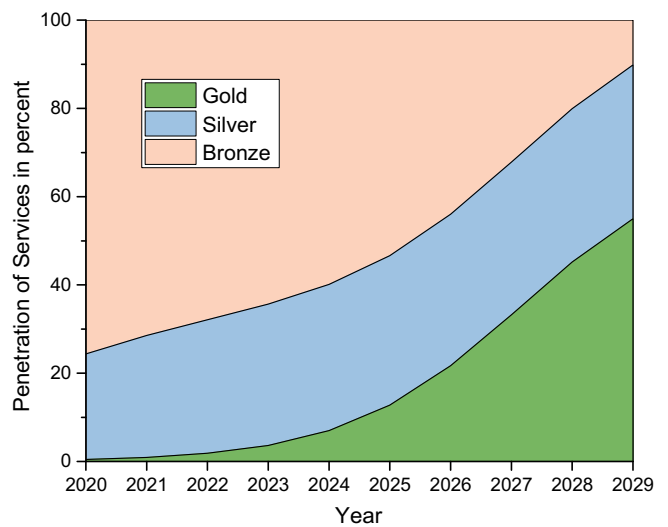


FIGURE 5 Services penetration forecast

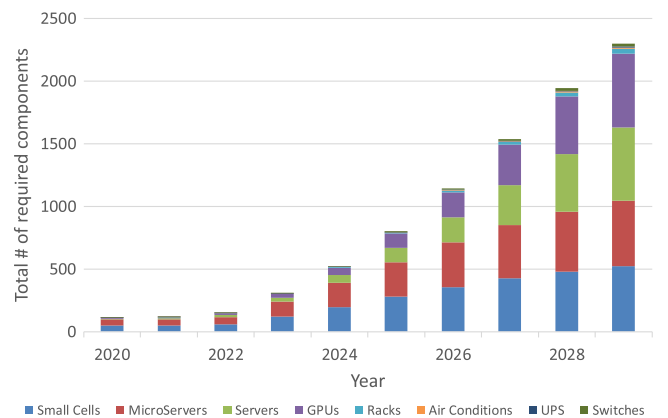
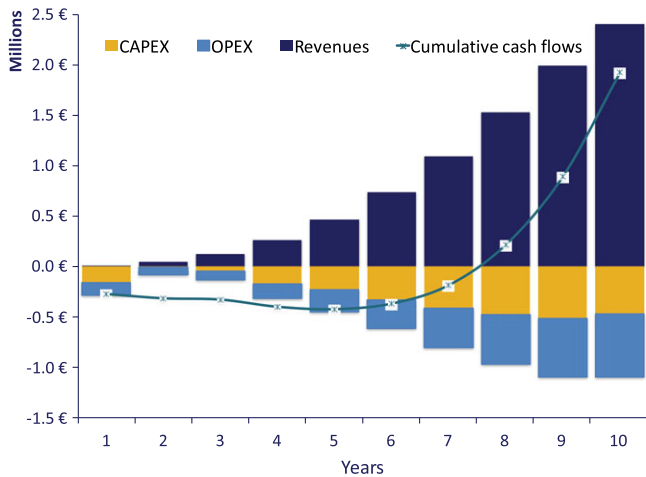


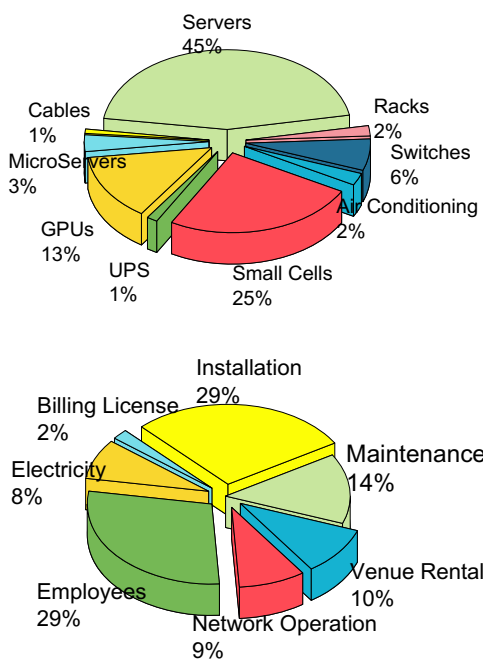
FIGURE 6 Required number of components during the lifetime of the project: total network

During the 10-year study period, the SCNO will have to invest in equipment such as racks, switches, servers, microservers, and small cells in order to provide wholesale access to MNOs. In Figure 6, there is a detailed rendition of the quantities of these investments throughout the study. As shown in Figure 6, SCNO will progressively deploy the network as 5G and IMS adoption evolves.

The basic economic results are presented in Figure 7 for the considered stadium. The cumulative discounted cash flow, widely known as the cash balance, summarizes the total economics and presents the total financial evolution of the business case. The cash flow depicts the balance for each year. The first observation is that the balance is initially negative.



**FIGURE 7** Annual revenues and cost analysis of the fifth-generation investment



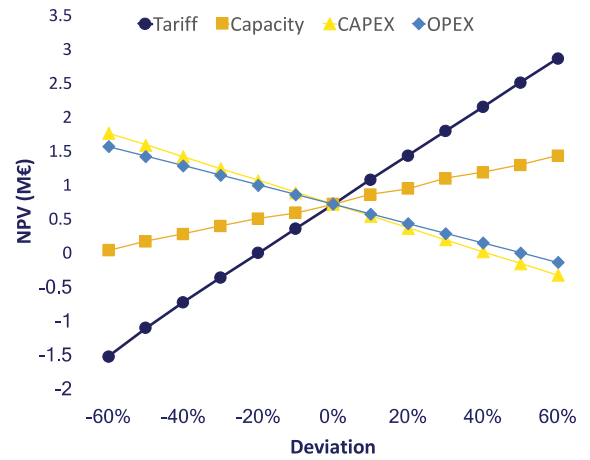
**FIGURE 8** Breakdown of investments (up) and running costs (down)

This is the usual case for telecom investments and can be attributed to the required high initial investments of the first years to actually deploy the network together with the relatively lower demand for 5G services. However, as the demand for 5G increases, the revenues increase making the investment profitable within the study period. Our calculations show that the payback period for the studied investment is expected to be approximately 6.5 years with an IRR equal to 27% and a NPV of approximately €705 thousand.

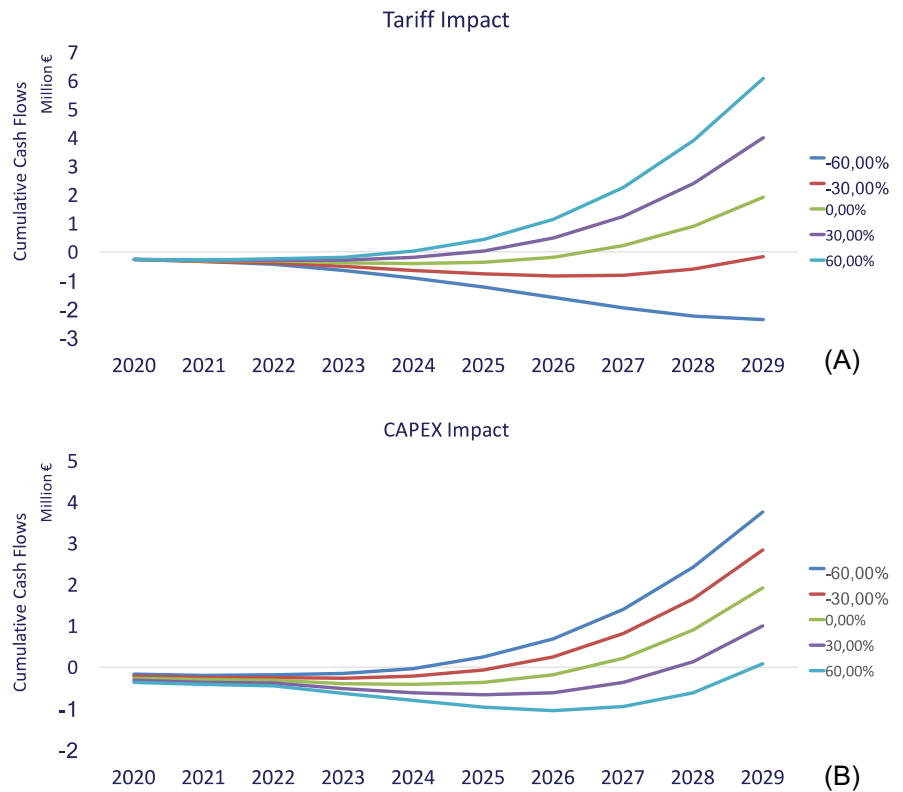
To have a better insight of the expenditures, a breakdown of investments and running costs (OPEX) for the whole IMS-capable CESC infrastructure is illustrated in Figure 8. Regarding investments, it can be concluded that the major contributors to cost are the following: (1) servers (44.9%), (2) small cells (25.1%), and (3) GPUs (13.5%) that are used to accelerate video transcoding. It should also be highlighted that the analysis revealed that the cost contribution of CESC microservers is low (3%). This is in line with SESAME targets of using general-purpose cost-effective servers ( $\mu$ -servers) to provide basic connectivity.

The costs related to installations and employees are dominant and are equal to 58% of the total OPEX. Although the cost for electricity is quite high, it results lower than other cases/projects since IT and radio resources are switched off before and after a CE.

In order to understand how the project's NPV would be affected, should one alter input parameters such as bundle tariffs, stadium capacity, CAPEX, and OPEX, a series of sensitivity analyses were performed.<sup>39</sup> The chosen parameters were altered within an interval of  $\pm 60\%$  of their initially assumed values. Figure 9 shows the results of the sensitivity



**FIGURE 9** Sensitivity analysis: dependence of NPVs on tariff, capacity, CAPEX and OPEX.<sup>39</sup> NPV, net present value



**FIGURE 10** Sensitivity analyses on cumulative cash flows in order to assess: A, Tariffs' B, CAPEX impact

analysis. The most sensitive parameter for the project's NPV appears to be the services' fees (tariffs). As shown, NPV can be improved by as much as €2 million in case of a 60% increase of the tariffs (ie, €1.6, €0.8, and €0.4/month in the case of gold, silver, and bronze subscriber, respectively). On the other hand, if the service tariff gets reduced by 20% (eg, €0.8, €0.4, and €0.2/month in the case of gold, silver, and bronze subscriber, respectively), then this leads to a marginally negative NPV.

To get a deeper insight on the impact of the tariff and CAPEX parameters, sensitivity analyses on cumulative cash flows were performed. Simulation results from 2020 to 2029 show that the business case under investigation appears more sensitive to services tariffs, which needs to be kept at or above their 2020 levels for a positive NPV to exist (Figure 10A). When it comes to the sensitivity of CAPEX, it could increase by as much as approximately 30% and still get a nonnegative NPV (Figure 10B).

## 7 | CONCLUSIONS AND FUTURE WORK

We presented in this paper a techno-economic analysis to explore the viability of deploying a cloud-enabled small cell network capable to provide immersive video services during CEs. Forecasts predict approximately 20% and 50% market

penetration for IMS in 4- and 7-year duration, respectively. The main factors contributing to CAPEX are servers (45%), small cells (25%), and GPU (14%) that must be installed in the edge infrastructure. Installation and employee costs are the two dominant factors that jointly cover approximately 58% of OPEX. Our study shows that an investment needs approximately 6.5 years to reach a breakeven point. Considering the effective functional period of telecom infrastructures which is 15 to 20 years, the return on investment happens at a relatively reasonable point making it profitable.

The techno-economic analysis presented in this paper is an original and novel development, based on results obtained in three funded research projects, namely, SESAME for the planning model, and IST-TONIC and CELTIC-ECOSYS for the economic part. The obtained tools have been successfully applied in various preliminary trials carried out in the SESAME project. However, a rigorous scientific validation of the overall model applied to a real-life network and a real CE is an extremely difficult activity which falls outside the scope of this paper. The experimental validation of the proposed techno-economic analysis is an ongoing research activity that will be presented in future publications.

## ACKNOWLEDGMENTS

Authors would like to thank Mr Nick Johnson and Mr Alan Whitehead from ip.access for their support. This research has been supported by the H2020 5G-PPP project SESAME (No. 671596).

## ORCID

Pietro Paglierani  <https://orcid.org/0000-0002-7786-0848>

## REFERENCES

- Hossain E, Hasan M. 5G cellular: key enabling technologies and research challenges. *IEEE Instrum Meas Mag.* 2015;18(3):11-21. <https://doi.org/10.1109/MIM.2015.7108393>
- Cisco. Cisco visual networking index: forecast and methodology, 2017-2022. 2017.
- Ericsson. Ericsson mobility report. 2018. <https://www.ericsson.com/en/mobility-report>. Retrieved February 28, 2019.
- Amdocs. State of the RAN. 2015. [http://solutions.amdocs.com/rs/amdocs1/images/Amdocs\\_2015\\_State\\_of\\_RAN.pdf](http://solutions.amdocs.com/rs/amdocs1/images/Amdocs_2015_State_of_RAN.pdf). Retrieved February 28, 2019.
- Ericsson. Rio-aiming-higher-report. 2016. <https://www.ericsson.com/assets/local/narratives/networks/documents/rio-aiming-higher-report.pdf>. Retrieved February 28, 2019.
- Nokia. Smart Venues. 2017. <https://onestore.nokia.com/asset/201454>. Retrieved February 28, 2019.
- SESAME. 2016. <http://www.sesame-h2020-5g-ppp.eu/>. Retrieved August 31, 2017.
- 5G-PPP. 2017. <https://5g-ppp.eu/>. Retrieved August 31, 2017.
- Small Cell Forum. Multi-operator and neutral host small cell. Version 191.08.02. December 2016.
- Taleb T, Samdanis K, Mada B, Flinck H, Dutta S, Sabella D. On multi-access edge computing: a survey of the emerging 5G network edge cloud architecture and orchestration. *IEEE Commun Surv Tutor.* 2017;19:1657-1681.
- 5GEssence. <http://www.5g-essence-h2020.eu/>. Retrieved February 28, 2019.
- 5GCity. <https://www.5gcity.eu/>. Retrieved February 28, 2019.
- Giannoulakis I, Kafetzakis E, Trajkovska I, et al. The emergence of operator-neutral small cells as a strong case for cloud computing at the mobile edge. *Trans Emerg Telecommun Technol.* 2016;27:1152-1159. <https://doi.org/10.1002/ett.3077>
- Oughton EJ, Frias Z. The cost, coverage and rollout implications of 5G infrastructure in Britain. *Telecommunications Policy.* 2018;42(8):636-652, ISSN 0308-5961, <https://doi.org/10.1016/j.telpol.2017.07.009>.
- Oughton EJ, Frias Z, van derGaast S, van derBerg R. Assessing the capacity, coverage and cost of 5G infrastructure strategies: analysis of the Netherlands. *Telematics Inform.* 2019. <https://doi.org/10.1016/j.tele.2019.01.003>
- Neokosmidis I, Rokkas T, Parker MC, et al. Assessment of socio-techno-economic factors affecting the market adoption and evolution of 5G networks: evidence from the 5G-PPP CHARISMA project. *Telematics Inform.* 2017;34(5):572-589. <https://doi.org/10.1016/j.tele.2016.11.007>
- Fajardo JO, Liberal F, Giannoulakis I, et al. Introducing mobile edge computing capabilities through distributed 5G cloud enabled small cells. *Mob Netw Appl.* 2016;21(4):564-574. <https://doi.org/10.1007/s11036-016-0752-2>
- GS NFV-MAN 001 - V1.1.1 - Network functions virtualisation (NFV); management and orchestration. 2014. [http://www.etsi.org/deliver/etsi\\_gs/NFV-MAN/001\\_099/001/01.01.01\\_60/gs\\_NFV-MAN001v010101p.pdf](http://www.etsi.org/deliver/etsi_gs/NFV-MAN/001_099/001/01.01.01_60/gs_NFV-MAN001v010101p.pdf)
- Jain R, Paul S. Network virtualization and software defined networking for cloud computing: a survey. *IEEE Commun Mag.* 2013;51(11):24-31. <https://doi.org/10.1109/MCOM.2013.6658648>
- Paglierani P, Meani C, Albanese A, Crosta P. Immersive video services at the edge: an energy-aware approach. *IARIA Int J Adv Telecommun.* 2017;10.
- Cattaneo G, Giust F, Meani C, Munaretto D, Paglierani P. Deploying CPU-intensive applications on MEC in NFV systems: the immersive video use case. *Computers.* 2018;7:55.

22. ITU-T. Advanced video coding for generic audiovisual services. 2013. <http://www.itu.int/ITU-T/recommendations/rec.aspx?rec=11466&lang=en>. Retrieved August 31, 2017.
23. Wiegand T, Sullivan GJ, Bjontegaard G, and Luthra A. Overview of the H.264/AVC video coding standard. *IEEE Trans Circuits Syst Video Technol.* 2003;13(7):560-576. <https://doi.org/10.1109/TCSVT.2003.815165>
24. Comi P, Crosta P, Beccari M, et al. Hardware-accelerated high-resolution video coding in virtual network functions. In: Proceedings of the 2016 European Conference on Networks and Communications (EuCNC); 2016; Athens, Greece.
25. Zukerman M. *Queueing Theory and Stochastic Teletraffic Models*. 2017. <https://arxiv.org/pdf/1307.2968.pdf>
26. SESAME D7.3. Experimental integration results of HW/SW modules of the overall SESAME framework. 2017.
27. Norhborg M. LTE overview. <http://www.3gpp.org/technologies/keywords-acronyms/98-lte>
28. Minokuchi A, Isobe S, Takahashi H, Nagata S. 5G standardization trends at 3GPP. *NTT DOCOMO Tech J.* 2018;19(3):5-12.
29. SESAME D3.1. CESC prototype design specifications and initial studies on self-X and virtualization aspects. 2016. [http://www.sesame-h2020-5g-ppp.eu/Portals/0/Deliverables/SESAME\\_GANo.671596\\_Deliverable3.1\\_v1.0\\_final.pdf?ver=2016-10-24-162425-640](http://www.sesame-h2020-5g-ppp.eu/Portals/0/Deliverables/SESAME_GANo.671596_Deliverable3.1_v1.0_final.pdf?ver=2016-10-24-162425-640)
30. SESAME D4.4. Light DC prototype. 2015. <http://www.sesame-h2020-5g-ppp.eu/Deliverables.aspx>
31. Monath T, Kristian N, Cadro P, Katsianis D, Varoutas D. Economics of fixed broadband access network strategies. *IEEE Commun Mag.* 2003;41(9):132-139. <https://doi.org/10.1109/MCOM.2003.1232248>
32. Rokkas T, Kamalakis T, Katsianis D, Varoutas D, Sphicopoulos T. Business prospects of wide-scale deployment of free space optical technology as a last-mile solution: a techno-economic evaluation. *J Opt Netw.* 2007;6(7). <https://doi.org/10.1364/JON.6.000860>
33. Fokas M, Katsianis D, Varoutas D, et al. Initial techno-economic results for OMEGA home gigabit networks. In: Proceedings of the 21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications; 2010; Istanbul, Turkey.
34. Ims LA, ed. *Broadband Access Networks*. Boston, MA: Springer; 1998. <https://doi.org/10.1007/978-1-4615-5795-1>
35. Fildes R, Kumar V. Telecommunications demand forecasting—a review. *Int J Forecast.* 2002;18(4):489-522. [https://doi.org/10.1016/S0169-2070\(02\)00064-X](https://doi.org/10.1016/S0169-2070(02)00064-X)
36. Michalakelis C, Varoutas D, Sphicopoulos T. Diffusion models of mobile telephony in Greece. *Telecommunications Policy.* 2008;32(3-4): 234-245. <https://doi.org/10.1016/j.telpol.2008.01.004>
37. Energy price statistics. 2017. [http://ec.europa.eu/eurostat/statistic-explained/index.php/Energy\\_price\\_statistics](http://ec.europa.eu/eurostat/statistic-explained/index.php/Energy_price_statistics)
38. Ovum. Ovum predicts 24 million 5G subscriptions by 2021. *TM Forum Inform.* 2016. <https://inform.tmforum.org/news/2016/06/ovum-predicts-24-million-5g-subscriptions-by-2021>. Retrieved August 31, 2017.
39. Neokosmidis I, Rokkas T, Paglierani P, et al. Techno economic assessment of immersive video services in 5G converged optical/wireless networks. In: Proceedings of the 2018 Optical Fiber Communications Conference and Exposition (OFC); 2018; San Diego, CA.

**How to cite this article:** Paglierani P, Neokosmidis I, Rokkas T, et al. Techno-economic analysis of 5G immersive media services in cloud-enabled small cell networks: The neutral host business model. Providing techno-economic guidelines for the successful provision of 5G innovative services in small cell networks. *Trans Emerging Tel Tech.* 2019;e3746. <https://doi.org/10.1002/ett.3746>

## APPENDIX

We consider  $Q$  single transcoding sessions  $p_0, p_1, \dots, p_{Q-1}$ , with rates  $\mu_0 > \mu_1 > \dots > \mu_{Q-1}$ , running in parallel, all started at time  $T_0$ . When all  $Q$  sessions are active and under the EPS assumption, the actual rate of the  $i$ -th session  $p_i$  is  $\mu_i/Q$ . At time  $T_1$ ,  $p_0$  terminates, having processed all the  $N_{\text{FILE}}$  frames contained in the input video file, whereas  $p_1, \dots, p_{Q-1}$  (running at lower rates) are still active. It also results

$$\Delta T_1 = T_1 - T_0 = N_{\text{FILE}} * Q / \mu_0.$$

At time  $T_1$ ,  $p_1$  has processed  $\Delta T_1 * \mu_1 / Q$  frames out of the total  $N_{\text{FILE}}$  frames of the input video file. After  $T_1$ , there are  $Q-1$  active tasks until  $T_2$ , when also  $p_1$  ends, having processed  $N_{\text{FILE}}$  frames. We can thus write

$$\Delta T_2 = T_2 - T_1 = (N_{\text{FILE}} - \Delta T_1 * \mu_1 / Q) * (Q - 1) / \mu_1 = N_{\text{FILE}} * (Q - 1) * (1 / \mu_1 - 1 / \mu_0) \quad (\text{A1})$$

and, by iterating the procedure,

$$\Delta T_3 = T_3 - T_2 = (N_{\text{FILE}} - \Delta T_1 * \mu_2 / Q - \Delta T_2 * \mu_2 / (Q - 1)) * (Q - 2) / \mu_2 = N_{\text{FILE}} * (Q - 2) * (1 / \mu_2 - 1 / \mu_1)$$



being  $T_3$  the time at which p2 ends. We thus assume that, up to  $i$ ,

$$\Delta T_i = T_i - T_{i-1} = N_{\text{FILE}} * (Q - i - 1) * (1/\mu_{i-1} - 1/\mu_{i-2}). \quad (\text{A2})$$

We then observe that  $\Delta T_{i+1}$  corresponds to the time it takes to  $p_i$  to process  $\Delta N_{i+1}$  frames, with

$$\begin{aligned} \Delta N_{i+1} &= N_{\text{FILE}} - \Delta T_1 * \mu_i / Q - \Delta T_2 * \mu_i / (Q - 1) - \dots - \Delta T_i * \mu_i / (Q - i - 1) \\ &= N_{\text{FILE}} - \mu_i * N_{\text{FILE}} * (1/\mu_0 + 1/\mu_1 - 1/\mu_0 + 1/\mu_2 - 1/\mu_1 + \dots + 1/\mu_{i-1} - 1/\mu_{i-2}) \\ &= N_{\text{FILE}} - \mu_i * N_{\text{FILE}} / \mu_{i-1}. \end{aligned} \quad (\text{A3})$$

In  $\Delta T_{i+1}$ , the rate of  $p_i$  is  $\mu_i / Q - i$ , since there are still  $Q - i$  active tasks. Hence,

$$\Delta T_{i+1} = T_{i+1} - T_i = \Delta N_{i+1} * (Q - i) / \mu_i. \quad (\text{A4})$$

Using A3 in A4, one finally obtains

$$\Delta T_{i+1} = N_{\text{FILE}} * (Q - i) * (1/\mu_i - 1/\mu_{i-1}), \quad (\text{A5})$$

which proves by mathematical induction the validity of A2 for any  $i$  such that  $2 \leq i \leq Q$ .

The overall time to terminate all the parallel sessions is

$$\Delta T = \Delta T_1 + \Delta T_2 + \dots + \Delta T_{M-1} + \Delta T_Q. \quad (\text{A6})$$

Substituting A2 in A6, one obtains

$$\begin{aligned} \Delta T &= N_{\text{FILE}} * [Q/\mu_0 + (Q - 1)(1/\mu_1 - 1/\mu_0) + \dots + 2/(1/\mu_{Q-2} - 1/\mu_{Q-3}) + 1/(1/\mu_{Q-1} - 1/\mu_{Q-2})] \\ &= N_{\text{FILE}} * [1/\mu_0 + 1/\mu_1 + \dots + 1/\mu_{Q-3} + 1/\mu_{Q-2} + 1/\mu_{Q-1}]. \end{aligned}$$

The overall transcoding rate  $\mu$  results are equal to

$$\mu = [1/\mu_0 + 1/\mu_1 + \dots + 1/\mu_{Q-3} + 1/\mu_{Q-2} + 1/\mu_{Q-1}]^{-1}.$$