

How to Migrate From Operational LTE/LTE-A Networks to C-RAN With Minimal Investment?

Davit Harutyunyan¹ and Roberto Riggio², *Senior Member, IEEE*

Abstract—By leveraging the fully-centralized and virtualized cloud radio access network (C-RAN) architecture over densely deployed small cells, mobile network operators (MNOs) are expected to meet the ever-increasing coverage and capacity demands. Towards this end, finding the optimal numbers, and locations of centralized unit (CU) pools, and centralizing the baseband units of eNBs at the optimal CU pools plays a pivotal role in curtailing the required investments in order to transit from legacy decentralized RAN (D-RAN) to C-RAN. In this paper, we propose an approach for MNOs to adopt the C-RAN architecture with minimal investment by using the available infrastructure (e.g., site locations and transmission links). Specifically, we propose a decentralized unit – CU (DU-CU) mapping algorithm, which effectively selects the quantity and the locations of CU pools and assigns the CU of each DU to the appropriate CU pool. We then compare the traffic aggregation gains of C-RAN and traditional D-RAN. Lastly, in order to quantify the total cost of ownership savings that can be obtained by employing the legacy network infrastructure while migrating to C-RAN, we compare this scenario with the C-RAN migration scenario in which there is no available infrastructure. In both scenarios, the mapping algorithms are formulated as virtual network embedding problems using integer linear programming techniques. The results of the simulations, conducted using data traffic of 26 eNBs (209 cells) of an operational LTE-A mobile network, reveal that significant saving can be obtained by employing the available mobile network infrastructure while migrating to C-RAN.

Index Terms—C-RAN, traffic aggregation, DU-CU mapping, multiplexing gain.

I. INTRODUCTION

MOBILE data traffic has been snowballing over the last few years due to various applications with their diverse requirements in terms of latency, data rates and traffic volume [1], [2]. By leveraging the fully-centralized and virtualized Cloud Radio Access Network (C-RAN) architecture over densely deployed small cells, Mobile Network Operators (MNOs) are expected to satisfy their coverage and capacity demands, which have recently been increasing at an unprecedented rate. In C-RAN, baseband units (termed Centralized Unit – CU) are decomposed from radio elements (termed

Distributed Unit – DU), consolidated in large data-centers (termed Centralized Unit pool – CU pool)¹ and are shared among multiple cells [7], [8]. The separation of the baseband processing functionalities in the RAN protocol stack between the CU pool and the DU is known as functional split. It is worthwhile to mention that only the classical C-RAN functional split, also referred as the PHY-RF functional split (option 8 in [3]), is considered in this work, although other functional splits between the DU and the CU pool are also possible [9], [10].

By decoupling baseband processing from radio elements, C-RAN can lower the Total Cost of Ownership (TCO) for MNOs. The vaunted benefits of C-RAN are enhanced radio resource (i.e., Radio Frequency (RF) bandwidth) utilization, coordination across multiple cells as well as the multiplexing gain in terms of baseband processing resources. The drawbacks of C-RAN lie in the tight bandwidth and latency requirements imposed on the fronthaul (i.e., the link/network interconnecting CU pools with DUs) where protocols such as Common Public Radio Interface (CPRI) [11] are typically used to carry the In-phase/Quadrature (IQ) samples over optical fiber, which is the most prevalent fronthauling option capable of carrying huge fronthaul bandwidth with low latency.

Nowadays, the baseband processing resources along with the radio resources of cellular networks are not used efficiently since MNOs allocate these resources to their eNBs in such a way as to be able to meet the peak hour traffic demand. Therefore, due to spatially and temporally fluctuating traffic, these resources are underutilized most of the time. Figure 1 is an example of a typical traffic utilization in a carrier/cell² of an eNB in residential and office areas of the operational LTE-A network considered in this study. It can be observed that the carrier load (i.e., traffic demand) varies significantly depending on the area and the time of the day. This traffic imbalance will be even more escalated with the network densification and with an increase in the volume of data traffic. By considering hourly traffic requirements at each cell of each eNB as well as the distance between DUs and CU pools, moving baseband processing of eNBs to the appropriate CU pools could provide

Manuscript received April 1, 2018; revised July 30, 2018; accepted October 8, 2018. Date of publication October 18, 2018; date of current version December 10, 2018. Research leading to these results received funding from the European Unions H2020 Research and Innovation Action under Grant Agreement H2020-ICT-761592 (5G-ESSENCE Project). The associate editor coordinating the review of this paper and approving it for publication was W. Kellerer. (Corresponding author: Davit Harutyunyan.)

The authors are with the Wireless and Networked Systems Unit, FBK CREATE-NET, 38123 Trento, Italy (e-mail: d.harutyunyan@fbk.eu; rriggio@fbk.eu).

Digital Object Identifier 10.1109/TNSM.2018.2876666

¹Notice that the 3GPP [3] terminology with a slight modification is used throughout this article. Specifically, the term CU is used for a BaseBand Unit (BBU) and the term CU pool is used as a BBU pool. However, also other terminologies such as Remote Radio Head (RRH) and BBU pool, Remote Radio Unit (RRU) and Radio Cloud Center (RCC), and Radio Unit (RU) and Digital Unit (DU) for the DU and CU can be found in the technical documents of, respectively, SCF [4], NGFI [5], NGMN [6].

²Notice that each carrier is considered as one cell in this article. The definition of a cell, a carrier and a sector is given in Section III-A.

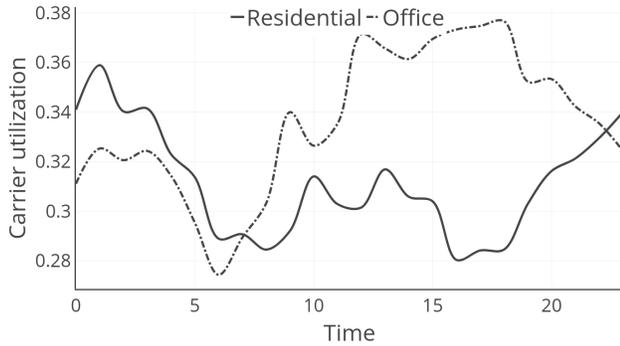


Fig. 1. Traffic demand variation in office and residential areas.

significant multiplexing gain in terms of both radio resources and baseband processing resources, entailing a reduction in TCO of the network.

A sizable body of work has been published on the DU–CU mapping problem in the last few years (see Section II). However, most of the studies make assumptions that would not be feasible/efficient to be applied to real mobile networks. For example, some studies assume that a direct optical line exists from all DUs to a single CU pool, while others assume that a CU pool could be deployed in any area where there is an eNB and they consider only the distance between the DUs and the CU pool while mapping DUs to CUs. One question that needs to be asked, however, is how to migrate from legacy networks to future 5G networks with the C–RAN architecture with minimal investment by exploiting the available mobile network infrastructure and the statistics of the hourly traffic demand per cell?

The contribution of this paper is threefold.

- First, we propose a DU–CU mapping algorithm in order to facilitate MNOs’ transition from their legacy Decentralized RAN (D–RAN) architecture of LTE networks to the C–RAN architecture. The algorithm computes the number of CU pools and determines their locations by taking into account the information about the available inter–eNB transmission links, the distance between the DUs and the potential CU pools, considering the available transport network, and the hourly traffic demand per cell of the mobile network.
- Second, we compare the traffic aggregation gain of C–RAN, which is obtained as a result of the inter–sector intra–carrier traffic aggregation, with the one of D–RAN, which is obtained by activating the intra–sector inter–carrier traffic aggregation feature.
- Third, in order to quantify the economic advantage of using the available infrastructure while transiting to C–RAN, we compare two C–RAN migration scenarios (i.e., DU–CU mappings): with and without using the available infrastructure. In both scenarios, the mapping problems are modeled as Virtual Network Embedding (VNE) problems, formulated and solved employing Integer Linear Programming (ILP) techniques.

The rest of this paper is structured as follows. The related work is discussed in Section II. The substrate and the virtual network models are detailed in Section III. The input sets,

parameters and binary decision variables used in the problem formulations are defined in Section IV. The problem statements and the ILP problem formulations for the intra–sector inter–carrier traffic aggregation problem and the DU–CU mapping problem are presented in, respectively, Sections V and VI. The migration cost computations are presented in Section VII. The numerical results are reported in Section VIII. Finally, Section IX draws the conclusions, pointing out the future work.

II. RELATED WORK

A considerable amount of literature has been published on the DU–CU mapping problem [12]–[20]. An optimization algorithm is presented in [12] for placing CUs over Fixed/Mobile Converged optical networks. The authors formulate an ILP problem, which calculates the required minimum number of CU pools, taking into account only the maximum allowed distance between DUs and their CU pools. The same authors put forward an energy–efficient CU placement algorithm for optical networks in [13], aiming to minimize the Aggregation Infrastructure Power.

Traffic– and interference–aware dynamic DU–CU mapping algorithm is proposed in [14]. Mutual coupling loss, which characterizes Cross–carrier Co–channel Interference (CCI) between DUs, is taken into account in order to find the most optimal DU–CU mapping, which apart from load balancing CUs and minimizing power consumption, would also minimize the CCI between DUs. Semi–static and adaptive DU–CU switching schemes are proposed in [15]. These schemes, although have the same objective of minimizing the required number of CU pools in order to meet the traffic demand at each DU, differ in terms of the DU–CU switching interval. Namba *et al.* [16] elaborate more on the DU–CU switching plan, considering also the signaling load caused by users’ handover while making DU–CU switching decisions. In all aforementioned works, however, the authors do not study how to select CU pool locations and how to assign DUs to CU pools in the case of multiple CU pools in order to get the highest statistical multiplexing gain of resources, since it is important to consider not only spatially and temporally fluctuating traffic, but also the distance between DUs and CU pools.

An energy–efficient DU–CU mapping algorithm is proposed in [17]. Aiming at minimizing the energy consumption at the CU pools, the computing resource requirement of the DUs and the inter–DU traffic exchange are considered for assigning the DUs to the CU pools. While a reconfigurable millimeter wave wireless fronthaul network is used in [21] with the goal of reducing the network–wide power consumption in the CU placement problem, Wang *et al.* [22] propose an energy–efficient scheme for the optical–transport–enabled C–RAN networks by introducing the concept of a virtual base station and enabling baseband processing resource sharing of CUs and line cards of optical line terminators.

A CU placement problem is studied in [23], considering different LTE–A configurations and investigating the impact of different CU centralization levels on both CAPEX and OPEX

in an optical network-supported C-RAN. Chen *et al.* [18] propose a dynamic DU-CU mapping scheme employing a borrow-and-lend approach. The key idea is to migrate the DUs assigned to a highly utilized CU to a less utilized CU, having the objective of maximizing the utilization of every single CU inside the CU pool. However, only one CU pool is considered in these studies without tackling the problem of selecting the number and the locations of the CU pools in order to cater the traffic demand of all the DUs in the network. Moreover, the authors do not consider the CU placement problem simply assuming that all DUs are assigned to a single CU pool via direct links.

An analytical model is derived in [19] for finding the optimal ratio between optical fibers and microwave links, which would reduce the CAPEX required to build the fronthaul network and, at the same time, meet the traffic requirement at each cell site. Holm *et al.* [20] study the problem of minimizing the CAPEX for those MNOs who want to design a mobile network from scratch, adopting the C-RAN architecture. Specifically, the authors study the trade-offs between the multiplexing gain in terms of baseband processing resource, which would increase by assigning more DUs to the same CU pool, and the fronthaul network deployment cost, which would reduce if more CU pools were available for DUs to be associated with. However, the authors make some simplistic assumptions which would be unreasonable to be applied to existing mobile infrastructures. For example, they assume that the fronthaul links are directly connected to the CU pool. They also categorize base stations into two types, office and residential, and assume that the same number of office cells are allocated to the CU pools. Nevertheless, they do not go into the granularity of hourly traffic requirement of each cell in order to better understand the CU composition of which DUs would provide the highest multiplexing gain in terms of both radio resource and baseband processing resource, since the traffic demand in some office and some residential areas might be such that they would not provide high multiplexing gain being assigned to the same CU pool. An optimization problem similar to ours is presented in [24]. A CU placement problem is formulated aiming to find the optimal quantity and locations of CU pools with the goal of minimizing TCO. The migration cost is computed from an optical D-RAN and a Microwave D-RAN to C-RAN, considering greenfield and brownfield optical networks for macrocell, microcell and, nanocell deployments. The main difference compared to our work, however, is that the authors do not consider the real traffic demand at each cell of the traditional D-RAN network, which plays a pivotal role in selecting the number and the locations of the CU pools. While in [25], a CU placement problem is studied for the C-RAN network with Wavelength Division Multiplexing (WDM) aggregation networks. The authors formulate a joint and an independent CU and electronic switch placement problems, considering their placement possibilities in different parts of the Optical Transport Network and the Overlay fronthaul transport network.

The research to date has tended to focus on building future mobile networks from scratch based on the traffic demand. A common characteristic of the aforementioned works is that

none of them has studied how MNOs, owning legacy LTE networks, can upgrade the network by adopting the C-RAN architecture with the minimal cost by employing the available site locations, transmission links, the knowledge of the CU pool candidate locations, and hourly traffic demand per cell. This is a relevant problem since we believe that a few MNOs would be willing to invest a huge amount of money in building C-RAN from scratch, when they can just reuse the legacy infrastructure, and therefore, significantly curtail the required investments in order to deploy C-RAN.

III. NETWORK MODEL

This section details the substrate and the virtual network models. The parameters (e.g., the locations of eNBs, the number of sectors per eNB, the number of carriers/cells per sector and the hourly traffic demand per cell), used in the models, are taken from an operational LTE-A network.

A. Definition of Basic Elements

Before introducing the substrate and the virtual network models, let us provide a more precise definition of the terms carrier, cell and sector.

- *Carrier*: The RF bandwidth (in MHz) owned by the MNO is divided into smaller RF bands called carriers. For example, if an MNO owns 45MHz of RF bandwidth then such bandwidth may be distributed across a 10MHz, a 15MHz, and a 20MHz carrier.
- *Cell*: An eNB can have a different number of cells depending upon its configuration. In this work, we assume that each cell is assigned one and only one carrier, regardless of its RF bandwidth and the LTE band that it belongs to.
- *Sector*: The coverage area of the antenna beam. Each eNB may have one or more sectors, and each sector, in turn, may have one or more cells. In this work, each sector has a maximum of 3 cells. For example, if the total coverage of an eNB is divided into 3 sectors, and each sector has 3 carriers (10MHz, 15MHz and 20MHz) then it means that such eNB has a total of 9 cells.

B. Substrate Network Model

The considered substrate network is a small cluster (5 Km^2) of an operational LTE-A network composed of 26 eNBs and in a total of 209 cells (see Fig. 2). Let $G_s = (N_s, E_s)$ be an *undirected* graph modeling the physical network (i.e., the network to be transformed to C-RAN), where $N_s = N_s^{du} = N_{enb}$ is the set of $m = |N_s|$ DUs/eNBs. Each $n \in N_s^{du}$ DU has a variable number of sectors $|N_{sct}^n|$ in the set of $\{1, 2, 3, 4\}$, each $s \in N_{sct}^n$ sector has a variable number of carriers/cells $|N_{car}^{n,s}|$ in the set of $\{1, 2, 3\}$, while each carrier $c \in N_{car}^{n,s}$ has its maximum supportable downlink throughput $\omega_{s,c}^{tmax}$ (see Table III). A subset of the eNBs $N_s^{cu} \subset N_s$ can be candidates for CU pools. More specifically, the candidate CU pools are selected as follows. The eNBs are sorted in descending order according to the number of inter-eNB transmission links, whereas if some of them have equal number of inter-eNB links, they are sorted according to their total link capacity, and the first

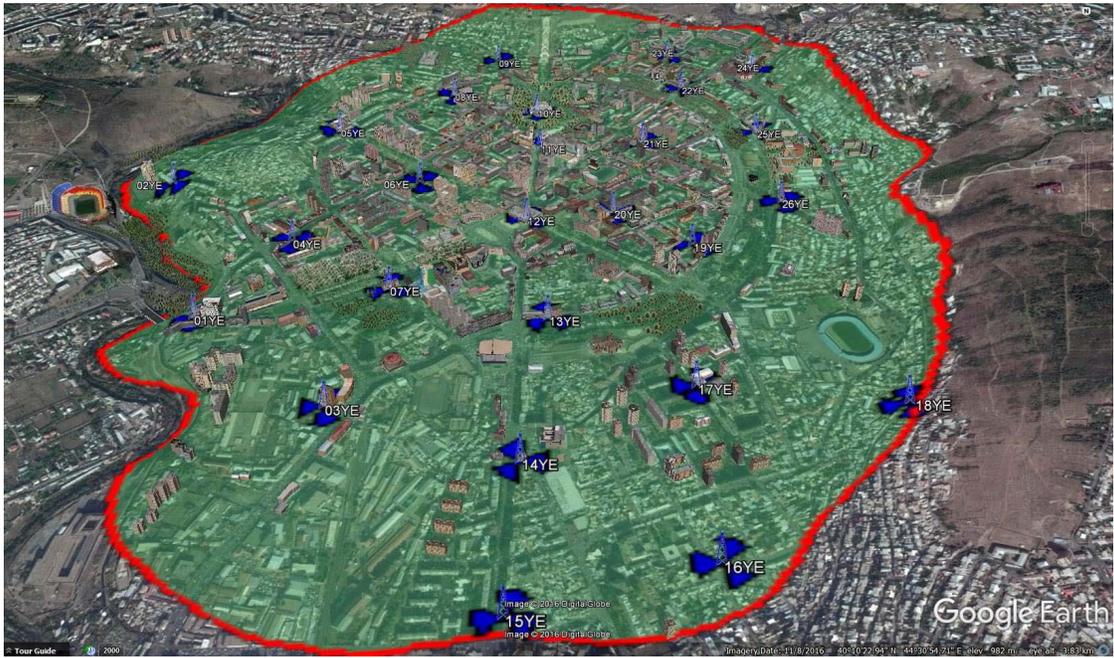


Fig. 2. An operational LTE-A network with 26 eNBs (in total 209 cells) in the city center of Yerevan. Each eNBs is composed of variable number of sectors, which, in turn, is composed of variable number of carriers employing 20MHz, 15MHz and 10MHz RF bandwidths.

four eNBs are considered as CU pool candidates. The required number of CU pool(s) are then picked starting from the first eNB in the sorted list. Thus, the CU pool candidates (i.e., *anchor* eNBs) are the eNBs that interconnect multiple eNBs and serve as a relay for them to transport their signals to the core network. Selecting an *anchor* eNB as a CU pool enables MNOs to exploit the available transport network (i.e., backhaul links) in an efficient manner without having to invest *too* much in building the fronthaul network while migrating to the C-RAN architecture. It is important to mention that, since in the considered small cluster of the mobile network only optical backhaul links are available, we assume that those backhaul links can be used as fronthaul links in the C-RAN architecture.

Each substrate node (i.e., DU/eNB) is also associated with a geographic location $loc(n)$, as x, y coordinates. In order to mimic the real physical topology (i.e., the accurate inter-eNB distance) of the considered LTE-A network, x, y coordinates of the nodes are obtained by converting the real locations (longitude and latitude) of the eNBs. Lastly, let E_s model the set of inter-eNB links of the real network. An edge $e^{nm} \in E_s$ if and only if a connection exists between DUs/eNBs $n, m \in N_s$. The substrate network parameters can be found in Section IV.

C. Virtual Network Model

The considered mobile network (D-RAN) is modeled as a virtual network, which has to be mapped to the substrate network (C-RAN). Let $G_v = (N_v, E_v)$ be an *undirected* graph, where $N_v = N_v^{du} \cup N_v^{cu}$ is the set of $m_1 = |N_v^{du}|$ virtual DUs and $m_2 = |N_v^{cu}|$ virtual CUs. Notice that since in C-RAN an eNB is decomposed into a DU and a CU then each $m \in N_v^{du}$ has to have its CU $m' \in N_v^{cu}$. Therefore, the number of virtual DUs is equal to the number of virtual CUs and is equal to the number of substrate DUs $m_1 = m_2 = m$.

In essence, each $m \in N_v^{du}$ virtual DU has its corresponding $n \in N_s^{du}$ substrate DU, and they have the same number of sectors, carriers per sector and the same location.³ Additionally, at each hour $h \in N_{hr}$, each carrier $c \in N_{car}^{n,s}$ of each sector $s \in N_{set}^n$ of each virtual DU $n \in N_v^{du}$ has its traffic demand $\omega_{v,t}^c(h)$, which is taken from the traffic demand statistics of the considered mobile network.

As opposed to the substrate network model, edges $e^{nm} \in E_v$ in the virtual network request represents the logical mapping between virtual DUs and their corresponding CUs. As an additional constraint, we require each virtual CU to be mapped to one and only one substrate CU pool. Conversely, different virtual CUs from different DUs can be mapped to the same CU pool. This enables advanced interference control algorithms such as Joint Transmission/Reception to be employed [26], which is one of the prominent advantages of C-RAN. The virtual network parameters can be found in Section IV.

D. WDM-PON

We assume that the C-RAN fronthaul network is a WDM Passive Optical Network (PON), which is composed of two main components, Optical Network Unit (ONU) and Optical Line Terminator (OLT), performing electrical to an optical signal and reverse conversion. The former is located at DUs while the latter is located at CU pools.

It is assumed that each CU pool has one OLT rack and one OLT shelf, which has eight OLT access modules. It is also assumed that each OLT access module, through WDM

³Notice that a single notation is used for those parameters that are the same for the substrate and the virtual network (e.g., the location of the DUs, the number of sectors and the number of carriers).

Mux/DeMux, is connected to a passive splitter by a single fiber link that supports four wavelengths each with $\lambda_b = 10\text{Gbps}$ capacity [27]. There is one passive splitter with 16 ports at each eNB site, while the number of ONUs at each eNB depends on the fronthaul bandwidth requirement of each DU at each eNB. For example, if an eNB has three sectors each having one 20MHz cell and one 15MHz cell (overall 6 cells) then the fronthaul bandwidth requirement, regardless of the cell utilization level,⁴ in total would be 7.37Gbps and 5.53Gbps for, respectively, 20MHz cells and 15MHz cells, assuming that CPRI protocol is used and that each cell has 2×2 MIMO antenna configuration. This would require two wavelengths ($\lambda_b = 2$), in order to meet the fronthaul bandwidth demand, which translates to two ONUs since there is one to one mapping between an ONU and a λ_b while four to one mapping between an ONU and an OLT access module.

IV. DEFINITION OF INPUT SETS, PARAMETERS AND BINARY DECISION VARIABLES

In this section, we define the input sets, parameters and binary decision variables used in the substrate and virtual network models of the ILP problem formulations.

A. Input Sets and Parameters

G_s	Substrate network graph.
G_v	Virtual network graph.
E_s	Set of substrate links in G_s .
E_v	Set of virtual links in G_v .
N_{enb}	Set of eNBs in G_s .
N_{sct}^n	Set of sectors of $n \in N_{enb}$ eNB.
$N_{car}^{n,s}$	Set of carriers of $s \in N_{sct}^n$ sector of $n \in N_{enb}$ eNB.
N_s^{cu}	Set of predefined CU pool candidates in G_s .
N_v^{cu}	Set of virtual CUs in G_v .
N_s^{du}	Set of substrate DUs in G_s .
N_v^{du}	Set of virtual DUs in G_v .
N_{hr}	Set of hours in a day.
N_{du}^*	Number of DUs not co-located with a CU pool.
$\omega_{v,t}^c(h)$	Traffic on $c \in N_{car}^{n,s}$ virtual carrier at $h \in N_{hr}$ hour.
$\omega_{s,C}^{t,max}$	Maximum traffic on $C \in N_{car}^{n,s}$ substrate carrier.
$L_{len}(e)$	Length of $e \in E_s$ substrate link [in Km].
L_{len}	Overall length of the substrate links [in Km].
L_{len}^*	Overall length of the built fronthaul links [in Km].
$loc(n)$	Geographical location of n virtual/substrate node.
W_C	Bandwidth of $C \in N_{car}^{n,s}$ carrier [in MHz].
μ_b	Big positive constant.
μ_s	Small positive constant.
λ_b	Lightpath capacity [in Gbps].

B. Binary Decision Variables

Φ_C	Shows if $C \in N_{car}^{n,s}$ carrier of $s \in N_{sct}$ sector of $n \in N_{enb}$ eNB has been selected for traffic aggregation.
----------	--

⁴For a given cell/carrier, fronthaul bandwidth requirement is fixed for the traditional PHY-RF split in the C-RAN architecture and does not depend on the traffic requirement of the considered cell as long as the cell is active [9].

$\Phi_C^{c,t,h}$	Shows if $\omega_{v,t}^c(h)$ traffic of $c \in N_{car}^{n,s}$ virtual carrier at $h \in N_{hr}$ time has been aggregated on $C \in N_{car}^{n,s}$ substrate carrier.
Φ_m	Shows if $m \in N_s^{cu}$ candidate CU pool has been selected as a CU pool.
$\Phi_m^{m'}$	Shows if $m' \in N_v^{cu}$ virtual CU has been mapped to $m \in N_s^{cu}$ substrate CU pool.
$\Phi_n^{n'}$	Shows if $n' \in N_v^{du}$ virtual DU has been mapped to $n \in N_s^{du}$ substrate DU.
$\Phi_e^{e'}$	Shows if $e' \in E_v$ virtual link has been mapped to $e \in E_s$ substrate link.

V. INTRA-SECTOR INTER-CARRIER TRAFFIC AGGREGATION

In this section, we formally state the intra-sector inter-carrier traffic aggregation problem and present its ILP formulation.

A. Problem Statement: Intra-sector Inter-Carrier Traffic Aggregation in D-RAN

In order to show how much the radio resource utilization efficiency of the eNBs can be increased thanks to the statistical multiplexing gain obtained by the C-RAN architecture, we first need to quantify the radio resource utilization level of the eNBs of the current D-RAN architecture over the considered period of time.

In mobile networks, data traffic undergoes significant fluctuations depending upon the location of eNBs and the time of a day. When the traffic demand is low on carriers/cells in traditional LTE networks, intra-sector inter-carrier traffic aggregation, as a feature, can be activated with the goal of reducing power consumption by switching off unnecessary carriers in sectors, and at the same time, ensuring that users traffic demand is met. The upper part of Fig. 3 illustrates an example of the intra-sector inter-carrier traffic aggregation technique. Two three-sector eNBs are considered each having three carriers/cell under each sector. The tables below the eNBs show their corresponding average carrier/cell utilization per RF bandwidth per sector over the given period of time. After performing intra-sector inter-carrier traffic aggregation per eNB, considering the maximum capacity of each carrier/cell (see Table III), we can observe that the algorithm aggregated the traffic of all carriers onto overall six carriers (three carrier per eNB, which are marked in gray in the tables). Notice that this is the minimum number of carriers since, as the name of the employed traffic aggregation technique suggests, the traffic aggregation takes place under each sector separately. As a side effect, this increases the radio resource utilization of the *active* carriers (i.e., the carriers on which the traffic of some other carriers under the same sector have been aggregated). The radio resource utilization of the eNBs is computed after activating the intra-sector inter-carrier traffic aggregation feature, which has been formulated as an ILP problem that can be formally stated as follows:

Given: a small cluster of an operational LTE-A network with the location of eNBs, the sectors per eNB and the

carriers/cells per sector with one-month statistics of their hourly traffic demand.

Find: the number of carriers/cells that need to be active at each eNB such that users hourly traffic demand is satisfied.

Objective: through intra-sector inter-carrier traffic aggregation, curtail the number of active carriers per sector per eNB, thus reducing the power consumption in the network.

B. ILP Formulation: Intra-Sector Inter-Carrier Traffic Aggregation in D-RAN

1) *Objective Function:* The objective function of the intra-sector inter-carrier traffic aggregation techniques is the following:

$$\text{minimize} \sum_{n \in N_{enb}} \sum_{s \in N_{sct}^n} \sum_{C \in N_{car}^{n,s}} \Phi_C V_C \quad (1)$$

where V_C is the cost for using the carrier $C \in N_{car}^{n,s}$ for traffic aggregation. V_C is chosen to be proportional to the size of the carrier bandwidth. Given that the carrier has enough capacity to support the aggregated traffic, the wider is the carrier bandwidth, the more expensive is its cost to be used for traffic aggregation. This is because it is assumed that the wider is the bandwidth of an active carrier, the more is the consumed power. Note that by changing V_C , MNOs can give more/less priority to the carriers that they want to be used in the traffic aggregation.

2) *Constraints:* In order to effectively achieve the aforementioned objective, all the following constraints have to be respected.

$$\sum_{c \in N_{car}^{n,s}} \omega_{v,t}^c(h) \Phi_C^{c,t,h} \leq \omega_{s,C}^{t,max} \quad (2)$$

$$\sum_{c \in N_{car}^{n,s}} \sum_{h \in N_{hr}} \Phi_C^{c,t,h} - \mu_b \Phi_C \leq 0 \quad (3)$$

$$\sum_{c \in N_{car}^{n,s}} \Phi_C^{c,t,h} = 1 \quad \forall n \in N_{enb} \quad (4)$$

Constraint (2) ensures that data traffic on the carriers at which the traffic of other carriers have been aggregated is at most equal to the maximum traffic capacity of the host carriers. Constraint (3) guarantees that if the traffic of any carrier of any sector of any eNB at any time is mapped to a carrier of the same sector at the same time then the host carrier is selected in mappings. Notice that this constraint allows data traffic of a carrier at different times to be mapped to different carriers belonging to the same sector of the same eNB. Lastly, Constraint (4) enforces the traffic of all carriers of all sectors of all eNBs to be mapped/aggregated on the host carriers. In other words, it makes sure that users' traffic demand at any time is met.

VI. DU-CU MAPPING

In this section, we formally state the DU-CU mapping problem and present its ILP formulation.

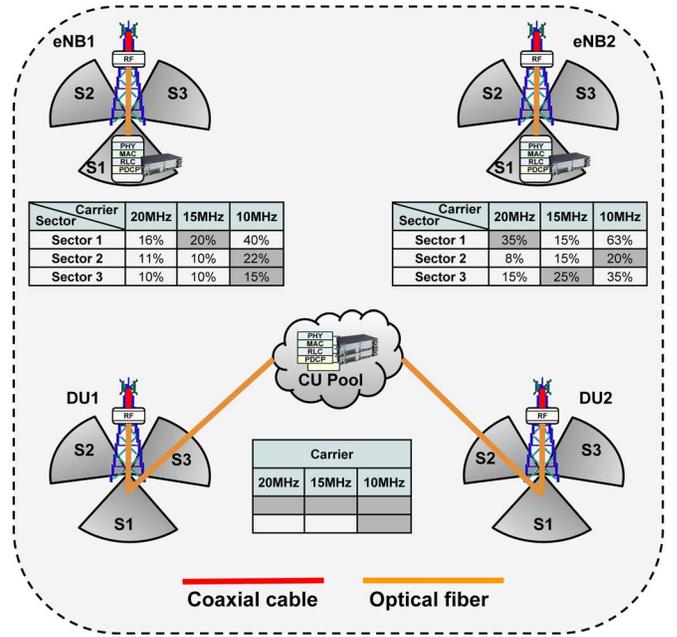


Fig. 3. Examples of the intra-sector inter-carrier traffic aggregation (the upper part) and the inter-sector intra-carrier traffic aggregation (the lower part) techniques.

A. Problem Statement: DU-CU Mapping

In the DU-CU mapping problem, an inter-sector intra-carrier traffic aggregation is taking place, which is the aggregation of the traffic of the carriers that have the same RF bandwidth and belong to the same LTE band. Thus, the traffic on a carrier of a DU/eNB can be aggregated with the traffic on a carrier of another DU/eNB only if the CUs of those DUs are mapped to the same CU pool (i.e., their baseband signal processing is taking place at the same CU pool), those carriers have the same RF bandwidth (i.e., either 20MHz or 15MHz or 10MHz) and they are from the same LTE band. The rationale behind this approach is to guarantee a seamless transition from the D-RAN architecture to the C-RAN architecture, making sure that users experience no channel quality degradation during this transition. The lower part of Fig. 3 illustrates an example of the inter-sector intra-carrier traffic aggregation technique. The entire baseband processing of the eNBs is performed at the same CU pool, harvesting the multiplexing gain in terms of the baseband processing resource as well as the radio resource. As a result, network-wide traffic aggregation is taking place. The traffic utilization of each cell/carrier per eNB before the inter-sector intra-traffic aggregation is reported in the tables (see the upper part of Fig. 3). Notice that, as opposed to the intra-sector inter-carrier traffic aggregation that resulted in six carriers/cells being active at the eNBs (one carrier per eNB per sector), after employing inter-sector intra-carrier traffic aggregation, the traffic of all the carriers is aggregated on four carriers. Thus, the overall number of active carriers/cells is reduced compared to the previous traffic aggregation technique. More details on the inter-sector intra-carrier traffic aggregation is provided in Section VIII-C2. The DU-CU mapping problem can be stated as follows:

Given: a small cluster of an operational LTE-A network with the location of eNBs, the transport network topology with the capacity of each link, the sectors per eNB, the carriers/cells per eNB with one-month statistics of its hourly traffic demand and the candidate locations for CU pools.

Find: the overall number of carriers/cells required to meet users traffic demand, the number and location of CU pools, and DU-CU mappings.

Objective: through the inter-sector intra-carrier traffic aggregation minimize the overall number of carriers/cells, the number of CU pools required to support users hourly traffic demand as well as minimize the fronthaul latency for each virtual link by mapping it onto the shortest substrate path.

B. ILP Formulation: DU-CU Mapping

The available network topology with traffic demand on the carriers of each sector of each eNB is modeled as a virtual network request. Upon arrival of the virtual network request, the substrate network must find the optimal mapping, aiming to minimize the objective function. Efficient mapping of virtual network requests onto a substrate network is known as a VNE problem [28]. The problem is NP-hard and has been studied extensively in [29]–[31]. The embedding process consists of two steps: the node embedding and the link embedding. In the node embedding step, each virtual node (i.e., a virtual DU or a virtual CU) in the request is mapped to a substrate node (i.e., substrate DU or substrate CU pool). In the link embedding step, each virtual link is mapped to a single substrate path. In both steps, nodes and link constraints must be satisfied.

1) *Objective Function:* The DU-CU mapping problem has been formulated as an ILP problem whose objective function (see (5)) aims at minimizing the TCO for those MNOs who own an LTE/LTE-A network and want to migrate to the C-RAN architecture. Specifically, the objective function is composed of three parts:

- The first part aims at minimizing the deployment cost of the CU pools by minimizing the required number of CU pools. The candidate CU pool locations are pre-selected based on the inter-eNB connectivity ranking of each eNB. The more optical transmission links an eNB has with its neighbors, the higher is the likelihood of the eNB to become a CU pool. The rationale behind this approach is to reuse the available transmission links as much as possible.
- The second part aims at minimizing the fronthaul deployment cost by exploiting the available transmission links, and therefore, curtailing the investments required to build the fronthaul network for the C-RAN architecture. It also minimizes the fronthaul delay for each virtual link by mapping it onto the shortest substrate path from the substrate DU, on which the virtual DU has been mapped, to the substrate CU pool that has hosted the baseband processing (virtual CU) of the virtual DU.
- The last part of the objective function minimizes the required number of carriers in order to meet the traffic demand on each carrier of each virtual DU at any time.

This radio resource multiplexing gain is achieved by considering the hourly traffic demand on each carrier/cell and aggregating the traffic of low utilized carriers into a fewer carriers. Notice that contrary to the intra-sector inter-carrier traffic aggregation (see Section V), in this case, an inter-sector intra-carrier traffic aggregation is taking place.

$$\begin{aligned} \text{minimize} \quad & \sum_{m \in N_s^{cu}} V_{cu} \Phi_m + \sum_{e \in E_s} \sum_{e' \in E_v} \mu_s L_{len}(e) \Phi_e^{e'} \\ & + \sum_{n \in N_s^{du}} \sum_{s \in N_{sect}^n} \sum_{C \in N_{car}^{n,s}} W_C V_{mhz} \Phi_C \end{aligned} \quad (5)$$

where V_{cu} is a CU pool built-out cost while V_{mhz} is the annual spectrum fee per MHz (see Table I).

It is worthwhile to note that the second argument of the objective function is very small compared to the rest of the arguments. This is because in the CAPEX savings computation (see Section VII-A1) we consider no fiber rollout cost since in our case the backhaul links of the legacy mobile network are employed as fronthaul links in the C-RAN deployment. Nevertheless, the second argument, although negligible, still exists in order to find the shortest path among the available substrate paths for mapping the virtual links.

2) *Constraints:*

$$\sum_{m' \in N_v^{cu}} \Phi_{m'} - \mu_b \Phi_m \leq 0 \quad \forall m \in N_s^{cu} \quad (6)$$

$$\begin{aligned} \sum_{n \in N_{enb}} \sum_{s \in N_{sect}^n} \omega_{v,t}^c(h) \Phi_C^{c,t,h} &\leq \omega_{s,C}^{t,max} \\ \forall h \in N_{hr}, \quad \forall c = C \in N_{car}^{n,s} \end{aligned} \quad (7)$$

$$\sum_{n \in N_s^{du}} \Phi_n^{n'} = 1 \quad \forall n' \in N_v^{du} \quad (8)$$

$$\sum_{m \in N_s^{cu}} \Phi_m^{m'} = 1 \quad \forall m' \in N_v^{cu} \quad (9)$$

$$\sum_{e \in E_s^{i*}} \Phi_e^{enm} - \sum_{e \in E_s^{i*}} \Phi_e^{enm} = \begin{cases} -1 & \text{if } i = n \\ 1 & \text{if } i = m \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

$$\begin{aligned} \sum_{c \in N_{car}^{n,s}} \Phi_C^{c,t,h} &= 1 \quad \forall n \in N_v^{cu} \\ \forall s \in N_{sect}^n, \quad \forall h \in N_{hr}, \quad \forall C \in N_{car}^{n,s} \end{aligned} \quad (11)$$

Constraint (6) makes sure that a CU pool candidate is selected as a CU pool as long as it has assigned at least one virtual CU. Notice that the case in which $\sum_{m' \in N_v^{cu}} \Phi_{m'} = 0$ and $\Phi_m = 1$ is excluded since the objective function (5) also aims at minimizing the number of CU pools. Constraint (7) guarantees that traffic capacity limit of the host carriers is not violated after traffic aggregation. While Constraints (8) and (9) make sure that each virtual DU and CU are, respectively, mapped to their corresponding substrate DU and CU pool, Constraint (10) enforces for each virtual link $e^{nm} \in E_v$ to be a continuous path established between the pair of the substrate DU and the CU pool on top of which the virtual DU $n \in N_v^{du}$ and the virtual CU $m \in N_v^{cu}$ have been mapped.

In Constraint (10), E_s^{*i} is the set of the fronthaul links that originate from any node and directly arrive at the node $i \in N_s$, while E_s^{i*} is set of the fronthaul links that originate from the node $i \in N_s$ and arrive at any node directly connected to i . Lastly, Constraint (11) ensures that the traffic on all carriers for all sectors of all virtual DUs are mapped; in other words, it is guaranteed that users' traffic demand at each moment is satisfied.

Notice that, although C-RAN has a stringent fronthaul latency requirement, which translates to a maximum admissible length of fronthaul links, which typically ranges between 20 and 40 Km [7], we do not use the fronthaul latency constraint in the ILP formulation since in the considered operational LTE-A network the maximum length of any possible fronthaul route for CPRI flows is far smaller than the mentioned maximum admissible length. Also, notice that we do not have a fronthaul link bandwidth capacity constraint. Initially, given the fronthaul network topology, it is assumed that all the fronthaul links have infinite capacity. This is because our goal is (i) to reuse the available links and (ii) to compute the additional bandwidth, therefore, the overall number OLT access modules and ONUs required in order to meet the network-wide fronthaul bandwidth demand in the C-RAN architecture.

VII. MIGRATION COST COMPUTATION

In this section, we first analyze the TCO savings obtained by migrating from legacy D-RAN to C-RAN. We then compare the migration costs of two C-RAN migration scenarios: the migration cost of C-RAN when the available transport network is exploited and the migration cost of the C-RAN in which the fronthaul infrastructure must be deployed from scratch.

A. From D-RAN to C-RAN Migration Savings Computation

There are many advantages of adopting the C-RAN architecture [7]. These advantages can be mainly divided into two groups: feature- and cost-related advantages. The possibility to exploit advanced interference avoidance/cancellation algorithms such as FeICIC [32] or a coordinated scheduling algorithm [33] are examples of feature-related advantages; whereas, the cost-related advantages are the reduction of CAPEX and OPEX.

In this work, our focus is on the CAPEX and OPEX savings that can be obtained by using the available transport network (i.e., the links that are used to transport the backhaul traffic of the eNBs of legacy mobile networks) and the network knowledge while transiting to C-RAN from legacy D-RAN.

1) *CAPEX Savings*: After mapping the virtual network request onto the substrate network, we start the computation of CAPEX savings (C_{cpx}^{save}), which is computed as the summation of the cost of the available OLT access modules, ONUs and the cost of the available fronthaul transport network (i.e., fiber rollout cost). This is because the mentioned components are reused in the C-RAN deployment. Thus, CAPEX savings

can be computed as follows:

$$C_{cpx}^{save} = \sum_{m \in N_s^{cu}} \Phi_m V_{olt} + |N_{enb}| V_{onu} + L_{len} V_{rol} \quad (12)$$

It is important to mention that while all the eNBs possess a single ONU, only a few of them (i.e., the anchor eNBs) possess an OLT access module.

2) *OPEX Savings*: The OPEX savings (C_{opx}^{save}), obtained as a result of migrating from D-RAN to C-RAN, is the summation of the following two costs: the annual fee for the spectrum usage per MHz per link and the annual cost of renting cell sites. The former savings, thanks to the radio resource multiplexing gain of the C-RAN architecture, is obtained by curtailing the required number of carriers with different RF bandwidths, making sure that mobile data traffic demand is met at any hour. Whereas, the latter is the result of reducing the cell site rent. Indeed, in the case of C-RAN, DUs being compact devices can be easily deployed in the street furniture (e.g., on a lamp post near to the original location of the eNB in order to provide seamless coverage to users), and therefore, the cell site rent can be curtailed on average by a factor⁵ of $\alpha = 0.2$, which is our assumption; while the cell site rent remains the same for the DUs that have a CU pool co-located. Thus, the annual OPEX savings can be computed as follows:

$$C_{opx}^{save} = \sum_{n \in N_{enb}} \sum_{s \in N_{sct}^n} \sum_{C \in N_{car}^{n,s}} (1 - \Phi_C) W_C V_{mhz} + \alpha N_{*}^{du} V_{rent} \quad (13)$$

Notice that, although the OPEX reduction of C-RAN is also contributed by less power consumption compared to D-RAN [7], it is not considered in this study. There are several studies modeling and comparing the power consumption of C-RAN with the one of the traditional D-RAN [22], [24]. Table I summarizes all the cost parameters defined in the equations.

B. Migration Cost Computation of Two C-RAN Migration Scenarios

In order to show the advantage of using the available infrastructure in the C-RAN deployment in terms of CAPEX (OPEX is the same in both C-RAN migration scenarios), we compare two C-RAN migration scenarios: the infrastructure-unaware C-RAN migration and the infrastructure-aware C-RAN migration. In the infrastructure-unaware C-RAN migration, we assume that the fronthaul network of the C-RAN deployment is designed without taking into account the available optical transport network of the legacy mobile network; thus, no transport network exists, and therefore, it should be designed from scratch. Conversely, in the infrastructure-aware C-RAN migration scenario, the fronthaul network is available, which actually is the backhaul network of the legacy D-RAN. Most of the links, however, do not have

⁵The exact value of α depends on the country, the exact location where the DUs are deployed, and on several other factors. However, it is a fact that in the traditional C-RAN architecture (i.e., the PHY-RF functional split) curtails the site rent for the DUs due to their small space requirement for deployment with respect to the that of eNBs [7].

TABLE I
COST ASSUMPTIONS TAKEN FROM [27]

Component	Parameter	Cost [k€]
CU pool build-out	V_{cu}	75
OLT access module, 4*10GPON ports	V_{olt}	6.5
ONU	V_{onu}	0.25
Fiber rollout per Km	V_{rol}	3.8
Annual spectrum fee per MHz and link	V_{mhz}	0.06
Annual cell site rent	V_{rent}	8

enough capacity in order to support the fronthaul bandwidth requirement of C-RAN since there were originally designed to carry the backhaul traffic of the legacy network, which is much smaller compared to the fronthaul bandwidth requirement of C-RAN. Therefore, the capacity of the fronthaul links should be increased by adding the required number of OLT access modules and ONUs.

For the infrastructure-aware C-RAN migration scenario, we use the ILP formulation presented in Section VI-B. In this case, in order to compute the CAPEX (C_{cpx}^{with}), we first need to compute the fronthaul bandwidth $B_{fh}(n)$ requirement at each $\forall n \in N_{enb}/N_{v}^{du}$ eNB by using the equation in [34]:

$$B_{fh}(n) = \sum_{s \in N_{sct}^n} \sum_{c \in N_{car}^{n,s}} 2f_s(c)N_o N_Q N_R \quad \forall n \in N_{v}^{du} \quad (14)$$

where 2 accounts for the complex nature of the IQ samples, while the other parameters are reported in Table II. We can now compute the CAPEX:

$$C_{cpx}^{with} = \sum_{n \in N_{v}^{du}} \left(\left\lceil \frac{B_{fh}(n)}{\lambda_b} \right\rceil - 1 \right) V_{onu} + \sum_{m \in N_s^{cu}} \left(\left\lceil \frac{\sum_{m' \in N_{v}^{cu}} B_{fh}(m'_*) \Phi_m^{m'}}{4\lambda_b} \right\rceil - 1 \right) \Phi_m V_{olt} \quad (15)$$

where m'_* is the corresponding virtual DU/eNB of m' virtual CU. The first and the second fractions calculate the additional number of, respectively, ONUs in DUs and OLT access modules in CU pools required in order to support the network-wide fronthaul bandwidth demand. It is important to mention that, since in this case, the optical transport network is available, there is no fiber rollout cost in the CAPEX computation. Notice that we do not consider the cost of building CU pools since, as we will see in Section VIII, the number of CU pools (not the candidate CU pools) after embedding is the same in both migration scenarios.

For the infrastructure-unaware C-RAN migration scenario, we use the ILP formulation presented in Section VI-B with a slight modification in the objective function (5), while keeping the constraints the same. The objective function for this scenario is the following:

$$\begin{aligned} \text{minimize} \quad & \sum_{m \in N_s} V_{cu} \Phi_m + \sum_{e \in E_s} \sum_{e' \in E_v} V_{rol} L_{len}(e) \Phi_e^{e'} \\ & + \sum_{n \in N_s^{du}} \sum_{s \in N_{sct}^n} \sum_{C \in N_{car}^{n,s}} W_C V_{mhz} \Phi_C \quad (16) \end{aligned}$$

TABLE II
PARAMETERS FOR CALCULATING FRONTHAUL DATA RATES

Parameter	Symbol	Value
Bandwidth	B	[20MHz, 15MHz, 10MHz]
Sampling frequency	f_s	[30.72MHz, 23.04MHz, 15.36MHz]
Quantization bits per I/Q	N_Q	10
RX antennas	N_R	2
Oversampling factor	N_o	2

Notice that as opposed to (5), here the second argument is not negligible and accounts for the fiber rollout cost of the fronthaul network, since in this case it is assumed that the fronthaul network is not given. Notice also that, as opposed to the infrastructure-aware C-RAN migration scenario, in which the candidates for CU pools are predefined, in this scenario, any eNB is a potential candidate for a CU pool. Whereas like in the first scenario, also here it is initially assumed that the mapped links have infinite capacity, meaning that there are enough number of OLT access modules and ONUs to host the required fronthaul bandwidth on any fronthaul link. However, the exact fronthaul capacity of the mapped fiber links, which is, the exact number of OLT access modules and ONUs is computed by considering the fronthaul requirements of the mapped virtual links.

Considering that the fronthaul bandwidth requirement ($B_{fh}(n)$) for each DU $n \in N_{v}^{du}$ is the same in both scenarios, the CAPEX (C_{cpx}^{wout}) in the infrastructure-unaware C-RAN migration scenario can be computed as follows:

$$\begin{aligned} C_{cpx}^{wout} = & \sum_{n \in N_{v}^{du}} \left\lceil \frac{B_{fh}(n)}{\lambda_b} \right\rceil V_{onu} \\ & + \sum_{m \in N_s} \left\lceil \frac{\sum_{m' \in N_{v}^{cu}} B_{fh}(m'_*) \Phi_m^{m'}}{4\lambda_b} \right\rceil V_{olt} \\ & + L_{len}^* V_{rol} \quad (17) \end{aligned}$$

The first argument computes the cost of ONUs, the second argument computes the cost of OLT access modules, while the last argument computes the fiber rollout cost.

VIII. EVALUATION

The goal of this section is to compare the presented traffic aggregation algorithms and the C-RAN migration scenarios. We shall first describe the simulation environment used in our study. We will then report on the outcomes of the numerical simulations carried out using a discrete event simulator implemented in MATLAB.

A. Simulation Parameters

A small cluster (26 eNBs deployed on rooftops with 209 cells in total) of an operational LTE-A mobile network in the city center of Yerevan is considered in our simulations (see Fig. 2). The cluster provides mobile coverage in the area of 5Km². The average number of RRC connected users per eNB varies between 450 and 2000 depending on the location of eNBs, their carriers and the time of the day. Whereas, the

TABLE III
LTE-A NETWORK PARAMETERS

Band.	Av. modul.	Max. DL traff.	RF config.	Over.
20 MHz	16 QAM	100 Mbps	2×2 MIMO	0.25 %
15 MHz	16 QAM	75 Mbps	2×2 MIMO	0.25 %
10 MHz	16 QAM	50 Mbps	2×2 MIMO	0.25 %

number of sectors per eNB, as well as the number of carriers/cells per sector, vary in the set of, respectively, {1, 2, 3, 4} and {1, 2, 3}, depending upon the need for providing either coverage or extra capacity in the given area. Three LTE carriers, 20MHz, 15MHz and 10MHz, are used in the network, and only optical fiber links are used to connect the eNBs to the core network. This is a representative of a dense urban mobile network deployment scenario.

The maximum downlink traffic $\omega_{s,C}^{t_{max}}$ per carrier/cell $C \in N_{car}^{n,s}$, which has either 20MHz or 15MHz or 10MHz RF bandwidth, is derived considering 2×2 MIMO antenna configuration in every sector, the average modulation order, which is assumed to be 16 QAM since eNBs are deployed densely, and 25% overhead such as PDCCH, reference signal, synchronization signals, PBCH and channel coding. Whereas, hourly downlink traffic demands per carrier $\omega_{v,t}^c(h)$ is derived from the traffic demand statistics of the RRC connected users in the considered LTE-A network. Table III shows the parameters used to derive the maximum downlink throughput per carrier. The simulations are conducted for each day separately, and the reported results are the average of 30 simulations (one month) with 95% confidence intervals.

B. Simulation Results

As it has been mentioned, the objective function of the intra-sector inter-carrier traffic aggregation problem (see formula (1)) aims at curtailing the number of active carriers by aggregating the traffic of low-utilized carriers on fewer carriers, and therefore, enabling MNOs to switch off the unnecessary carriers. For MNOs, the effect of minimizing this objective function (i.e., the activation of the intra-sector inter-carrier traffic aggregation feature) is the OPEX savings obtained in the power consumption bills. However, we will look at this objective from the perspective of the load at eNBs, since curtailing the number of active carriers increases the utilization of the remaining active carriers after switching off the low-utilized carries.

Figure 4 displays the average traffic load per eNB over 24 hours averaged for one month. The traffic load at each hour at each eNB is the summation of the carrier loads of that eNB. It can be observed that by activating the intra-sector inter-carrier traffic aggregation feature, the load of the active carriers per sector can be significantly increased. This is a consequence of aggregating the traffic of the low-utilized and already inactive carriers to the active carriers. It can also be observed, however, that there is still room for increasing the load at eNBs, and therefore, resulting in a more efficient carrier utilization. Towards this end, we adopt the C-RAN architecture and study different scenarios for migration from legacy D-RAN to C-RAN.

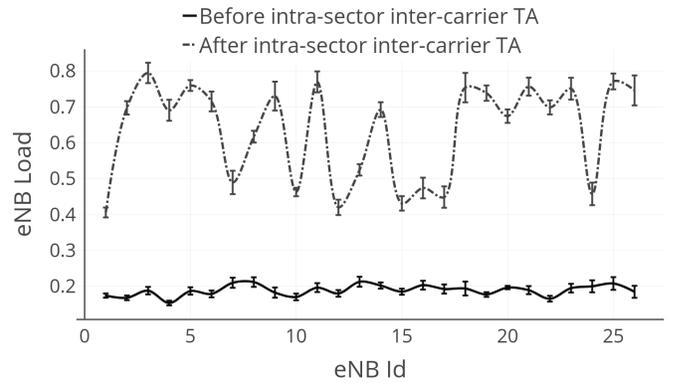
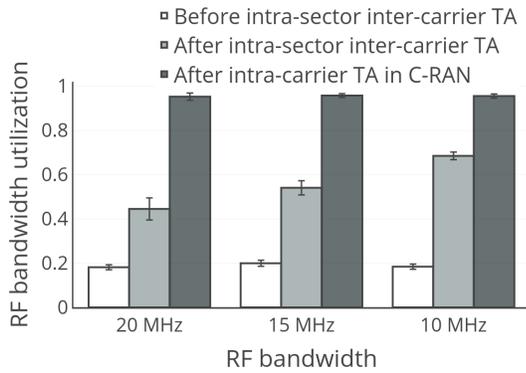


Fig. 4. Traffic load per eNB.

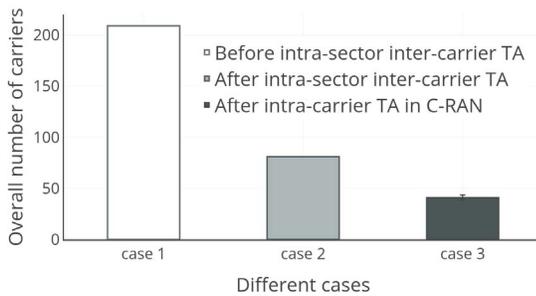
Figure 5 illustrates the RF bandwidth utilization, the overall number of carriers, the number of carriers per RF bandwidth and the execution time of the considered traffic aggregation before the intra-sector inter-carrier traffic aggregation (*case 1*), after the intra-sector inter-carrier traffic aggregation (*case 2*) and after the inter-sector intra-carrier traffic aggregation in C-RAN (*case 3*). We can observe that in the considered period of time the RF bandwidth utilization barely reaches 20% at all the considered carriers before the intra-sector inter-carrier traffic aggregation (see Fig. 5a). Although the intra-sector inter-carrier traffic aggregation feature increases the utilization of the active carriers, as we have also seen in Fig. 4, it can still be significantly increased since 20MHz, 15MHz and 10MHz carriers are underutilized by, respectively, 56%, 46% and 31% with the maximum 5% of difference from the mean values in their confidence intervals. We can observe that after adopting the C-RAN architecture the utilization of all the carriers with different RF bandwidths is increased up to approximately 95%.

The overall number of carriers for all the cases is depicted in Fig. 5b. Notice that, although the number of active carriers after intra-sector inter-carrier traffic aggregation (*case 2*) is reduced by 61%, this just curtails the power consumption bills without exempting the MNO from paying the fee for using the spectrum of temporarily unused carriers. Whereas, after adopting the C-RAN architecture (*case 3*), the number of active carriers is curtailed by 80%, which not only reduces the power consumption but also significantly lowers the overall fee for using the spectrum. It is worthwhile to note that, regardless of different traffic requirements on the carriers in different days, there is no change in the overall number of carriers after the intra-sector inter-carrier traffic aggregation (i.e., the confidence interval is zero in *case 2*). This stems from the fact that unlike the inter-sector intra-carrier traffic aggregation in C-RAN (*case 3*), the intra-sector inter-carrier traffic aggregation technique aggregated the traffic of carriers per sector individually.

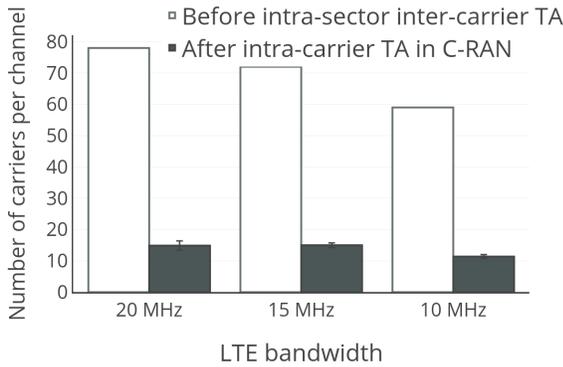
Figure 5c shows the number of carriers per RF bandwidth for *case 1* and *case 3*. We remind the reader that each sector has variable number of carriers (i.e., 20MHz, 15MHz or 10MHz RF bandwidth, maximum one carrier from each RF bandwidth). As expected, C-RAN curtails the number



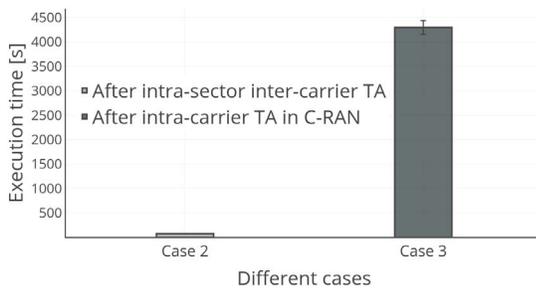
(a) RF bandwidth utilization.



(b) Overall number of carriers.



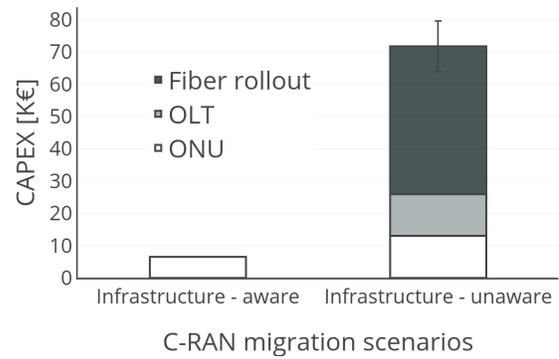
(c) Number of carriers per RF bandwidth.



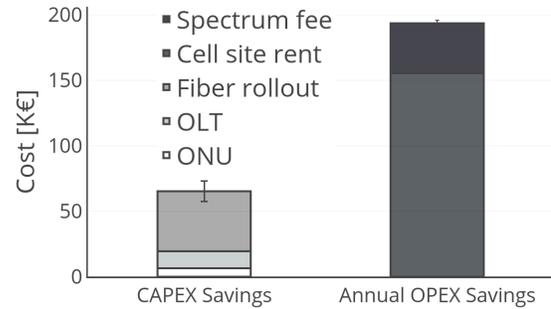
(d) Execution time.

Fig. 5. RF bandwidth utilization, overall number of carriers, number of carriers per RF bandwidth and execution time for the considered cases.

of carriers in all RF bandwidths and distributes the overall traffic more uniformly across the different RF bandwidths. Lastly, Fig. 5d compares the time required to execute the



(a) CAPEX of two C-RAN migration scenarios.



(b) TCO savings.

Fig. 6. CAPEX in infrastructure-aware and infrastructure-unaware C-RAN migration scenarios, and TCO savings in the former scenario.

intra-sector inter-carrier traffic aggregation (*case 2*) and the time to execute the inter-sector intra-carrier traffic aggregation in C-RAN (*case 3*). It can be seen that in *case 2* the execution time is extremely shorter compared to the execution time in *case 3*. This is justified by the fact that, as opposed to *case 3* in which a global inter-sector intra-carrier traffic aggregation is taking place, in *case 2* the intra-sector inter-carrier traffic aggregation is confined within each sector of each eNB. Even though one could suggest resorting to heuristics to address large instances of the problem, in our opinion, MNOs may agree to wait even a week in order to find the optimal mapping solution for their CU-DU mapping problem rather than find a suboptimal solution very fast. This is because there is no need for performing this kind of DU-CU mappings (inter-sector intra-carrier traffic aggregation) very frequently.

In order to get an insight into what is the advantage of using an infrastructure-aware C-RAN migration and how much the MNO can gain in terms of CAPEX and OPEX savings while migrating from their D-RAN to C-RAN, let us analyze Fig. 6. Figure 6a compares CAPEX of two C-RAN migration scenarios: infrastructure-aware and infrastructure-unaware migration from D-RAN to C-RAN. We remind the reader that the former scenario implies that the available infrastructure of D-RAN (e.g., optical backhaul links, OLT access modules, ONUs, etc.) is used in the C-RAN. Whereas in the latter scenario, it is assumed that only the eNB site locations are available and the fronthaul infrastructure should be designed from scratch. Notice that the CU pool build-out cost is not

considered in CAPEX savings since in both C-RAN cases the ILP-based DU-CU mapping algorithms select two CU pools to be built in order to support centralized signal processing of all eNBs. We can observe that around seven times more capital investment is required if the available infrastructure is not considered the C-RAN deployment. The greatest share of this CAPEX constitutes the cost of building the optical fronthaul network. Additionally, we can observe that, as opposed to the infrastructure-unaware migration, the infrastructure-aware migration requires no OLT deployment cost in this particular network setup since after the DU-CU mapping the capacity of the OLTs available in the legacy LTE-A network is enough to meet the fronthaul traffic demand in C-RAN. It is worthwhile to note that while there is no change in the OLT and the ONU shares in CAPEX for both migration scenarios, which essentially means that the required number of ONUs and OLTs is the same regardless of traffic variation in different days at different carriers, the fiber rollout cost share in CAPEX changes due to different fronthaul link mappings in different days. Therefore, the infrastructure-unaware fronthaul network design requires careful analysis of the traffic pattern change at different days at the eNBs.

Let us now analyze how much the MNO can gain from the considered small part of the LTE-A network in terms of CAPEX and annual OPEX savings (see Fig. 6b) as a result of employing the available network infrastructure. CAPEX savings come from three components: fiber rollout cost, ONUs and OLT access modules. As it is expected, the greatest part of CAPEX savings is obtained from the fronthaul deployment cost. The cost of OLT access modules and ONUs is significantly lower compared to the fiber rollout cost. Whereas, annual OPEX savings comes from reduced spectrum fee and cell site rent. It can be observed that in this particular scenario the annual cell site rent of the curtailed cell sites is around four times higher than the annual spectrum fee of the curtailed carriers. It can also be observed that apart from the fiber rollout cost variation, there is a negligible OPEX variation also in the spectrum fee. This is due to the fact that in the C-RAN architecture, after the inter-sector intra-carrier traffic aggregation, the overall number of carriers and their RF bandwidth depends on the traffic demand per carrier before the aggregation, which varies in different days.

C. Discussion

1) *Intra-Sector Inter-Carrier Traffic Aggregation in D-RAN*: The intra-sector inter-carrier traffic aggregation feature has several pros. For instance, if the traffic demand on the carriers is low then by activating this feature in current LTE/LTE-A networks the overall power consumed by carriers can be reduced by deactivating the unused carriers (see the upper part of Fig. 3). Whereas if the required traffic on the carriers is high, the inactive carriers can then be reactivated, meeting the traffic demand and providing the possibility of carrier aggregation, which is one of the prominent features of LTE-A technology. Nonetheless, the cons of the intra-sector inter-carrier traffic aggregation lay in the fact that the overall OPEX can be reduced only in terms of power consumption

cost for the period in which some of the carriers are inactive. Moreover, by aggregating the intra-sector inter-carrier traffic, the host carriers may still be underutilized (see Fig. 4).

2) *Inter-Sector Intra-Carrier Traffic Aggregation in C-RAN*: The inter-sector intra-carrier traffic aggregation in C-RAN (*case 3*) has an important advantage over the intra-sector inter-carrier traffic aggregation (*case 2*). The carriers at different frequency bands (e.g., Band 3, Band 5, Band 7) with different RF bandwidths (e.g., 10MHz, 15MHz and 20MHz) have their peculiarities. For example, high-frequency bands (e.g., Band 7, 2620-2690MHz) are more beneficial for LTE networks construction in the regions with a large population where high speed of data transfer is required. Whereas, LTE network deployments in the low-frequency bands (e.g., Band 5, 869-894MHz) is very appealing from the cost viewpoint and is ideal for the regions with low population (e.g., suburban areas, villages). The pros of such a deployment are costs, better penetration inside buildings and coverage of large territories. Thus, some users may experience performance degradation while forcing them to handover from one carrier to another under the same sector. Whereas, this problem does not exist in *case 3*, since in this case an inter-sector intra-carrier traffic aggregation is taking place.

3) *Infrastructure-Aware C-RAN Migration*: In Section VIII, we have seen that significant TCO savings can be obtained by exploiting the legacy mobile network infrastructure while migrating to C-RAN. In this work, however, just a small part of an operational LTE-A network is considered. The TCO savings will increase with the increase of the network size. More accurate TCO savings estimation depends not only on the size of the network but also on the geographical areas (e.g., urban, suburban, rural), the link types of the transport network, population, traffic demand and on several other factors. Thus, although the considered scenario the obtained results cannot be extrapolated to all kind of deployments, it is, however, a good example of an urban mobile network deployment.

IX. CONCLUSION

Recently, C-RAN has come to the fore as a promising way to use the precious baseband processing and radio resources elastically and efficiently based on the actual need and, through better inter-cell coordination, overcome all the possible performance degradation that may be entailed by network densification.

In this paper, we propose an ILP-based algorithm to allow MNOs to transit to the C-RAN architecture with minimal investment by employing the available infrastructure in an efficient manner, and therefore, curtailing the required investments to adopt the C-RAN architecture. In order to quantify the TCO saving, we compare infrastructure-aware and infrastructure-unaware C-RAN migration scenarios, showing the significant savings that can be obtained by using the former approach.

Since only the classical C-RAN functional split is considered in this study, the fronthaul network itself does not provide any multiplexing gain. As a future work, we are planning

to extend the problem formulation and, based on the traffic demand statistics per cell and the availability of the transmission links, consider the possibility of flexible functional split at the RANs and study the impact of functional splits to the fronthaul multiplexing gain.

REFERENCES

- [1] *Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper*, Cisco, Vis. Netw. Index, San Jose, CA, USA, 2017.
- [2] Y. Q. Bian. (2014). *Small Cells Big Opportunities*. [Online]. Available: <https://www.huawei.com/ilink/en/download/HW330984>
- [3] “Technical specification group radio access network; study on new radio access technology; radio access architecture and interfaces, V2.0.0,” 3GPP, Sophia Antipolis, France, Rep. TR 38.801, 2017.
- [4] “Small cell virtualization functional splits and use cases, V7.0,” Small Cell Forum, London, U.K., Rep. 159.07.02, 2016.
- [5] J. Huang and Y. Yuan. *White Paper of Next Generation Fronthaul Interface, Version 1*. Accessed: Oct. 31, 2016. [Online]. Available: <https://labs.chinamobile.com/cran>
- [6] “Further study on critical C-RAN technologies, V1.0,” Next Gener. Mobile Netw. Alliance, Frankfurt, Germany, Rep., 2015.
- [7] K. Chen and R. Duan, “C-RAN the road towards green RAN, version 2.5,” Beijing, China, China Mobile Res. Inst., White Paper, 2011.
- [8] A. Checko *et al.*, “Cloud ran for mobile network—A technology overview,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, 1st Quart., 2015.
- [9] U. Dötsch *et al.*, “Quantitative analysis of split base station processing and determination of advantageous architectures for LTE,” *Bell Labs Tech. J.*, vol. 18, no. 1, pp. 105–128, Jun. 2013.
- [10] D. Harutyunyan and R. Riggio, “Flex5G: Flexible functional split in 5G networks,” *IEEE Trans. Netw. Service Manag.*, vol. 15, no. 3, pp. 961–975, Sep. 2018.
- [11] “Common public radio interface, interface specification, V7.0,” CPRI, Bengaluru, India, Rep., 2015.
- [12] N. Carapellese, M. Tornatore, and A. Pattavina, “Placement of base-band units (BBUs) over fixed/mobile converged multi-stage WDM-PONs,” in *Proc. IEEE ONDM*, Brest, France, 2013.
- [13] N. Carapellese, M. Tornatore, and A. Pattavina, “Energy-efficient base-band unit placement in a fixed/mobile converged WDM aggregation network,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 8, pp. 1542–1551, Aug. 2014.
- [14] D. Zhu and M. Lei, “Traffic and interference-aware dynamic BBU-RRU mapping in C-RAN TDD with cross-subframe coordinated scheduling/beamforming,” in *Proc. IEEE ICC*, Budapest, Hungary, 2013, pp. 884–889.
- [15] S. Namba, T. Warabino, and S. Kaneko, “BBU-RRH switching schemes for centralized RAN,” in *Proc. IEEE CHINACOM*, Kunming, China, 2012, pp. 762–766.
- [16] S. Namba, T. Matsunaka, T. Warabino, S. Kaneko, and Y. Kishi, “Colony-RAN architecture for future cellular network,” in *Proc. IEEE FutureNetw*, Berlin, Germany, 2012, pp. 1–8.
- [17] B. J. R. Sahu, S. Dash, N. Saxena, and A. Roy, “Energy efficient BBU allocation for green C-RAN,” *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1637–1640, Jul. 2017.
- [18] Y.-S. Chen, W.-L. Chiang, and M.-C. Shih, “A dynamic BBU-RRH mapping scheme using borrow-and-lend approach in cloud radio access networks,” *IEEE Syst. J.*, vol. 12, no. 2, pp. 1632–1643, Jun. 2018.
- [19] R. Al-Obaidi, A. Checko, H. Holm, and H. Christiansen, “Optimizing cloud-RAN deployments in real-life scenarios using microwave radio,” in *Proc. IEEE EuCNC*, Paris, France, 2015, pp. 159–163.
- [20] H. Holm, A. Checko, R. Al-Obaidi, and H. Christiansen, “Optimal assignment of cells in C-RAN deployments with multiple BBU pools,” in *Proc. EuCNC*, Paris, France, 2015, pp. 205–209.
- [21] R. Riggio, D. Harutyunyan, A. Bradai, S. Kuklinski, and T. Ahmed, “SWAN: Base-band units placement over reconfigurable wireless fronthauls,” in *Proc. CNSM*, Montreal, QC, Canada, 2016, pp. 28–36.
- [22] X. Wang *et al.*, “Green virtual base station in optical-access-enabled cloud-RAN,” in *Proc. ICC*, London, U.K., 2015, pp. 5002–5006.
- [23] A. Asensio, P. Saengudomlert, M. Ruiz, and L. Velasco, “Study of the centralization level of optical network-supported cloud RAN,” in *Proc. ONDM*, Cartagena, Spain, 2016, pp. 1–6.
- [24] S. S. Lisi, A. Alabbasi, M. Tornatore, and C. Cavdar, “Cost-effective migration towards C-RAN with optimal fronthaul design,” in *Proc. ICC*, Paris, France, 2017, pp. 1–7.
- [25] F. Musumeci *et al.*, “Optimal BBU placement for 5G C-RAN deployment over WDM aggregation networks,” *J. Lightw. Technol.*, vol. 34, no. 8, pp. 1963–1970, Apr. 15, 2016.
- [26] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS Long Term Evolution: From Theory to Practice*. Chichester, U.K.: Wiley, 2011.
- [27] A. A. W. Ahmed, K. Chatzimichail, J. Markendahl, and C. Cavdar, “Techno-economics of green mobile networks considering backhauling,” in *Proc. Eur. Wireless*, Barcelona, Spain, 2014, pp. 1–6.
- [28] A. Fischer, J. F. Botero, M. T. Beck, H. De Meer, and X. Hesselbach, “Virtual network embedding: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1888–1906, 4th Quart., 2013.
- [29] Z. Despotovic *et al.*, “VNetMapper: A fast and scalable approach to virtual networks embedding,” in *Proc. IEEE ICCCN*, Shanghai, China, 2014, pp. 1–6.
- [30] M. Chowdhury, M. R. Rahman, and R. Boutaba, “ViNEYard: Virtual network embedding algorithms with coordinated node and link mapping,” *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 206–219, Feb. 2012.
- [31] R. Riggio, A. Bradai, D. Harutyunyan, T. Rasheed, and T. Ahmed, “Scheduling wireless virtual networks functions,” *IEEE Trans. Netw. Service Manag.*, vol. 13, no. 2, pp. 240–252, Jun. 2016.
- [32] H. Hu, J. Weng, and J. Zhang, “Coverage performance analysis of FeICIC low-power subframes,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5603–5614, Aug. 2016.
- [33] O. D. Ramos-Cantor, J. Belschner, G. Hegde, and M. Pesavento, “Centralized coordinated scheduling in LTE-advanced networks,” *EURASIP J. Wireless Commun. Netw.*, vol. 2017, no. 1, p. 122, 2017.
- [34] D. Wubben *et al.*, “Benefits and impact of cloud computing on 5G signal processing: Flexible centralization through cloud-RAN,” *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 35–44, Nov. 2014.



Davit Harutyunyan received the bachelor's and master's degrees (Hons.) in telecommunication engineering from the National Polytechnic University of Armenia in 2011 and 2015, respectively. He is currently pursuing the Ph.D. degree with the University of Trento. He was a Radio Access Network Optimization Engineer with Orange Armenia. He has published nine papers in internationally recognized journals/conferences. Since 2015, he has been a member of the Wireless and Networked Systems Research Unit, FBK CREATE-NET. His main research interests include cellular networks and software-defined mobile networking. He was a recipient of the IEEE CNSM 2017 Best Student Paper Award.



Roberto Riggio (SM'16) is currently the Head of the Wireless and Networked Systems Research Unit, FBK CREATE-NET. He is the Creator of 5G-EmPOWER the first network operating system for 5G networks, which is currently used in several 5GPPP Projects. He has one granted patent, 81 papers published in internationally refereed journals and conferences, and has generated over 1.5M in competitive funding. He has extensive experience in the technical and project management of European and industrial projects and he is currently a Project Manager of the H2020 Vital Project. His research interests include network operating systems for wireless and mobile networks, active network slicing in 5G systems, and distributed management and orchestration of network services. He was a recipient of several awards, including the IEEE INFOCOM 2013 Best Demo Award, the IEEE ManFI 2015 Best Paper Award, and the IEEE CNSM 2015 Best Paper Award. He serves in the TPC/OC of leading conferences in the networking field and an Associate Editor for the *International Journal of Network Management* (Wiley), *Wireless Networks* (Springer), and the *IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT*. He is the Co-Founder of the IEEE 5GMan Workshop. He is a member of ACM.