SWAN: Base–Band Units Placement over Reconfigurable Wireless Front–Hauls

Roberto Riggio*, Davit Harutyunyan*, Abbas Bradai§, Slawomir Kuklinski§, Toufik Ahmed¶
*CREATE-NET, Italy; Email: rriggio,dharutyunyan@create-net.org
§XLIM Institute, University of Poitiers, Poitiers, France; Email: abbas.bradai@univ-poitiers.fr
§Orange Polska, Warsaw, Poland; Email: slawomir.kuklinski@orange.com
¶CNRS-LaBRI, University of Bordeaux, Bordeaux, France; Email: tad@labri.fr

***Abstract*—Small–cells are rapidly emerging as the mobile operators' choice to provide additional capacity in current and future mobile networks. However, in order to fully deliver on their promises, small–cells need to address severe interference control and coordination challenges. By centralizing base–band processing in large high–volume computing infrastructures, Cloud–RAN can effectively enable advanced coordination features for dense small–cells deployments. Unfortunately, Cloud–RAN tight bandwidth and latency requirements have made optical fiber the most common solution for the links interconnecting remote radio heads (RRHs) with the base band units (BBUs), i.e. the fronthaul. Recent advances in microwave communications are making wireless fronthauls a viable option especially in dense urban environments where fiber fronthauls could be too rigid for accommodating highly dynamic traffic patterns. In this paper, we provide a novel formulation for the BBU Placement problem where BBU pools are placed at the edges of the network, possibly co–located with macro–cells, and a reconfigurable wireless fronthaul is used in order to provide RRHs with connectivity. To the best of our knowledge this is the first work to tackle the BBU placement problem over a reconfigurable substrate network with mmWave links. We also propose a BBU Placement heuristics, and we evaluate it using a numerical simulator.**

***Index Terms*—Mobile Networks, Cloud RAN, BBU Placement, Wireless Fronthaul, mmWave, Mesh Network**

I. Introduction

Mobile data traffic has been growing exponentially over the last few years. Cisco's Visual Network Index shows that the mobile traffic increased dramatically in 2015, with a growth ranging from 52% (in Western Europe) up to 117% (in Middle East and Africa). Overall mobile data traffic is expected to grow to 30.6 exabytes per month by 2020, an eight–fold increase over 2015 [1]. This trend is forcing mobile network operators (MNOs) to perform costly network upgrades in a time when the average revenue per user is decreasing.

In a traditional mobile network, the radio and base–band processing units, which compose base stations, are placed in close proximity. This is done to mitigate the high signal losses associated with the RF cables that are typically used for their interconnection. In order to circumvent these limitations MNOs moved to the Distributed RAN architecture (D–RAN), where RF cables are replaced with optical fiber and a digital interface is used to carry the IQ (in–phase/quadrature) signals between the base–band units (BBU) and the radio elements, named Remote Radio Head (RRH). Cloud RAN (C–RAN) has emerged as a solution capable of reducing the deployment and operational costs of mobile networks while at the same time enhancing network capacity, coverage and power consumption. C–RAN [2] achieves such goals by consolidating BBUs in large high–volume computing infrastructure, named BBU pools, and by sharing them among multiple sites.

BBU pools can run on a number of general purpose off-the-shelf servers deployed in one or multiple centralized locations with virtualization capabilities. This approach seems very promising to MNOs looking to introduce new services without impacting all the system components. C–RAN can be used to cope with spikes in traffic demand by dynamically deploying additional BBU pools when and where needed. Furthermore, C–RAN can improve the experience of users at the edges of the cell or in dense environment by implementing advanced Inter-cell interference coordination (ICIC) techniques.

C–RAN main drawbacks lie in the tight bandwidth and latency requirements imposed on the fronthaul (i.e. the links interconnecting BBUs with RRHs) where protocols like the Common Public Radio Interface (CPRI) [3] are typically used to carry the IQ samples. In fact, being a digital representation of a very high frequency waveform, this protocol requires very high data rates; for example a 20 MHz LTE FDD channel using a $2x2$ MIMO antenna configuration can result in a CPRI rate of ≈ 2.5 Gbps. As a result, providing fiber–based CPRI links for the tens or hundreds of small cells that are expected to be deployed in dense urban scenarios can simply be not a viable option for mobile operators. This consideration is made even more true by the fact that, being characterized by a coverage radius in the order of hundreds of meters, small cells can suffer of severe under utilization in case of changing traffic patterns.

A particularly interesting solutions combining low deployment and operational costs with the benefits of C–RAN is represented by wireless fronthauling. Recent advances in microwave communications allow for up to a few Gbps of bandwidth over short distances (less than one Km) in the E–band ($70-80$ GHz) making it suitable as fronthaul technology for dense small cells. Moreover due to the short wavelength, devices operating in the E–band, commonly referred to as millimeter wave (MMW), can leverage on compact antennas, allowing to pack several interfaces in a small form factor.

In this paper we formalize and solve a novel BBU Placement problem where BBU pools are placed at the edges of the net-

Research leading to these results received funding from the European Unions H2020 Research and Innovation Action under Grant Agreement H2020-ICT-644843 (COHERENT Project).

978-3-901882-85-2

work, possibly co–located with macro–cells and/or distributed clouds while a reconfigurable MMW fronthaul is used in order to provide RRHs with CPRI connectivity. The MMW fronthaul leverages on steerable directional antennas in order to adapt its topology to different usage scenarios, e.g. to reducing energy consumption. We formulate the BBU placement problem as an integer linear programming (ILP) problem and we propose a placement heuristic named *SWAN* to solve the problem. To the best of our knowledge this is the first work to tackle the BBU placement problems over a reconfigurable substrate network with wireless (MMW) links. Such technology imposes several constraints that are not found in traditional wired backhauls. For example, the capacity and the availability of a link depends on both its length and on previously embedded requests, as opposed to wired backhauls where the wiring media is the only constraint to the BBU placement.

The rest of this paper is structured as follows. In Sec. II we discuss the related work. The substrate network model and the virtual request model are detailed in Sec. III. The ILP problem and the heuristic are introduced in Sec. IV. The numerical results are reported in Sec. V. Finally, Sec. VI draws the conclusions pointing out future work.

II. Related Work

As mobile networks data traffic keeps growing, traditional cellular architectures have become overloaded due to the lack of core network capacity. Data traffic growth also substantially impacts power consumption since most of the power is consumed at evolved Node Bs (around 80% as estimated in [4]). In addition, the baseband resources of current mobile network operators deployments are not used efficiently, since operators allocate resources to their evolved Node Bs in such a way to be able to meet peak hour traffic demand. This means that, due to variable traffic profiles, those resources might be underutilized across the entire year. One way to overcome these challenges is to adopt the Cloud–RAN architecture [2], [5], [6], [7].

BBU Placement. Sizeable body of work has been published on BBU placement and Cloud-RAN technology in recent years [4], [8], [9], [10], [11]. In [8] the authors propose a Colony–RAN architecture for cellular systems, which is able to change the cell layout by dynamically adapting the connections between BBUs and RRHs according to traffic demand, user distributions and user mobility. This architecture significantly reduces the number of BBUs thanks to statistical multiplexing effect. An optimization algorithm is presented in [9] for the BBU Placement problem over Fixed/Mobile Converged optical networks. The authors formulate an ILP problem, which efficiently calculates the minimum number of BBU pools taking into account maximum allowed distance between RRHs and their BBUs. The same authors put forward an energy efficient BBU Placement algorithm in optical networks in [4] aiming to minimize the Aggregation Infrastructure Power. An ILP optimization problem is formalized [10] for optimizing cells assignment to different BBU pools. Statistical multiplexing gain and required fiber length are used as key performance indicators. An analytical model is derived in [11], which optimizes C-RAN deployments by finding the most efficient relationship between using optical fiber or microwave links in fronthaul of mobile networks. In [12] cost and energy consumption reduction in C–RAN is compared to the traditional D–RAN. Authors in [13] compare fiber–based networks to microwave in terms of cost deployments. In Rural areas, fiber–based fronthauls are more cost effective for over distances (less than 500m) while MMW–based fronthauls take the advantage at around 1.6 Km and above. However, in urban areas MMW links deployment are much more effective even for very short distances.

VNF Placement. The VNF placement problem is conceptually similar to component placement in data–centers and clouds. The amount of literature in this domain is thus humbling [14], [15], [16], [17]. A survey on resource management in cloud computing environments can be found in [18]. In [14] the authors study the problem of placing virtual machine instances on physical containers in such a way to reduce communication overhead and latency. In [15] the author propose a novel design for a scalable hierarchical application components placement for cloud resource allocation. The proposed solution operates in a distributed fashion, ensuring scalability, while providing performances very close to that of a centralized algorithm. This work is extended in [16] where several algorithms for efficient data management of component-based applications in cloud environments are proposed. In [17] the elasticity overhead and the trade–off between bandwidth and host resource consumption are jointly considered by the authors when formulating the VNF placement problem. In [19], [20] a joint node and link mapping algorithm is proposed. While the authors of [21], [22], [23] tackle the problem of dynamic VNF placement. A VNF placement problem is proposed in [24] for the radio access network. In [25] an online VNF scheduling and mapping problem is formulated. The authors propose three greedy algorithms and a tabu search-based heuristic. These algorithms are compared using criteria such as cost, revenue and service processing time without considering links, bandwidth requirements and the associated transmission delay between VNFs.

Functional Split. Recently, flexible small cell functional splits has attracted a great deal of attention by MNOs, industry and academy. There are different possible functional splits between the Physical (PHY) and the Packet Data Convergence Protocol (PDCP) layers. A number of factors (e.g., traffic demand, energy efficiency, and latency constraints) have to be taken into account to decide the actual split point. For example LTE's Hybrid Automatic Repeat Request (HARQ) and MAC scheduling impose strict latency requirements which can be mitigated at the price of reduced peak data rate and fronthaul requirements. The latter can be relaxed, at the cost of reduced centralized processing benefits, by moving forward the split point within PHY layer or towards the upper layers.

A detailed discussion on various functional splits can be found in [26], [27], [28], [29], [30]. The authors of [26] propose a novel RAN as a Service (RANaaS) concept in which centralization of management and processing is flexible (i.e., partially centralization of functionalities may be executed) and can be adapted to the actual service demands. Several functional splits are introduced and numerical results on the

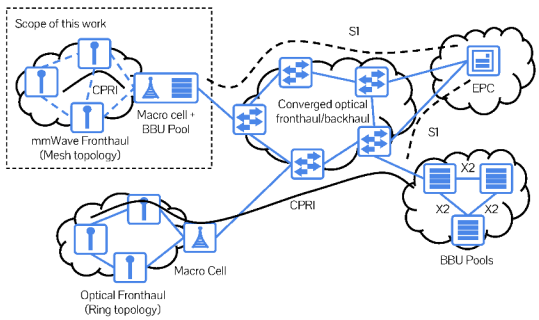


Fig. 1: Reference network architecture for our BBU Placement problem. CPRI links are represented as solid lines while S1 links are represented as dashed lines. Notice how in the case of the MMW fronthaul BBU pools are co–located with macro–cells improving computational locality and shortening fiber–based CPRI links.

required backhaul data rates for each envisioned split case are provided in [27]. The authors of [31] put forward a graph–based algorithm for analyzing some of the base–band function splittings within the PHY. The authors of [28] explore all the possible wired/wireless transport fronthauling technologies as well as associated bandwidth and latency requirements for the different functional splits.

The PHY layer split enables to exploit spatial diversity and achieve inter–cell interference mitigation through the implementation of advanced signal processing mechanisms, such as Coordinated Multi-Point Transmission and Reception (e.i., joint processing, coordinated scheduling/beamforming) [32]. A detailed investigation on the various functional splits at the PHY layer has been conducted in [29], [30].

III. Network Model

Figure 1 depicts the reference network architecture. In the lower part of the figure we can see a traditional C–RAN deployment where all the BBUs are centralized in a large high–volume computing facility and long CPRI links carried over an optical fronthaul are used in order to connect BBUs with RRHs. In the upper part of the figure instead we can see the architecture envisioned in this work. In this case BBU pools are co–located with macro–cells and a reconfigurable MMW fronthaul is used to carry the CPRI links. A traditional S1 link is used in order to deliver connectivity to the core network. The latter approach allows to reduce the length of the (expensive) CPRI links while still allowing for improved control and coordination and statistical multiplexing of BBU resources (although in this case the multiplexing gain is lower compared to a fully centralized BBU pool).

In this section we shall first detail the notation used for the substrate network model and for the virtual network request. Then in the next section we will introduce the optimal ILP formulation for the BBU Placement problem and a scalable heuristic. It is worth noticing that this work will focus on the BBU Placement problem in the MMW wireless fronthaul, thus leaving the problem of BBU Placement over converged wireless/optical network as future work.

TABLE I: Substrate network parameters

Variable	Description
G_s	Substrate network graph.
N_s	Substrate nodes in G_s.
N_s^1	Substrate RRH sites in G_s.
N_s^2	Substrate BBU pools in G_s.
E_s	Substrate links in G_s.
$\omega_a^s(n)$	Number of RF frontends available at RRH site $n \in N_s^1$.
$\omega_i^s(n)$	Number of MMW interfaces available at RRH site $n \in N_s^1$.
$\omega_c^s(n)$	Number of BBUs supported by BBU pool $n \in N_s^2$.
$loc(n)$	Geographical location of node $n \in N_s$ (x, y).
$\delta^s(n)$	Coverage radius of node $n \in N_s$ (in meters).
$\omega_b^s(e^{nm})$	Capacity of the MMW link $e^{nm} \in E_s$ (in Gbps).

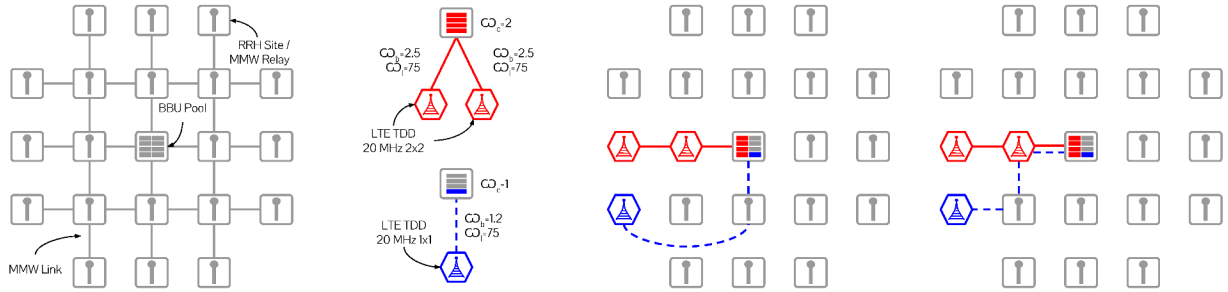
A. Substrate Network Model

The substrate network considered in this work encompasses both computational and networking resources. The former consists of micro datacenters, possibly co–located with macro–cells. The latter consists of a reconfigurable wireless fronthaul and multiple RRH sites. The reconfigurable MMW fronthaul consists of MMW routers equipped with a variable number of MMW interfaces and highly directional and steerable antennas. Finally each RRH site is equipped with a variable number of RF frontends. RRH sites and MMW routers are assumed to be co–located.

Let $G_s = (N_s, E_s)$ be an *undirected* graph modeling the physical network, where $N_s = N_s^1 \cup N_s^2$ is the set of $n_1 = |N_s^1|$ RRHs sites/MMW Relays and $n_2 = |N_s^2|$ BBU pools and E_s is the set of fronthaul links. Notice how N_s^1 nodes in the substrate network can act as both RRH and MMW relays, i.e. they can both serve end–user terminals over, for example an LTE air interface, and act as relays in MMW fronthaul, while N_s^2 nodes, i.e. the ones co–located with macro–cells, can act only as BBU pools. A wireless edge $e^{nm} \in E_s$ if and only if a line–of–sight connection exists between $n, m \in N_n$.

Three weights, $\omega_a^s(n)$, $\omega_i^s(n)$ and $\omega_c^s(n)$, are assigned to each node $n \in N_s : \omega_{a,i,c}^s(n) \in \mathbb{N}^+$ representing, respectively, the number of RF frontends and MMW interfaces available at the RRH site, and the number of BBUs supported by the BBU pool. Each substrate node is also associated with a geographic location $loc(n^s)$, as x, y coordinates, and a coverage radius $\delta^s(n)$, in meters indicating the coverage area of the small cell centered on RRH n. Another weight $\omega_b^s(e^{nm})$ is assigned to each link $e^{nm} \in E_s : \omega_b^s(e^{nm}) \in \mathbb{N}^+$ representing the capacity (in Gbps) of the link connecting the two nodes. Finally, let P_s be the set of all loop free substrate paths $\in N_s$ and $P_s(s,t)$ the shortest path between $s, t \in N_s$. Table I summarizes the substrate parameters.

A sample substrate network is sketched in Fig. 2. This 21–nodes network is representative of a scenario where RRH sites are deployed at road intersections in a dense Manhattan–like urban area. Outer RRHs are equipped with a single interface, while the single BBU pool is equipped with 4 interfaces. The remaining RRHs are equipped with 2 interfaces. It is worth noticing that, such a topology is to be considered merely as an example and that the problem formulation itself is generic and can accommodate other types of topologies.



(a) Substrate network. (b) Virtual request. (c) Valid BBU Placement. (d) Invalid BBU Placement.

Fig. 2: Sample substrate network, virtual request, and BBU Placement. Substrate node spacing is uniform. Substrate nodes and links weights are not depicted in order to improve readability. Notice how the BBU Placement in Fig. 2c is valid since the interface constraint is satisfied on all relay nodes. Conversely, the BBU Placement in Fig. 2d is invalid since it would require 4 interfaces in relaying RRH node.

B. Virtual Network Model

Requests are formulated as *undirected* graphs $G_v = (N_v, E_v)$, where $N_v = N_v^1 \cup N_v^2$ is the set of $n_1 = |N_v^1|$ RRHs and $n_2 = |N_v^2|$ BBU pools and $E_s \subseteq N_v^1 \times N_v^2$ is the set of fronthaul links. Notice how, in the case of virtual requests, nodes consist of RRHs and BBU pools, i.e. user do not request MMW Relays. An edge $e^{nm} \in E_v$ if, and only if, the BBU pool m is mapped to the RRH n. Thus, as opposed to the previous substrate network model, edges in the virtual requests represent the logical mapping between RRHs and their BBUs. As an additional constraint, we require that each RRH to be mapped to one, and only one, BBU pool. Conversely, different RRHs can be mapped to the same BBU pool. This formulation allows users to specify requests where a group RRHs are mapped to the same BBU pool enabling advanced interference control features like CoMP and eICIC.

Nodes in the virtual request have two weights $\omega_c^v(n)$ and $\omega_a^v(n)$ indicating respectively the number of BBUs requested by the node $n \in N_v^2$ and the number of RF frontends request by the cell $n \in N_v^1$. Each cell $n \in N_v^1$ is also associated with a geographic location $loc(n)$, as x, y coordinates. This information is used, together with the substrate node location and its coverage radius, to express how far a cell $n \in N_v^2$ can be placed from the preferred location specified by $loc(n)$. A weight $\omega_b^v(e^{nm})$ is assigned to each link $e^{nm} \in E_v : \omega_b^s(e^{nm}) \in \mathbb{N}^+$ denoting the requested capacity of the links connecting the two nodes. The weight $\omega_b^v(e^{nm})$ can be easily derived starting from the aggregated cell bandwidth requirements. For example, in order to support 150 Mbps of bandwidth, a 20 MHz FDD LTE channel with a $2x2$ MIMO antenna configuration is required. This translate to a CPRI bitrate of $\approx 2.5 Gbps$. Table II summarizes the virtual network request parameters.

A sample virtual network request is sketched in Fig. 2b. The request consists of 3 small cells. Two of them (the red ones) require their BBUs to be co–located at the same BBU pool. While for the third small cell (the blue one) there is no such a constraint. Notice also that, the blue small cell is characterized by a lower CPRI bandwidth requirement, i.e. 1.2 Gbps Vs. 2.5 Gbps, and that only one BBU is requested.

TABLE II: Virtual network request parameters

Variable	Description
G_v	Virtual network graph.
N_v	Virtual nodes in G_v.
N_v^1	Virtual RRH sites in G_v.
N_v^2	Virtual BBU pools in G_v.
E_v	Virtual links in G_v.
$\omega_a^v(n)$	Number of RF frontends required on RRH $n \in N_v^1$.
$\omega_c^v(n)$	Number of BBUs required by RRH $n \in N_v^2$.
$loc(n)$	Desired geographical location for node $n \in N_v^1$ (x, y).
$\omega_b^v(e^{nm})$	Requested capacity for the CPRI link $e^{nm} \in E_v$ (in Gbps).

IV. BBU Placement

A. Overview

Upon arrival of a virtual network request, the substrate network must decide if it can be supported or if it shall be rejected. The embedding process consists of two steps: the node embedding, and the link embedding. In the first step (node embedding), each node (BBU and small cells) in the request is mapped to a different substrate node. In the second step (link embedding), each link is mapped to a single substrate path. In both cases nodes and links constraints must be satisfied.

Due to the fact that MMW links require highly directional antennas and that MMW Relays can be equipped with a limited number of interfaces, only a subset of the viable substrate links can be used at a given time. For example consider the BBU Placement depicted in Fig. 2c. In this case three RRHs are deployed over the substrate network. The red ones are assumed to require two CPRI Option 3 links (for a total of 5 Gbps), as a result one MMW interface is required on the BBU pool to serve the request, while two interfaces are required on the relay node: one for serving the local small cell and one for relaying the CPRI link of the other small cells. On the other hand the blue small cell is assumed to require just a CPRI Option 2 link (for a total of 1.2 Gbps), as a result a longer MMW link can be used minimizing the number of relaying nodes required to serve the request. The relationship between CPRI capacity and MMW lenght will be further discussed in the evaluation section. Notice how,

in the alternative BBU Placement depicted in Fig. 2d, the constraint on the maximum number of interfaces utilized on the relaying node is violated (4 would be required to support this configuration while only 2 are actually available).

B. ILP Formulation

In order to properly map the location constraint, we need to modify the substrate network. Every RRH $n \in N_v^2$ in the virtual request has a location constraint $loc(n)$, likewise every substrate RRH/Relay $n' \in N_s^1$ has both a location $loc(n')$ and a coverage radius $\delta(n)$. We can then define for each virtual node n the coverage cluster $\Omega(n)$:

$$\Omega(n) = \left\{ n' \in N_s^2 | dis(loc(n), loc(n')) \leq \delta(n') \right\} \quad (1)$$

We can now provide the optimal ILP formulation for the BBU Placement problem. The overall objective is to compute the optimal BBU Placement based on the available computational and fronthaul radio resources under a certain cost function. In our formulation we chose to minimize the overall number of substrate links, and thus MMW interfaces, utilized to support the virtual network requests. The rationale here is to reduce the number of active MMW interfaces in order to minimize the overall energy consumption of the MMW fronthaul. Other objective functions are however possible, optimizing other aspects of the system. The chosen objective function is:

$$minimize \quad \sum_{e \in E_s} \sum_{e' \in E_v} \omega_b^v(e') \Phi_e^{e'}$$

where $\Phi_e^{e'} \in 0, 1$ is a binary variables indicating if the virtual link $e' \in E_v$ has been mapped to the substrate link $e \in E_s$. Similarly, the binary variable $\Phi_n^{n'}$ indicates if the virtual node $n' \in N_v$ has been mapped to the substrate node $n \in N_s$.

A valid solution is the one where the BBU resources utilized by the virtual request are at most equal to the available resources on the substrate BBU pools nodes and links:

$$\sum_{n' \in N_v^2} \omega_c^v(n') \Phi_n^{n'} \leq \omega_c^s(n) \quad \forall n \in N_s^2 \quad (2)$$

$$\sum_{e' \in E_v} \omega_b^v(e') \Phi_e^{e'} \leq \omega_b^s(e) \quad \forall e \in E_{s'} \quad (3)$$

$$\sum_{n' \in N_v^2} \omega_a^v(n') \Phi_n^{n'} \leq \omega_a^s(n) \quad \forall n \in N_s^2 \quad (4)$$

Every node in the request shall be mapped only once:

$$\sum_{n \in N_s} \Phi_n^{n'} = 1 \quad \forall n' \in N_v \quad (5)$$

Every RRH in the request shall be mapped only on substrate nodes in its coverage cluster:

$$\sum_{n \in N_s \setminus \Omega(n')} \Phi_n^{n'} = 0 \quad \forall n' \in N_v \quad (6)$$

The sum of used substrate links originating from, or terminating to, each substrate node must be equal to, or less than, the number of MMW interfaces available on that node:

$$\sum_{e^{ij} \in E_v} \Phi_{e^{nm}}^{e^{ij}} + \sum_{e^{ij} \in E_v} \Phi_{e^{mn}}^{e^{ij}} \leq \omega_i^s(n) \quad \forall n \in N_s \quad (7)$$

Finally, the following constraint enforces that for each link $e^{nm} \in E_v$ there must be a continuous path allocated between the pair of physical nodes on top of which the virtual nodes $n, m \in N_v$ have been mapped:

$$\sum_{j \in N_s}^{j>i} \Phi_{e^{ij}}^{e^{nm}} - \sum_{j \in N_s}^{j<i} \Phi_{e^{ji}}^{e^{nm}} = \Phi_i^n - \Phi_i^m \quad (8)$$
$$\forall i \in N_s \quad \forall e^{nm} \in E_v$$

C. Heuristic

The ILP formulation, described in the previous sections, cannot be applied to realistic scenarios due to its limited scalability. For example, embedding a 4–nodes request (1 BBU and 3 small cells) over a $k = 7$ grid–size substrate topology can take up to 1 day on Intel Core i7 laptop (3.0 GHz CPU, 16 Gb RAM) using the Matlab® ILP solver (intlinprog). In this section we present a heuristic, named *SWAN*, that can handle similar requests in less than 10 milliseconds.

The proposed greedy heuristic is composed of three steps implementing a joint node *and* link embedding strategy (see pseudo code in Alg. 1). Let $m_1 = |N_s^1|$ and $m_2 = |N_s^2|$ be the number of, respectively, substrate RRH sites and substrate BBU pools, with $m = m_1 + m_2$. Similarly, let $n_1 = |N_v^1|$ and $n_2 = |N_v^2|$ be the number of, respectively, virtual RRH sites and virtual BBU pools. Finally, let $k = |E_s|$ be the number of edges in the substrate network.

In the first step for each virtual node $n \in N_v$ the heuristic loops over the substrate nodes and computes the list of candidate nodes $candidates(n)$. These are the substrate nodes that can support the virtual nodes in the request given the input capacity and location constraints. This process takes $O(n_1 m_1 + n_2 m_2)$ time.

In the second step, the list of virtual BBU nodes is traversed starting with the virtual BBU nodes $n \in N_v^1$ with more embedding opportunities. For each of the candidate substrate BBU pools $p \in candidates(n)$, the heuristic considers all the neighboring nodes $m \in N_v^2$ of the virtual node n. The heuristic then computes how much it would cost to embed each virtual node pairs n, m including the cost to embed the virtual edge e^{nm} (line 26 through 32 in the pseudocode). The heuristic then assigns the node n to the substrate node $p \in candidates(n)$ with the lowest mapping cost (line 35 and 36 in the pseudocode). The rationale here is to place a BBU node on the BBU pool that can support all of its RF front-ends at the minimal cost. This process requires $O(n_2 m_2 n_1 (m_1 - 1) k \log_{10} m)$ time.

In the third and final step, the list of virtual BBU nodes is traversed again. For each virtual BBU node $n \in N_v^1$, the heuristic considers its neighbors, i.e. the RF front–ends. Each front–end $m \in N_v^2$ is placed on the substrate node with the lowest mapping cost (line 41 through 49 in the pseudocode). Once the RF front–end is placed the heuristic allocates the path $P_s(mapped(n), mapped(m))$. This results in virtual nodes in a request to be placed close to each other over the substrate network, which in time means that less substrate resources are needed to support a given number of requests. Step 3 takes $O(n_2 n_1 (m_1 - 1) k \log_{10} m)$ time.

Algorithm 1 Nodes and links assignment

procedure *SWAN*(G_s, G_v)
  Step 1: Compute list of candidates.
  for $n \in N_v^1$ **do** ▷ RF front–ends.
    for $p \in N_s^1$ **do** ▷ RRH sites.
      $d \leftarrow dis(loc(n), loc(p))$ ▷ Distance in meters.
      if $d \leq \delta(m)$ and $\omega_a^v(n) \leq \omega_a^s(p)$ **then**
        $candidates(n) \leftarrow p$
      end if
    end for
  end for
  for $m \in N_v^2$ **do** ▷ Virtual BBUs.
    for $q \in N_s^2$ **do** ▷ Substrate BBU pools.
      if $\omega_c^v(m) \leq \omega_c^s(q)$ **then**
        $candidates(m) \leftarrow q$
      end if
    end for
  end for
  Step 2: Perform BBU Placement.
  for $n \in N_v^2$ **do** ▷ Virtual BBUs.
    for $i \in N_s$ **do** ▷ Initialize mapping cost array.
      $m_c(i) \leftarrow 0$
    end for
    for $p \in candidates(n)$ **do**
      for $m \in neighbors(n)$ **do**
        $cost \leftarrow +\infty$
        for $q \in candidates(m)$ **do**
          $c_{new} \leftarrow \sum_{e \in P_s(p,q)} \omega_e^v(e^{nm})$
          $c_{curr} \leftarrow min(c_{curr}, c_{new})$
        end for
        $m_c(p) \leftarrow c_{curr}$ ▷ Accumulate mapping cost.
      end for
    end for
    $p \leftarrow argmin(m_c(p))$
    $mapped(n) \leftarrow p$
  end for
  Step 3: Perform RF front–ends embedding.
  for $n \in N_v^2$ **do** ▷ Virtual BBUs.
    $p \leftarrow mapped(n)$
    for $m \in neighbors(n)$ **do**
      for $i \in N_s$ **do** ▷ Initialize mapping cost array.
        $m_c(i) \leftarrow 0$
      end for
      for $q \in candidates(m)$ **do**
        $m_c(q) \leftarrow \sum_{e \in P_s(p,q)} \omega_e^v(e^{nm})$
      end for
      $q \leftarrow argmin(m_c(q))$
      $mapped(m) \leftarrow q$
      Allocate path $P_s(p, q)$
    end for
  end for
end procedure

Thus, the overall time complexity of the *SWAN* algorithm is $O(n_1 m_1 + n_2 m_2 + [n_1 n_2 (m_1 - 1) k \log_{10} m](1 + m_2)$.

V. Evaluation

The goal of this section is to compare the performance of the ILP–based placement algorithm with the performance of the heuristic using different synthetic substrate networks and different virtual network requests. We shall first describe the simulation environment and the performance metrics used in our study. Then we will report on the outcome of the numerical simulations carried out in a discrete event simulator implemented in Matlab®.

TABLE III: CPRI link bandwidth per option

CPRI Option	CPRI Rate	IQ Sampling Rate	LTE Conf.
1	600 Mbps	400 Mbps	10 MHz, $1x1$
2	1.2 Gbps	0.9 Gbps	20 MHz, $1x1$
3	2.4 Gbps	1.8 Gbps	20 MHz, $2x2$
5	5 Gbps	3.6 Gbps	20 MHz, $4x4$

A. Simulation Environment

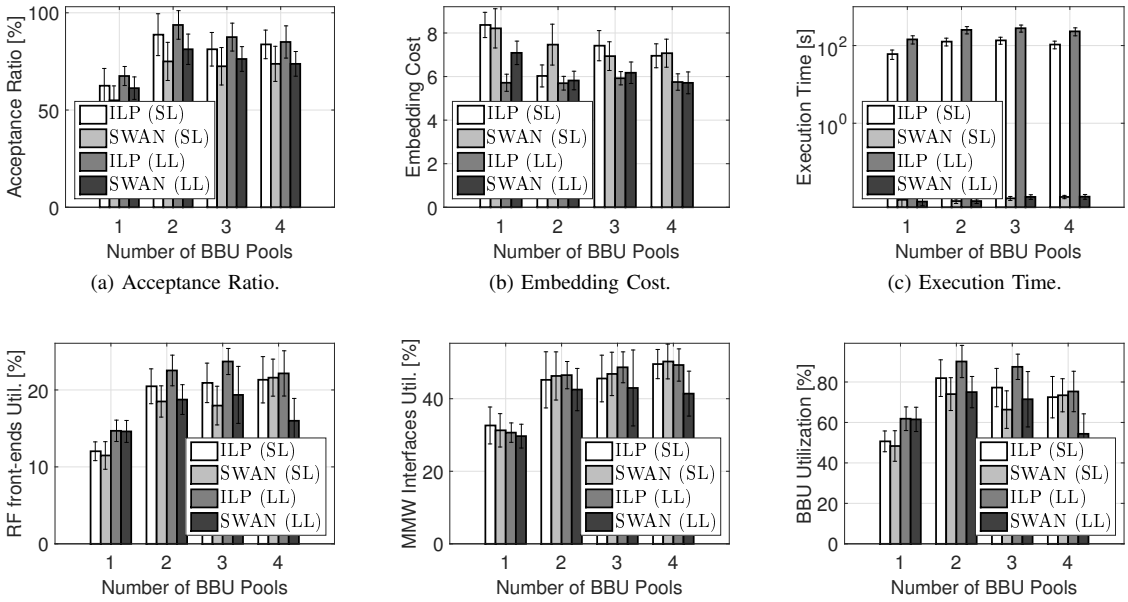
The simulation parameters and in particular the choices made for the substrate network characteristics originate from a number of works on MMW communications. In [33] the authors suggest that optimum coverage can be achieved by having 200 meters as distance between each RRH. In [34] the authors estimate 1 Km to be the typical coverage radius for MMW links in line–of–sight conditions. Finally, in [35], [36] the authors rely on empirical measurements to show that bitrates as high as 10 Gbps can be achieved with an outage probability of $\approx 11\%$, while 5 Gbps of bitrate can be achieved with an outage probability of $\approx 3\%$.

The ILP–based placement algorithm and the proposed placement heuristic are evaluated in two different scenarios differentiated by the MMW links length and by their performance (bandwidth). In the first scenario, named *short links* (SL), we assume that the maximum MMW line–of–sight distance is equal to 250m and that at this distance the link can deliver up to 5 Gbps. In the second scenario, named *long links* (LL), we assume that the maximum MMW line–of–sight distance is equal to 500m and that at this distance the link can deliver up to 2.5 Gbps. Notice that, the shorter, high bandwidth links are also available in this second scenario. Table III summarizes some of the most common CPRI setups providing some illustrative LTE configurations that can be supported by each option. As it can be seen, the *short links* and *long links* scenarios corresponds, respectively, to a CPRI Option 5 and to a CPRI Option 3 configuration.

The reference substrate network is a grid–shaped 2D lattice network with 5×5 similar to the one depicted in Fig. 2a. Nodes spacing is uniform and set to 250m. Each node can be either an RRH site or a BBU pool (they are all however MMW relays). The number of BBU pool is variable between 1 and 4. BBU pools are randomly deployed. RRH relays at the edges of the network are equipped with a single MMW interface. BBU pools are equipped with 8 MMW interfaces. All other MMW relays are equipped with 4 MMW interfaces.

Virtual network requests consist of star–shaped networks like the ones depicted in Fig. 2b. The number of RF front-ends in each request as well as their characteristics (LTE bandwidth) are randomly generated for each request. In particular, for each request we randomly generate between RF front–ends 1 and 4. Each of them may require either a CPRI option 3 or a CPRI option 5 link. Each request also contains a single BBU with a requested capacity ω_c set to the number of equivalent CPRI option 3 in the request (i.e. one CPRI Option 5 link equals to 2 CPRI Option 2 links).

In this study we assume that a fixed number of virtual requests are embedded sequentially. In particular, in each run



(a) Acceptance Ratio. (b) Embedding Cost. (c) Execution Time.

(d) Antenna Utilization. (e) Interface Utilization. (f) BBU Utilization.

Fig. 3: Performance of the ILP–based algorithm and of the heuristics with a different number of substrate BBU pools.

the simulator tries to embed 10 randomly generated request using either the ILP–based algorithm or the heuristic. Reported results are the average of 8 simulations.

B. Simulation Results

Figure 3 shows the performance of the ILP–based BBU placement algorithm and of the heuristic with an increasing number of substrate BBU pools and for the two scenarios being considered. As it can be seen the acceptance ratio (Fig. 3a) increases with the number of available substrate BBU pools. This is due to the fact that as the number of BBU pools increases more embedding opportunities become available.

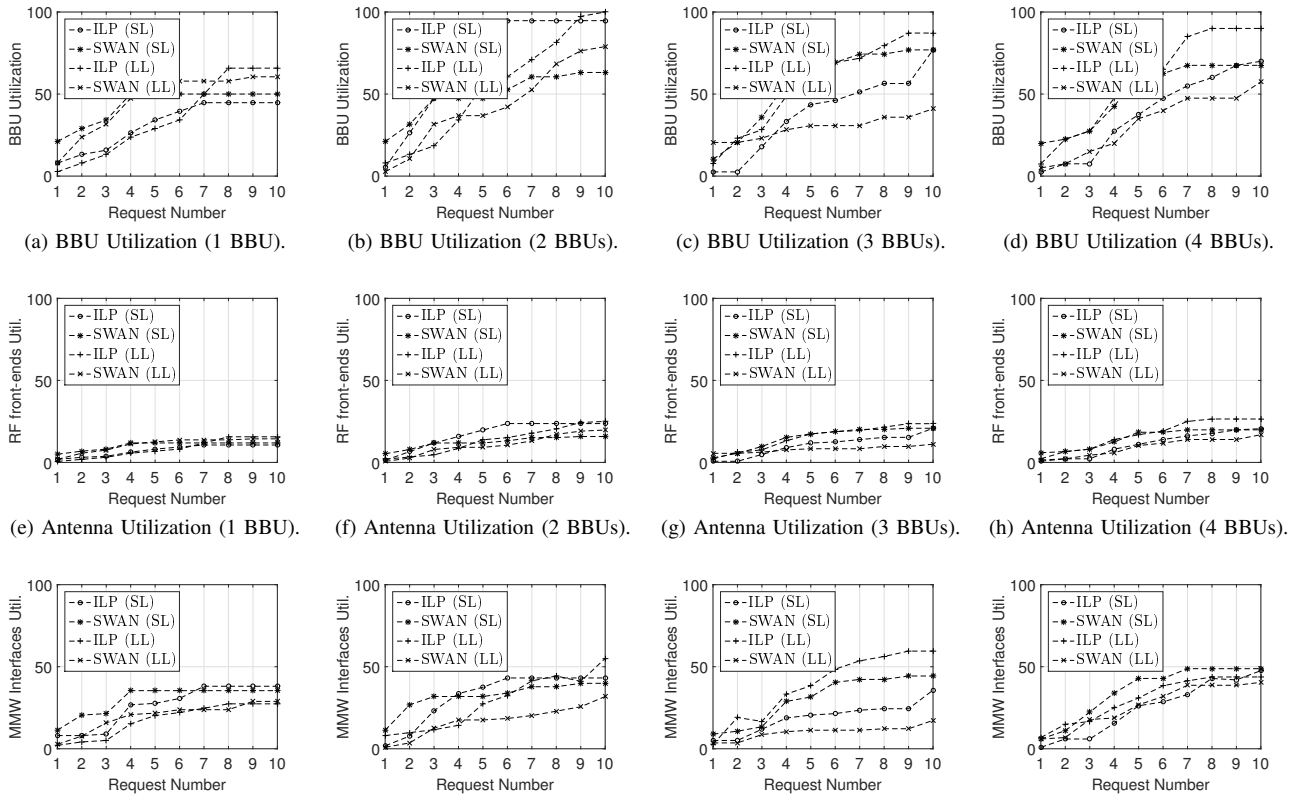
As expected the ILP–based placement algorithm is more efficient than the heuristic in mapping the incoming request. This can be seen in terms of both an higher number of accepted requests (Fig. 3a) and a lower average embedding cost (Fig. 3b). Notice that (Fig. 3a) even though the acceptance ratio of the ILP–based algorithm in both scenarios is approximately the same, the average embedding cost is smaller in the *long links* scenario. This means that fewer substrate resources, i.e. MMW interfaces, are used in the *long links* scenario.

Figure 3c shows that the average amount of time required to embed a *single* request using the ILP–based placement algorithm is significantly higher than the time required to embed the same request using the heuristic. The ILP problem becomes essentially intractable for substrate networks with more than a few tens of nodes, while the heuristic can effectively embed complex requests on substrate networks with hundreds of nodes in a limited amount of time. Although operators may prefer to wait even several weeks in order to have a optimal BBU placement, we argue that our heuristic could allow a faster service on–boarding time while the ILP–based placement algorithm could not use to periodically optimize the network configuration.

Figure 3d plots the final RF front–ends utilization. As it can be seen the utilization of the RF front–ends increases with the number of BBU pools. However a saturation point around 3 substrate BBU pools can be noticed. The same consideration can also be made for the average MMW interfaces utilization and for the BBU pool utilization. Notice also how both the RF front–ends and the MMW interfaces utilization never approaches 100%, this essentially means that the BBU placement does not fail due to lack of such resources, but rather due to a non–homogeneous utilization of the available resources. We leave as future work the task of analyzing how MMW interface density impacts on the acceptance ratio.

In order to gain an increased insight into how resources are actually utilized during the embedding process, we will now analyze in detail a single iteration of the simulator. We remind the reader that in each iteration the simulator tries to embed 10 randomly generated virtual network requests. Figure 4 plots the substrate resources utilization for the two scenarios. As it can be seen the ILP–based algorithm utilizes more substrate resources irrespective of the number of available substrate BBU pools. This can be explained by the fact that the ILP–based placement algorithm is capable of embedding a higher number of requests than the heuristic.

Notice also that both the RF front–ends as well as the MMW interfaces utilization increase with the number of BBU pools. The explanation for this behavior is twofold. On the one hand when additional BBU pools (which we remind the reader do not possess any RF front–ends) are added to the network the overall number of available front–ends decreases. However, since the BBU placement rarely fails due to unavailability of



(a) BBU Utilization (1 BBU). (b) BBU Utilization (2 BBUs). (c) BBU Utilization (3 BBUs). (d) BBU Utilization (4 BBUs).

(e) Antenna Utilization (1 BBU). (f) Antenna Utilization (2 BBUs). (g) Antenna Utilization (3 BBUs). (h) Antenna Utilization (4 BBUs).

(i) Interfaces Utilization (1 BBU). (j) Interfaces Utilization (2 BBUs).. (k) Interfaces Utilization (3 BBUs). (l) Interfaces Utilization (4 BBUs).

Fig. 4: Performance of the ILP–based algorithm and of the heuristics with a different number of substrate BBU pools.

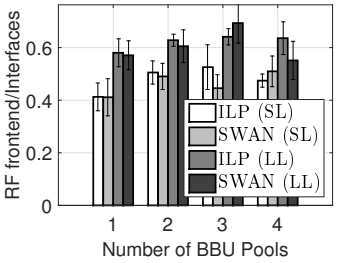


Fig. 5: Ratio between the number of requested RF front–ends and the utilized MMW interfaces.

either RF front–ends or MMW interfaces, when the number of BBU pools in the network increases also the number of embedding opportunities increase. This behavior can be seen also in terms of acceptance ratio in Fig. 3a. Notice in fact that, the acceptance ratio for the substrate network with 4 BBU pools is higher than the one with only 1 BBU pool.

Finally, in Fig. 5 we plot the ratio between the number of requested RF front–ends and the utilized MMW interfaces. As it can be expected in the *long links* scenario the ratio is higher than in the *short links* scenario. This is due to the fact that, when longer links are available the same distance can be covered in a single hop (using just 2 interfaces) rather than in two hops (using 4 interfaces).

VI. Conclusions

Small–cells are rapidly emerging as a cost–efficient solution to provide additional capacity in current and future mobile networks. However scalable and flexible fronthaul technologies are needed in order to make small–cells an economically viable option for MNOs. Among the many solution available, wireless front–hauls are one of the most promising.

In this paper we provide a novel formulation for the BBU Placement problem where BBU pools are placed at the edges of the network, possibly co–located with macro–cells, and a reconfigurable MMW wireless fronthaul is used in order to provide RRHs with connectivity. We introduce an ILP–based algorithm solving the placement problem for small networks and a BBU Placement heuristic for larger networks. We perform extensive numerical simulation in order to better understand the trade–offs involved in deploying wireless front–hauls in dense networks scenarios.

As future work we plan to extend the problem formulation to more complex scenarios. In particular we want to consider scenarios were the MMW wireless network is used as both fronhaul for the RRHs and as backhaul for other technologies, e.g. WiFi, LTE. We want also to develop a better channel model capable of accounting for both capacity and latency. Finally, we want to extended the problem formulation in order to account for different functional splits in the small–cells covering the full spectrum between D–RAN and C–RAN.

REFERENCES

[1] "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015-2020," Cisco, Tech. Rep., 2016.

[2] "C-RAN, The Road Towards Green RAN," China Mobile, Tech. Rep., 2011.

[3] "Common Public Radio Interface, Interface Specification V6.0," CPRI, Tech. Rep., August 2013.

[4] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in *Proc. of IEEE VTC 2009-Fall*, Anchorage, AK, USA, 2009.

[5] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud ran for mobile networksa technology overview," *Communications Surveys & Tutorials, IEEE*, vol. 17, no. 1, pp. 405–426, 2014.

[6] "The Benefits of Cloud-RAN Architecture in Mobile Network Expansion," Fujitsu, Tech. Rep., 2014.

[7] "Cloud RAN the benefits of virtualization, centralization and coordination," Ericsson, Tech. Rep., 2015.

[8] S. Namba, T. Matsunaka, T. Warabino, S. Kaneko, and Y. Kishi, "Colony-RAN architecture for future cellular network," in *Proc. of IEEE FutureNetw*, Berlin, Germany, 2012.

[9] N. Carapellese, M. Tornatore, and A. Pattavina, "Placement of base-band units (BBUs) over fixed/mobile converged multi-stage WDM-PONs," in *Proc. of IEEE ONDM*, Brest, France, 2013.

[10] H. Holm, A. Checko, R. Al-obaidi, and H. Christiansen, "Optimal assignment of cells in C-RAN deployments with multiple BBU pools," in *Proc. of EuCNC*, Paris, France, 2015.

[11] R. Al-obaidi, A. Checko, H. Holm, and H. Christiansen, "Optimizing Cloud-RAN deployments in real-life scenarios using Microwave Radio," in *Proc. of EuCNC*, Paris, France, 2015.

[12] C. Aleksandra, H. L. Christiansen, and M. S. Berger, "Evaluation of energy and cost savings in mobile Cloud RAN," in *Proc. of OPNETWORK*, 2013.

[13] "Mobile backhaul: Fiber vs. microwave," Ceragon Networks, Tech. Rep., 2009.

[14] D. Breitgand, A. Epstein, A. Glikson, A. Israel, and D. Raz, "Network aware virtual machine and image placement in a cloud," in *Proc. of IEEE CNSM*, Zurich, Switzerland, 2013.

[15] M. Barshan, H. Moens, and F. De Turck, "Design and evaluation of a scalable hierarchical application component placement algorithm for cloud resource allocation," in *Proc. of IEEE CNSM*, Rio, Brasil, 2014.

[16] M. Barshan, H. Moens, S. Latre, and F. De Turck, "Algorithms for efficient data management of component-based applications in cloud environments," in *Proc. of IEEE NOMS*, Krakow, Poland, 2014.

[17] M. Ghaznavi, A. Khan, N. Shahriar, K. Alsubhi, R. Ahmed, and R. Boutaba, "Elastic Virtual Network Function Placement," in *Proc. of IEEE CloudNet*, Niagara Falls, Canada, 2014.

[18] B. Jennings and R. Stadler, "Resource management in clouds: Survey and research challenges," *J. Netw. Syst. Manage.*, vol. 23, no. 3, pp. 567–619, Jul. 2015.

[19] R. Guerzoni, R. Trivisonno, I. Vaishnavi, Z. Despotovic, A. Hecker, S. Beker, and D. Soldani, "A novel approach to virtual networks embedding for sdn management and orchestration," in *in Proc. of IEEE NOMS*, Krakow, Poland, 2014.

[20] Z. Despotovic, A. Hecker, A. N. Malik, R. Guerzoni, I. Vaishnavi, R. Trivisonno, and S. A. Beker, "VNetMapper: A fast and scalable approach to virtual networks embedding," in *Proc. of IEEE ICCCN*, Shanghai, China, 2014.

[21] S. Clayman, E. Maini, A. Galis, A. Manzalini, and N. Mazzocca, "The dynamic placement of virtual network functions," in *Proc. of IEEE NOMS*, Krakow, Poland, 2014.

[22] H. Moens and F. De Turck, "VNF-P: A model for efficient placement of virtualized network functions," in *Proc. of IEEE CNSM*, Rio, Brasil, 2014.

[23] F. Bari, S. R. Chowdhury, R. Ahmed, and R. Boutaba, "On orchestrating virtual network functions," in *Proc. of IEEE CNSM*, Barcelona, Spain, 2015.

[24] R. Riggio, T. Rasheed, and R. Narayanan, "Virtual network functions orchestration in enterprise WLANs," in *Proc. of IEEE ManFI*, Ottawa, Canada, 2015.

[25] R. Mijumbi, J. Serrat, J. L. Gorricho, N. Bouten, F. D. Turck, and S. Davy, "Design and evaluation of algorithms for mapping and scheduling of virtual network functions," in *in Proc. of IEEE NetSoft*, Seoul, Korea, 2015.

[26] P. Rost, C. J. Bernardos, A. D. Domenico, M. D. Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wbben, "Cloud technologies for flexible 5g radio access networks," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 68–76, May 2014.

[27] D. Wubben, P. Rost, J. S. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, and G. Fettweis, "Benefits and impact of cloud computing on 5g signal processing: Flexible centralization through cloud-ran," *Signal Processing Magazine, IEEE*, vol. 31, no. 6, pp. 35–44, 2014.

[28] A. Maeder, M. Lalam, A. De Domenico, E. Pateromichelakis, D. Wubben, J. Bartelt, R. Fritzsche, and P. Rost, "Towards a flexible functional split for Cloud-RAN networks," in *Proc. of EuCNC*, Bologna, Italy, 2014.

[29] U. Dötsch, M. Doll, H.-P. Mayer, F. Schaich, J. Segel, and P. Sehier, "Quantitative analysis of split base station processing and determination of advantageous architectures for LTE," *Bell Labs Technical Journal*, vol. 18, no. 1, pp. 105–128, 2013.

[30] J. Bartelt, P. Rost, D. Wubben, J. Lessmann, B. Melis, and G. Fettweis, "Fronthaul and backhaul requirements of flexibly centralized radio access networks," *Wireless Communications, IEEE*, vol. 22, no. 5, pp. 105–111, 2015.

[31] J. Liu, S. Zhou, J. Gong, Z. Niu, and S. Xu, "Graph-based framework for flexible baseband function splitting and placement in C-RAN," in *Proc. of IEEE ICC*, London, United Kingdom, 2015.

[32] S. Sesia, I. Toufik, and M. Baker, *LTE - the UMTS long term evolution: from theory to practice*. Wiley, 2011.

[33] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[34] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, 2011.

[35] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1152–1163, 2014.

[36] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave cellular wireless networks: Potentials and challenges," *CoRR*, vol. abs/1401.2560, 2014.