Enhancing RAN Throughput by Optimized CoMP Controller Placement in Optical Metro Networks

Francesco Musumeci, *Member, IEEE,* Ernesto De Silva, and Massimo Tornatore, *Senior Member, IEEE*

***Abstract*—The 5th generation of mobile communications (5G) will target unprecedented network performance and Quality of Service for end users. Among the various aspects which will be addressed in 5G, advanced cell coordination is deemed as crucial to maximize network throughput. In particular, in this paper we refer to Coordinated Multipoint (CoMP) techniques, that allow to coordinate groups of cells (i.e., *clusters*) through a coordination controller, namely a Radio Controller Coordinator (RCC), to enhance the mobile-network throughput by reducing interference. We focus on the placement of RCCs in the metro optical network and on its impact on the performance of cell coordination. We provide strategies to perform an optimized placement of such controllers in metro optical networks in order to maximize network throughput via cell coordination. Several CoMP techniques have been designed, whose throughput gain is affected by various factors, e.g., gain increases with the cluster size, while it decreases for larger latencies between the RCC and cells. As current metro networks are characterized by a hierarchical architecture with different levels of central offices, the choice of where to place controllers to maximize throughput gain can be optimized according to several factors, i.e., network geographical dimension, cells density and available technology. In addition, selection of the most appropriate CoMP technique to be used in each cluster is not trivial, as the gain provided by the various techniques is differently affected by cluster size and latency between the RCC and cells. Our results show that under certain conditions optimized placement provides up to around 10% higher coordination gain with respect to fixed controller placement. Moreover, when adopting fronthaul technology, the coordination gain provided by an optimized controller placement may increase up to 20% in comparison to fixed placement.**

***Index Terms*—5G; Radio Access Networks; Coordinated Multipoint; WDM; optical metro-access networks.**

I. Introduction

FIFTH-GENERATION (5G) mobile networks target unprecedented performance in terms of devices density, data rate per user, latency and network-coordination [2]. To achieve these goals, several new technological solutions are being investigated and deployed, such as cell densification, advanced radio-coordination protocols (e.g., Coordinated Multipoint, CoMP), and Centralized Radio Access Network (C-RAN) architectures. The adoption of these technologies in the 5G RAN will strain the backhaul optical access-metro network, which will be required to effectively transport huge amount of mobile-backhaul data with sub-ms latency. In a general sense, the strict requirements of the 5G RAN will increasingly influence the design of future optical metro-access networks.

In particular, in this paper, we concentrate on the relation between the optical backhaul metro network and the performance of CoMP coordination. CoMP techniques are used to coordinate clusters of cell sites (e.g., LTE-A eNodeBs) via a Radio Controller Coordinator (RCC), with the objective of enhancing users' throughput by reducing signal interference, especially at cells edge. We observe that the metro network characteristics (i.e., its topology, capacity and switching capabilities) might impact significantly the decisions on the placement of CoMP controllers in the metro area. To model such impact, we consider a hierarchical optical metro-access networks, consisting of a multi-stages ring-and-spur topology, as shown in Fig. 1. Four levels of nodes are considered in this hierarchy, namely, Cell Sites (CSs in the figure), Access Central Offices (COs), Main COs and Core COs, interconnected via optical links. Due to this hierarchical organization of COs interconnected over multi-ring or meshed topologies, there are multiple possible network-node locations to place the CoMP RCCs. In general, the throughput gain (or equivalently, coordination gain) introduced by CoMP is influenced by several competing factors, mainly, the cluster size and the latency between cell sites and RCC [3]. On the one hand, the larger is the cluster size, i.e., the number of coordinated cells, the higher is the throughput gain. Since metro-access networks are typically organized in hierarchical topologies, to coordinate more cells, RCCs shall be placed at higher layers of this hierarchy, e.g., in Main COs, or even in the Core CO. On the other hand, placing RCC in nodes located in higher metro-network layers will increase the latency between the RCC and the cells it coordinates, thus decreasing the throughput again.

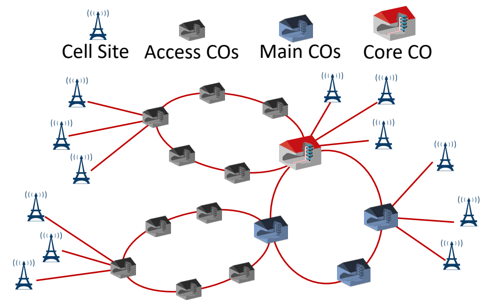


Fig. 1. 4-stages metro network architecture.

Hence, as for each cluster the location of its RCC impacts on the trade-off between clusters size and latency, CoMP throughput gain can be maximized by carefully choosing the RCCs placement according to network characteristics, such as the geographical size, cells distribution and density, available bandwidth etc. In addition, when choosing the optimal RCC placement, also the selection of proper coordination techniques (e.g., Coordinated Scheduling (CS) vs. Joint Transmission

A preliminary version of this paper is published in [1]. The authors are with Politecnico di Milano, Italy, e-mail: francesco.musumeci@polimi.it, ernesto.desilva@mail.polimi.it, massimo.tornatore@polimi.it.

(JT)) plays an important role as they provide substantially different throughput gain depending on cluster size and latency.

In this paper, we formally state and discuss the Clustering, Routing and RCC Placement (CRRP) optimization problem in an optical metro-access backhaul network. The problem has been originally introduced, and solved through Integer Linear Programming (ILP), in our previous work in [1]. Since ILP greatly limits problem scalability, in this paper we propose a scalable heuristic solution to solve the CRRP problem in much larger network instances. To confirm scalability of our approach, we evaluate our approach in a real mobile network topology, i.e., the Telecom Italia Mobile (TIM) metro-access network in Milan [4]. Moreover we show that, compared with the ILP-based methodology used in [1], our approach achieves near-optimal solution in all cases.

The main contributions of this paper are: $i)$ we model the CoMP throughput gain according to latency, cluster size and coordination technique; $ii)$ we describe in detail different backhaul/fronthaul solutions which can be used to implement CoMP in an optical metro-access network, emphasizing the different latency contributions which needs to be considered in the various cases; $iii)$ we formally state the CRRP problem and propose an efficient heuristic algorithm to solve it; $iv)$ we provide illustrative numerical results for various architectural and geographical scenarios, also including a real metro-access network topology, and show under which situations an optimized placement of RCCs across the metro-access network provides benefits with respect to a fixed placement.

We show that an optimized RCC placement with respect to network topology and traffic distribution may substantially enhance network throughput in comparison to a static placement of RCCs in Main COs or Core CO.

A. Paper Organization

The rest of the paper is organized as follows. In Sec. II we overview relevant related work on RAN and CoMP technology. In Sec. III we provide details of the RAN modeling and the CoMP coordination gain, showing the impact of different factors (such as network architecture, cells cluster size and latency) on network throughput. Section IV we formally state the CRRP problem and describe the heuristic approach used to solve it. Illustrative numerical results are shown and discussed in Sec. V. Finally, we draw paper conclusion in Sec. VI.

II. Related Work

Various recent studies have focused on the enhancement of throughput and, more in general, on effective resources utilization in 5G RANs. One main architectural trend for 5G RANs is represented by the C-RAN principle, which exploits the opportunity of separating baseband processing units (i.e., the BBUs) from basic radio processing, and concentrating BBUs in common locations. C-RAN enables significant CapEx/OpEx savings, due to facility sharing in COs and reduced maintenance and power costs at the antenna sites (see, e.g., [5] for a survey on the topic). However, C-RANs impose the transport of high-bit-rate traffic with strict latency requirements over the optical metro network (i.e., the *fronthaul*) [6]. Hence, a trade-off between network capacity and BBUs consolidation arises (see, e.g., [7] for an analysis of this problem). Many other aspects related to the challenges of fronthaul transport have been investigated: [8] investigates the power consumption, equipment cost and transport capacity needed in 5G fronthaul networks, [9] specifically focuses on fronthaul capacity requirements, and [10] focuses on C-RAN cost and energy savings, by proposing a strategy aiming at minimizing the number of active BBU pools (represented as sets of Virtual Machines), in Cloud RANs. More relevant to our work, authors of [11] show that better network capacity utilization can be obtained by jointly coordinating radio resources with optical transport resources.

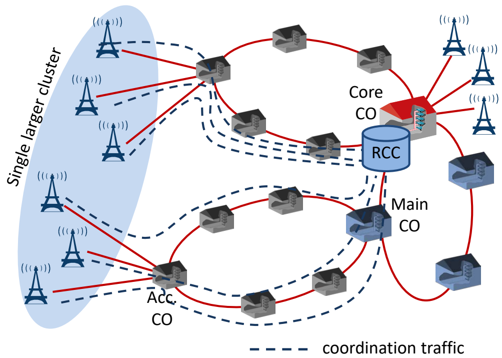
Other works have mainly focused on CoMP. E.g., in [12] the authors propose a mathematical model for resources allocation using CoMP in a Virtual Passive Optical Network to enhance the RAN throughput. CoMP performance is evaluated in [13], [14] for various CoMP techniques, and in [15], where the device-to-device communication is exploited to enhance network usage. The effectiveness of CoMP has been evaluated also in consortia of mobile network operators [16], [17]. A specific CoMP technique, i.e., Coordinated Scheduling is investigated in [18], where the authors study the impact of different Virtual Network Function deployments on the CoMP-enabled advantages, in terms of convergence delay (i.e., the time taken by an updated information to reach all the eNodeBs belonging to the cooperating set) and traffic overhead. Authors of [19] study the effectiveness of CoMP on an end-user level, i.e., they study the users-cells association and emitted power allocation required to maximize data rate provisioning for small cells users. The idea of dynamic association between RRH and BBU (namely, "any-RRH to any-BBU"), obtained by means of flexible optical fronthaul lightpaths reconfiguration is exploited in [20] to improve the performance of CoMP service. Finally, the impact of CoMP utilization and specifically of the required coordination traffic onto the C-RAN is investigated in [21], where the authors evaluate the benefit of data compression in the context of wireless fronthauling. To the best of our knowledge, no previous study has investigated how the impact of RCC placement over optical metro-access network on the CoMP throughput gain, which is the focus of our paper.

III. RAN and CoMP Coordination Gain Modeling

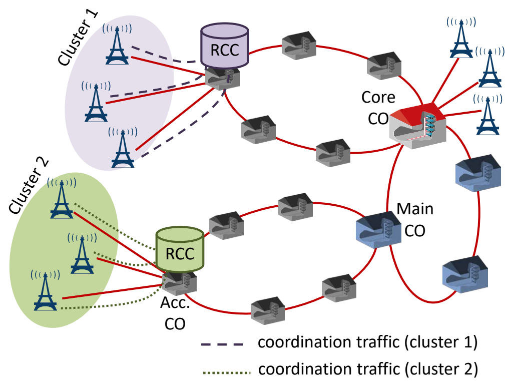
In the following, we model the CoMP throughput gain used for our analysis, and provide detailed discussion of the main factors affecting CoMP throughput gain, i.e.: $i)$ latency and cluster size impact on the gain, $ii)$ distributed vs centralized coordination, $iii)$ handling of X2 traffic.

A. Latency vs Cluster Size Trade-off

To highlight the trade-off between cluster size and latency, we use the example in Fig. 2, where two different RCC placement options are shown for the coordination of 6 cell sites. In one case (Fig. 2(a)) only one large cluster is formed, but higher latency is experienced by coordination traffic. Instead, in the case of Fig. 2(b), two smaller clusters are formed, allowing to place the RCCs closer to the cell sites.



(a)



(b)

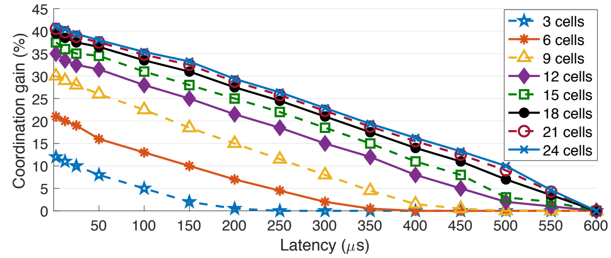
Fig. 2. Example of RCC placement and impact on cluster size and coordination traffic latency.

In this context, an optimized choice for 1) cells clustering, i.e., *how to group cells*, 2) RCCs placement, i.e., *where to place RCC for each cluster* and 3) CoMP technique selection, i.e., *which coordination strategy shall be used for each cluster* is crucial for network throughput maximization.

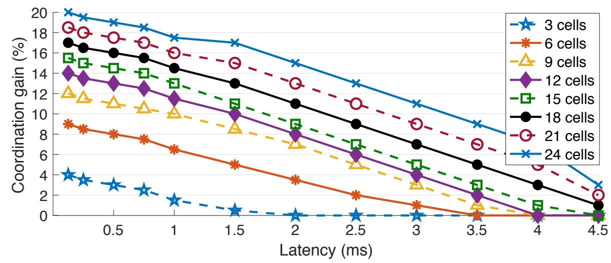
Several types of CoMP techniques have been defined [16], [22], that differ based on the amount and type of X2 coordination information, on their latency requirements and, most importantly, on the provided throughput gain (i.e., how much additional users traffic can be served thanks to coordination).

In the following, we model the CoMP throughput gain of two downlink CoMP techniques, i.e., Coordinated Scheduling (CS) and Joint Transmission (JT). To this end, we derived the throughput gain curves in Fig. 3, which quantify the percentage gain as a function of cluster size (from 3 to 24 cells) and X2 traffic latency, t_{X2} (i.e., maximum distance between RCC and corresponding cell sites). The curves have been obtained by interpolating the measured throughput gains reported in [16]. As expected, coordination gain increases for larger clusters and decreases for higher latency. Note that i) the observed latency range for CS (up to 4.5 ms) is larger than the one of JT (up to 600 μs), as JT has much stricter latency requirements, and ii) in general, JT has much higher throughput gain (up to around 40% for 24-cells clusters and 5 μs latency) in comparison to CS, whose gain achieves at most 20% for 24-cells clusters and 100 μs latency.

To evaluate the throughput gain for a given cluster c, we consider the gain achieved in terms of total S1 traffic generated by all nodes $n \in c$. Therefore, the overall throughput THR_c



(a) Joint Transmission - JT



(b) Coordinated Scheduling - CS

Fig. 3. Throughput gain for varying CoMP techniques, cluster size and latency (RTT delay, t_{X2}).

enabled by coordination gain in cluster c can be expressed as:

$$THR_c = Gain(c) \cdot T_c = Gain(c) \cdot \sum_{n \in c} t_n, \qquad (1)$$

where t_n is the backhaul (S1) traffic for a given node n, and T_c is the S1 traffic generated by all cells $n \in c$. The term $Gain(c)$ represents the *additional* percentage throughput for cluster c, compared to the case where no CoMP technique is used (in which case $Gain(c) = 0$). For cluster c, $Gain(c)$ is a function depending on cluster size $|c|$, latency t_{X2} and used coordination technique (CS vs JT), according to Fig. 3, i.e.:

$$Gain(c) = Gain(|c|, t_{X2}, tech.). \qquad (2)$$

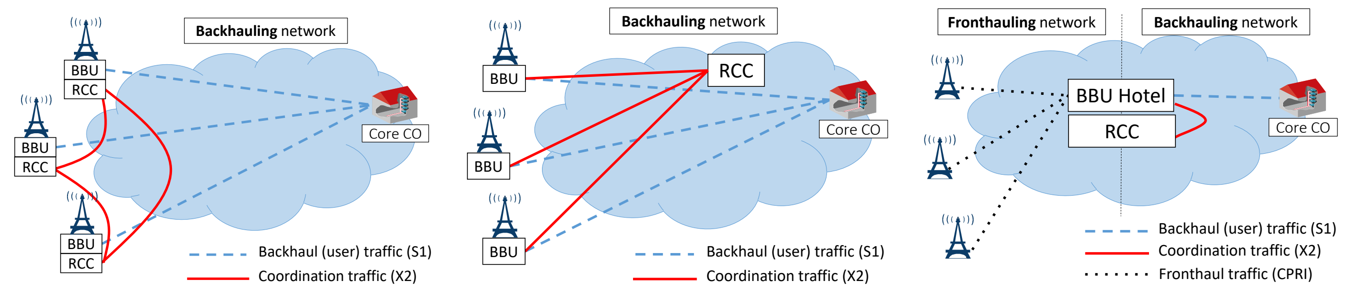
As an example, referring to Fig. 3, assume that a certain cluster $\hat{c}$ consists of 12 cells and the maximum X2 latency between any of these cells and the RCC of cluster $\hat{c}$ is $t_{X2} = 300\mu s$. If JT coordination is adopted, the percentage throughput gain is equal to $Gain(\hat{c}) = Gain(12, 300\mu s, JT) = 15\%$ (see Fig. 3(a)). On the other hand, if the RCC is located so that $t_{X2} = 450\mu s$, adopting JT coordination produces percentage throughput gain is equal to $Gain(\hat{c}) = Gain(12, 450\mu s, JT) = 5\%$. In this situation (i.e., cluster $\hat{c}$ with 12 cells and maximum X2 latency $t_{X2} = 450\mu s$), adopting CS coordination is more beneficial, as the throughput gain is in the order of $Gain(\hat{c}) = Gain(12, 450\mu s, CS) \simeq 13\%$.

Finally, the overall network throughput can be given as the sum of throughput values for all clusters in the network, i.e.:

$$THR = \sum_{c} THR_c \qquad (3)$$

B. Distributed vs Centralized Coordination Architectures

CoMP coordination can be achieved using *distributed* or *centralized* architectures. In the distributed architecture (Fig. 4(a)), each cell hosts a RCC, and the RCCs exchange coordination information (mainly on radio-channel quality) among themselves via the X2 interface. Instead, users' traffic is



(a) Distributed CoMP. (b) Centralized CoMP. (c) Centralized CoMP with BBU hotelling.

Fig. 4. Distributed and centralized CoMP architectures.

backhauled from cell sites to the Core CO using the S1 interface. Conversely, in a centralized architecture (see Fig. 4(b)), RCCs are placed within the metro-access network to collect coordination information from several cell sites and provide integrated coordination decisions. The centralized CoMP coordination solution can also be realized using a C-RAN architecture (see Fig. 4(c)). In this case, we assume that the RCC for a given cluster is co-located with the BBU hotel for the cells in that cluster. At the hotel, a BBU pool processes user and coordination traffic coming from several cell sites and transported via a fronthaul network (e.g., via CPRI interface). In such scenario, after implementing baseband signal processing, users traffic is backhauled towards the Core CO, while coordination traffic is transmitted to the RCC.

In this study we refer to centralized CoMP architecture, as our objective is to study how the placement of RCCs within the metro-access network impacts the CoMP throughput gain.

C. Coordination Traffic Latency Contributions

To model the CoMP coordination gain, the adopted network architectural solution (i.e., fronthaul vs backhaul) plays a key role, as in the various cases different latency contributions must be taken into account. In our evaluations we consider three different solutions for the transport of S1 and X2 traffic, that is: $i)$ *Backhaul-common*, where coordination and user traffic flows are aggregated into a common physical interface and then separated at the RCC node; $ii)$ *Backhaul-separate*, where coordination traffic is transported on a dedicated interface; $iii)$ *Fronthaul*, where both coordination and user information are transported via a fronthaul interface between cell sites and the BBU, where they are separated to reach the RCC and Core CO, respectively. The three solutions are shown in Fig. 5, along with the impact of latency in each scenario.

In case of *Backhaul-common* (Fig. 5(a)), an electronic switch, co-located with the RCC, separates S1 and X2 traffic. Thus, the total coordination latency can be expressed as:

$$t_{X2,backhaul-common} = \tau + t_{sw} + t_{RCC}. \quad (4)$$

The *propagation delay*, τ, depends on the length of the traversed links, t_{sw} is the *switching delay* introduced by the switch at the RCC node[1], and t_{RCC} is the RCC processing delay introduced when processing coordination information.

[1]As current technologies (e.g., Ethernet, OTN) feature switching delays that are too far from CoMP requirements, we assume "low-latency" switches, tailored for CoMP applications and introducing a fixed 20 μs delay, as in [7].

For the *Backhaul-separate* case, coordination traffic directly reaches the RCC with no need to be de-aggregated at the RCC node; so, the total latency of coordination traffic is (Fig. 5(b)):

$$t_{X2,backhaul-separate} = \tau + t_{RCC} \quad (5)$$

Finally, in the *Fronthaul* (i.e., C-RAN) scenario, a fronthaul interface is used to transport users and coordination traffic between cell sites and the BBUs, where flows are separated, so that coordination traffic can be directed to the co-located RCCs; in this case, as shown in Fig. 5(c), the only contribution to total coordination latency is the RCC processing delay, as propagation delay can be neglected due to the fact that the RCC and BBU are co-located, i.e.:

$$t_{X2,fronthaul} = t_{RCC} \quad (6)$$

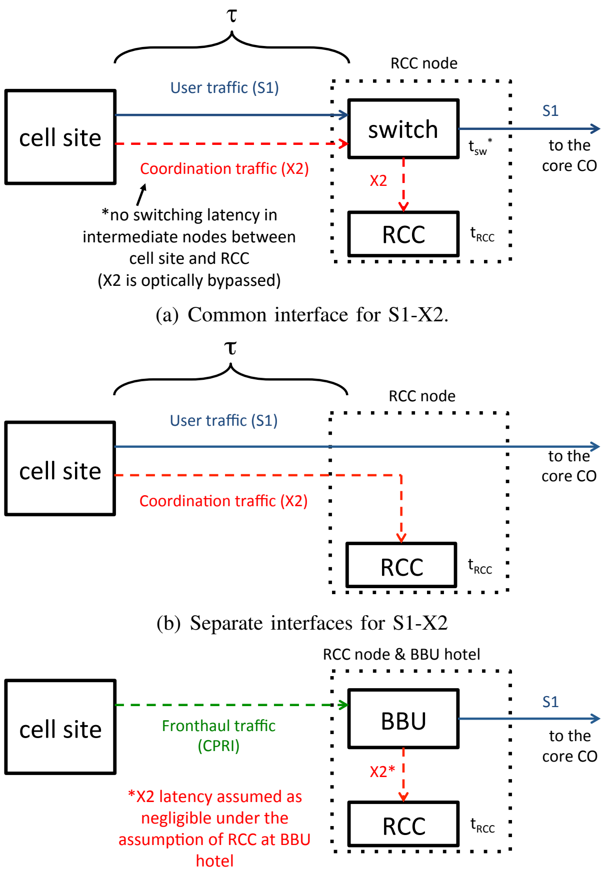
For our numerical evaluations we assume two alternative values for this contribution, i.e., $t_{RCC} = 0$ or $t_{RCC} = 250\mu s$, corresponding to the case when RCC processing time is negligible or to a conservative scenario in which higher processing delay must be considered due to high load, respectively.

As a future work, we plan to consider also the effect of intermediate solutions between backhaul and fronthaul architectures, i.e., the so-called *RAN splits*, for which different coordination latency values should be considered, thus producing a different impact on the coordination gain.

IV. The Clustering, Routing and Controller Placement Problem

A. Problem Statement

The Clustering, Routing and RCC Placement (CRRP) problem addressed in this paper can be summarized as follows. **Given** 1) a hierarchical multi-stage metro network topology, represented by a graph $G(N, E)$, where N is the set of nodes (i.e., COs and CSs) and E the set of fiber links, 2) the set of backhaul (S1) traffic requests directed to the Core CO, 3) the set of possible CoMP techniques and their throughput gain as function of cluster size and X2 latency, we **decide** the RCCs placement (in which node in G, the RCC shall be placed), the cells clustering, the coordination technique adopted for each cluster and the routing of S1 and X2 traffic, **maximizing** the overall throughput gain in the network, **constrained by** $i)$ network links capacity, $ii)$ latency between cell sites and RCCs, $iii)$ routing of all the requests.



(a) Common interface for S1-X2.

(b) Separate interfaces for S1-X2

(c) BBU hotelling

Fig. 5. Schematic representation of X2 latency contributions.

B. Heuristic Approach

CRRP is highly complex optimization problem. In [1] we provided a mathematical formulation for this problem and proposed an ILP-based approach where pre-computed clusters are fed to an ILP model that accomplishes RCC placement and (S1 and X2) traffic routing (please refer to [1] for details of the ILP). However, to guarantee that the optimal (i.e., maximum-throughput) solution is found, all the possible clusters shall be precomputed [1] and this leads to limitations in the size of the networks to be analyzed, as the number of available clustering options grows exponentially with number of nodes.

In this paper we provide a two-step heuristic algorithm performing the following tasks: 1) optimized clustering, RCC placement and CoMP technique selection; 2) S1 and X2 traffic routing. In the first step the algorithm gradually creates cells clusters aiming at maximizing the overall network throughput gain. In this process, information on the traffic originated by the cells are also included, so that the optimal RCC placement and CoMP technique to be selected for each cluster can be greedily adapted. In the second step, S1 (respectively, X2) traffic is routed with no constraints of network capacity, along the shortest-path between cell sites and Core CO (resp., cell sites and RCC). Details of step 1) are shown in Algorithm 1.

We start with an initial clustering where each cluster is composed by an Access CO (where we place the RCC) and all directly-connected cell sites[2]. Then, for each cluster c we

[2]Note that Main COs and the Core CO can have directly-connected cells.

Algorithm 1 Optimized clustering, RCC placement and CoMP technique selection.

INPUT: Network topology: E links, N nodes, N_c cells, N_a Access COs, N_m Main COs, 1 Core CO ($N = N_c + N_a$); S1 traffic, generated by each node $n \in N$, t_n; CoMP gain curves; K: nr. of candidate cell sites to consider for inclusion in a cluster.
OUTPUT: Clustering, RCC placement and CoMP technique selection.

<u>**Initialize clusters**</u>:
1: Start with N_a clusters, grouping each Access CO with all its directly-connected cells
2: Insert initial clusters in a list C
3: **for all** $c \in C$ **do**
4: Place RCC in each cluster at the Access CO
5: Compute tot. traffic of cells in cluster c, $T_c = \sum_{n \in c} t_n$
6: **end for**
7: Sort clusters in C in descending order of T_c
<u>**Increase clusters dimension**</u>:
8: **while** C is not empty **do**
9: Select the first cluster in C, $c*$
10: **while** $c*$ is in C **do**
11: Compute cluster gain $Gain(c*)$
12: Find the K nearest nodes to cluster $c*$ (nodes must
13: be in any of the clusters in C)
14: **for** $i = 1$ to K **do**
15: Find cluster of node i, c_i
16: Compute gain in the cluster $Gain(c_i)$
17: Create temporary clusters by including i in $c*$ and
18: removing it from c_i, $c*_{temp}$ and $c_{i,temp}$
19: **for** $c*_{temp}$ and $c_{i,temp}$ **do**
20: Place each RCC in the node which
21: minimizes the maximum latency
22: Set the best CoMP technique given new values
23: of cluster size, latency and curves in Fig. 3
24: Set $Gain_i = Gain(c*_{temp}) + Gain(c_{i,temp})$
25: **end for**
26: **end for**
27: Find $j = argmax_i(Gain_i)$ and set $Gain_j = Gain_i$
28: **if** $Gain_j > Gain(c*) + Gain(c_i)$ **then**
29: Set $c* = c*_{temp}$ and $c_i = c_{i,temp}$
30: **else**
31: Remove $c*$ from C
32: **end if**
33: **end while**
34: **end while**

compute the total traffic generated by its cells, T_c. Clusters are then stored in a list, where they are sorted in descending order of T_c (lines 1-7). Then, clusters enlarging takes place (lines 8-34), aiming at including new nodes to clusters with higher T_c. Starting from the first cluster in C, say $c*$, the K outer nearest nodes are considered[3](lines 12-13), and for each node $i = 1, 2, ...K$, we evaluate the gain of the clusters obtained by removing node i from its original cluster c_i and adding it to $c*$ (lines 14-26)[4]. After scanning all nodes $i = 1, 2, ...K$, we select the node providing the highest gain for the two clusters and compare it with the original solution, i.e., with the case where $c*$ has not been added any node (lines 27-28)[5]. If enlarging cluster $c*$ is beneficial, we update the two clusters

[3]Given cluster c, the outer nearest node is the one having the lowest average line-of-sight distance to the nodes in c.

[4]Note that $c*_{temp}$ and $c_{i,temp}$ are disjoint clusters.

[5]In our evaluation we consider $K = 10$, although very similar results can be obtained with $K = 4$ or $K = 5$ in most cases.

(line 29) and repeat the steps from line 11 for the updated cluster $c*$, otherwise we remove $c*$ from list C and consider $c*$ as permanently determined, i.e., we cannot remove/add any node from/to it (line 31). Otherwise, we keep unchanged clusters $c*$ and c_i and remove $c*$ from list C (line 31). Note that, in case the node included into $c*$ hosted the RCC in its original cluster c_i, we include all the nodes of c_i into $c*$.

To estimate the algorithm complexity consider that, after the initialization phase, which requires $O(N_a)$ iterations, the while cycle in line 8 is performed at most N_a times; then, for each cluster $c*$, cycle at line 10 is run until $c*$ is contained in list C, i.e., at most when all outer nodes are added to $c*$ (i.e., at most N times); finally, every time one node is added to a cluster, the K nearest cell sites shall be considered as candidates for inclusion in the cluster. Therefore, the overall algorithm complexity can be estimated as $O(N_a \cdot N \cdot K)$.

V. Illustrative Numerical Results

In this section we first validate our heuristic algorithm, comparing it with the ILP approach in [1], then we use the heuristic approach to solve the CRRP problem in larger network instances using the Net2Plan Java tool [23]. We compare the proposed optimized placement ("*RCC opt.*") with two static placement strategies: 1) "*RCC in main COs*", where RCCs are placed in all Main COs and each RCC coordinates the cells connected to that Main CO; 2) "*RCC in core CO*", where all RCCs are placed at the Core CO, i.e., we assume that computing capabilities are only available at that node.

To perform our evaluation we use the following two metrics: a) overall throughput gain [in %], G_p, calculated as the percentage increase of network throughput (see eq. (3)), obtained with respect to a baseline case where CoMP is not used; b) number of utilized wavelengths, W_u.

A. Case-Study Description

We consider different realizations of a hierarchical ring-and-spur topology as in Fig. 1. Each node in the network, including the COs, inserts (respectively, receives) mobile traffic (e.g., UMTS, LTE or LTE-A), directed to (resp., originated by) the Core CO. The amount of S1 traffic generated by each node is randomly selected in the range 155-400 Mb/s, in line with the traffic generated by an LTE site with three sectors, 20 MHz spectrum, 2×2 MIMO and 16 QAM modulation format. Moreover, we assume that, according to the selected CoMP technique, different amounts of coordination (X2) traffic must be transported between the cell sites and the corresponding RCCs. Namely, the coordination data for CS requires to exchange the LTE Channel State Information (CSI), so we considered an overhead of about 16% with respect to S1 traffic. As for JT CoMP technique, X2 traffic is assumed as 116% of S1 bit rate, as in this case, besides the CSI, also user traffic (S1) is transported with coordination traffic (X2) [16].

As a first step, to perform validation of the heuristic against ILP results, we consider small topologies with 12, 15 and 18 nodes (as in [1]), where nodes are distributed over a square area with size 15 km^2, corresponding to a urban geotype. For the 15-nodes topology, we also consider the sub-urban and rural geotypes [24]. Then, we consider several synthetically-generated topologies with 50 nodes in a urban scenario, i.e., with nodes distributed on a 166 km^2 area. To generalize our assessments, we consider four different interconnection degrees in the network, i.e., the 50-nodes ring-and-spur topologies are generated such that the number of metro rings (i.e., rings interconnecting Access COs as in Fig. 1) is increased from 1 to 4. In the following, we refer to each of these four cases to as a "scenario". For each scenario we consider 35 randomly generated topologies, each associated to three different traffic matrices, obtaining a total of 105 cases per scenario. Our plotted results are with 95% statistical confidence on a 5% confidence interval. For the scenario with 50 nodes and 4 metro-access rings, we also consider four different geotypes, namely dense-urban, urban, sub-urban and rural, corresponding to area with size 12, 166, 263, 625 km^2, respectively. Note that, for all the considered topologies, we identify the different geotypes according to the cell-sites density valus, i.e., i) dense-urban, density ranging between 1.2 and 4.47 BSs per km^2, ii) urban, between 0.3 and 1.2 BSs per km^2, iii) suburban, between 0.08 and 0.3 BSs per km^2, and iv) rural, between 0.019 and 0.08 BSs per km^2. Note that, for a given number of nodes in the network, changing the geotype has an impact on the distance between nodes, hence on the latency to be considered for X2 traffic transport. All synthetic topologies are generated by a hierarchical clustering procedure which runs the k-means algorithm, as explained in [25].

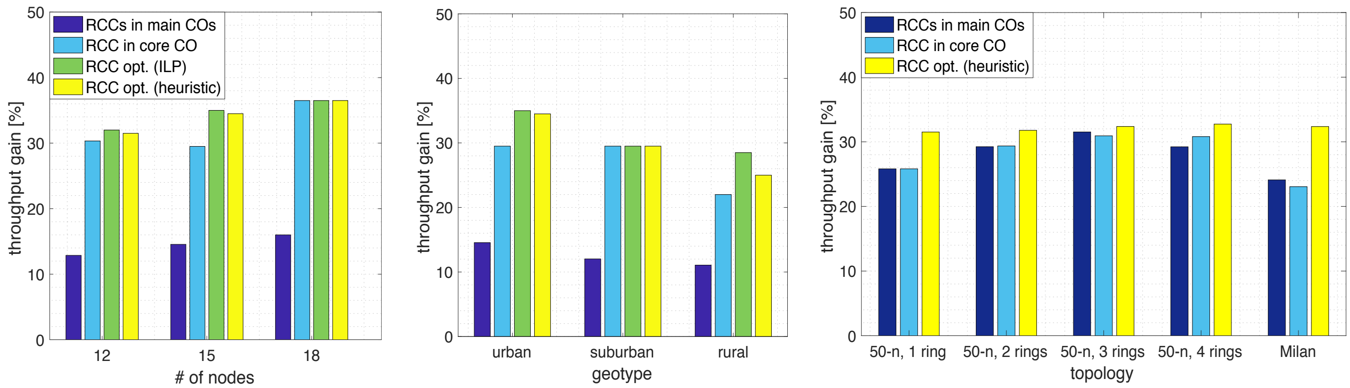
Finally, we consider a real network topology, i.e., the TIM metro-access network in Milan, consisting of 2548 cell sites, 110 Access COs, 15 Main COs, interconnected through a single ring to the Core CO [4]. This topology can be considered as a dense-urban topology, as it has a cell-site density of around 4.47 cells per km^2. For this topology, we consider a total of 30 different randomly-generated traffic matrices.

B. Discussion

For the initial evaluation, as the RAN architecture we consider the *Backhaul-common* described in III-C, and assume that RCC processing time is negligible, i.e., $t_{RCC} = 0$, so only propagation (τ) and switching (t_{sw}) delays are accounted for X2 latency. However, in a second phase we will also show the impact of a non-negligible RCC processing delay.

Heuristic approach validation

Figure 6(a) shows the throughput gain G_p of the proposed heuristic approach and of the optimal ILP-based solution of [1]. We also report G_p of the fixed placement strategies *RCC in main COs* and *RCC in core CO*, for increasing number of nodes in a urban scenario. We can observe that the proposed heuristic approach provides near-optimal solutions in all cases (i.e., the optimality gap is below 1%), especially for larger network instances, where the difference with respect to the ILP-based approach reaches zero. On the other hand, the heuristic approach provides substantial advantages in terms of solving time. For smaller network instances, the solution is obtained in less than one minute with the heuristic approach, whereas the ILP solution requires more than twenty minutes, and this difference rapidly increases with network size. Figure 6(b)



(a) Urban scenario, $t_{RCC} = 0$ (b) 15-nodes network, $t_{RCC} = 0$ (c) 50-nodes network and TIM Milan topology, $t_{RCC} = 0$

Fig. 6. Throughput gain for various RCC placement strategies. (a)-(b) Comparison ILP-vs-Heuristic for increasing network size in urban geotype (a) and for 15-nodes network in different geotypes (b). (c) Throughput gain for different RCC placement strategies in 50-nodes (50-n) and TIM Milan topologies.

shows the comparison between the aforementioned strategies in the case of 15-nodes network for the different geotypes. Also in this case the heuristic approach performs very close to the ILP approach, with a maximum gap, in the case of rural geotype, below 3.5%. Note that *RCC opt* always provides higher gains than *RCC in main COs* and *RCC in core CO*.

Effect of network topology

Now we move our analysis towards larger and more realistic network topologies using our heuristic approach. In Fig. 6(c) we compare the different RCC placement strategies for the 50-nodes topologies ("50-n" in the figure) in urban scenario and for the TIM Milan topology. Results show that, in all cases *RCC opt.* provides significantly higher gain with respect to fixed RCC placement strategies, with a G_p absolute increase ranging from 2% for 50-nodes networks up to 8% for Milan topology[6]. Moreover, we notice that, for 50-nodes topologies, higher gain improvement is achieved for larger number of rings interconnecting Access COs. This is due to the fact that, in topologies with only few rings (in an extreme case, only one ring), longer routes must be used to route coordination traffic, thus providing a negative effect to throughput gain due to higher propagation latency. Note that, in the case of only one metro-access ring, there is only one Main CO, i.e., the Core CO, therefore the results for *RCC in main COs* and *RCC in core CO* coincide. This motivates the deployment of more meshed metro networks for 5G backahul, which is in fact a trend being experienced by operators nowadays. It is worth noting that if the number of metro-access rings increases after a certain threshold (e.g., above 3 rings in the case under analysis), throughput gain tends to decrease for *RCC in core COs* and *RCC in main CO* strategies. In the former case, this decrement is due to the fact that having too many rings increases the maximum latency between the RCC, placed at the Core CO, and the farthest cell sites. In the latter case, more rings corresponds to a solution with a large number of very small clusters, that can only achieve lower gains.

Fig. 6(c) shows the case of Milan topology, corresponding to a dense-urban geotype. The *RCC opt.* strategy provides substantial advantages with respect to the other strategies. However, in this case the *RCC in core CO* has lower gain compared to *RCC in main COs* strategy, as it suffers from larger geographical dimensions of the physical topology, which only allows coordination among relatively small clusters.

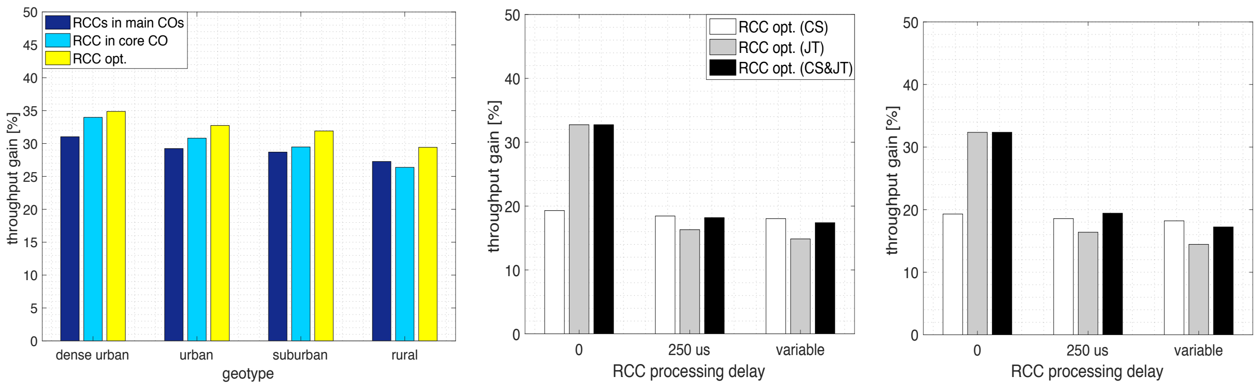
Effect of geotype and latency contributions

To evaluate the impact of cell-sites density we consider the 50-nodes topology with 4 metro-access rings and compare the throughput gain obtained with the three RCC placement strategies in different geotypes, as shown in Fig. 7(a). We observe that reducing cell sites density, i.e., moving towards sparser topologies as in rural geotypes, the throughput gain gradually decreases for all placement strategies. As expected, *RCC opt.* always outperforms the other RCC placement strategies. However, for larger and sparser networks, as in the suburban and rural geotypes, the *RCC in main CO* provides improvements in comparison to *RCC in core CO*, as the higher latency between RCCs and cell sites obtained with *RCC in core CO* is not compensated by the opportunity of increasing cluster size, due to the lower cell density.

To show the impact of RCC processing delay we compare three RCC processing delay values under the *RCC opt.* placement strategy, namely, $t_{RCC} = 0$, $t_{RCC} = 250\mu s$ and "variable" t_{RCC}, corresponding to a case where the RCC processing delay is assumed as proportional to the cluster size[7]. For this comparison we also consider three possibilities for the selection of CoMP coordination technique, i.e.: 1) "CS" and 2) "JT", where all the clusters are coordinated using only CS or JT CoMP technique, respectively, and 3) "CS&JT", where an optimal selection of CoMP technique is made on a cluster basis, according to the specific latency and size of each cluster. This evaluation is performed for the 50-nodes 4-rings synthetic topology in a urban scenario (Fig. 7(b)) and for the Milan topology (Fig. 7(c)). As expected, for both topologies the overall gain decreases when t_{RCC} is not negligible. Moreover, while for $t_{RCC} = 0$ JT is always the most convenient technique, when latency becomes stringent, the ability of choosing CoMP technique on a cluster basis enables higher flexibility and improved throughput gain.

[6]Note that, although in some cases throughput gain is limited, this gain is obtained with low or no additional investment for the network operator, as controller shall be anyway deployed.

[7]We assume that for a cluster with size $|c|$ the RCC processing delay is $t_{RCC} = (|c| - 1) \cdot 20\mu s$.



(a) 50-nodes, 4 rings, $t_{RCC} = 0$ (b) 50-nodes, 4 rings, urban scenario (c) TIM Milan topology

Fig. 7. Throughput gain for *RCC opt.* placement strategy. (a)-(b) 50-nodes networks: (a) impact of geotype; (b) impact of RCC processing delay and CoMP technique selection. (c) Impact of RCC processing delay and CoMP technique selection in TIM Milan topology.

Effect of RAN architecture

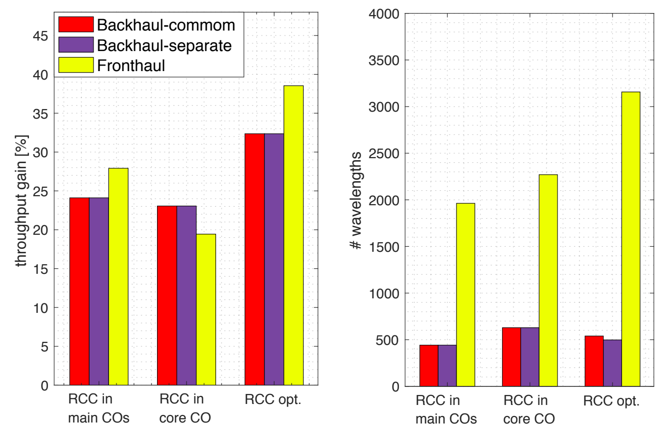
We now analyze the impact of CoMP architectures discussed in Sec. III-B. In this context we compare three different flavours of centralized CoMP techniques, namely: 1) Backhauled coordination with common S1-X2 interface (*Backhaul-common*), 2) Backhauled coordination with separate S1-X2 interfaces (*Backhaul-separate*), where we assume $t_{sw} = 0$, and 3) Fronthauled coordination (*Fronthaul*), where t_{RCC} is the only latency contribution impacting throughput gain.

Fig. 8 shows the throughput gain and number of wavelengths needed for the three CoMP architectures in the case of Milan topology and for the different RCC placement strategies.

As expected, when using the *RCC opt.* strategy, the throughput gain tends to increase for all CoMP architectures. Such increase is at least 10% (as absolute difference between the percentage gains) in case of backhaul solutions, and reaches 20% when adopting the fronthaul architecture.

Moreover, the lowest gain is obtained in all cases when adopting *RCC at core CO*, due to the fact that the considered topology allows only relatively small-size clusters with high latency between cells and each RCC. Despite the fronthauled solution provides in general the highest gain, due to the fact that only RCC processing delay impacts on throughput gain, this is not the case for the *RCC at core CO* strategy, as in this case only a small subset of cells can be coordinated by an RCC at the core CO, due to the highly stringent fronthaul latency constraint (to be applied between cell sites and the RCC/BBU hotel at the Core CO), which is in the order of few hundreds of μs [7]. For the considered topology, there is no evident throughput gain distinction between the two backhauled solutions as the switching latency contribution has a limited impact in a dense-urban scenario.

Also for the total number of wavelengths needed there is a very low difference between the two backhauled solutions. However, as expected, the number of wavelengths needed in the case of fronthaul architecture is much higher than in the other two cases, due to the fact that for the considered cell sites antenna configuration, a total of 7.5 Gb/s fronthaul traffic must be routed between each cell and the BBU hotel at the Core CO [7]. Moreover, for *RCC opt.*, much higher fronthaul traffic is required with respect to both other cases: compared to *RCC in core CO*, this is due to the larger number of clusters formed; on the other hand, *RCC opt.* requires more wavelengths with respect to *RCC in main COs* as in the former case placement of RCCs (and BBUs) in higher levels of the network hierarchy leads to fronthaul routing across more network links.



(a) (b)

Fig. 8. (a) Throughput gain and (b) nr. of wavelengths for different centralized CoMP architectures and RCC placement strategies (Milan topology).

Effect of parameter K in the heuristic algorithm

As described in Alg. 1, every time we try to increase the size of a cluster, we select one out of K candidate cell sites to include in the cluster. Hence, changing the value of K has an impact on the trade-off between solution optimality and the heuristic duration. To evaluate this trade-off, we perform a sensitivity analysis by increasing parameter K and considering the *RCC opt. (CS&JT)* scenario in the Milan topology (see Fig. 9). We observe that, increasing the value of K from 2 to 16, CoMP throughput gain increases from 27% up to around 33%, whereas heuristic execution time worsens substantially (i.e., it increases almost linearly with K). Moreover, significant improvement in throughput gain is observed up to $K = 10$, while after this point no relevant advantages are obtained. This suggests that, for the considered case study, values of K larger than 10 are not necessary.

VI. Conclusion

We address the Clustering Routing and Radio Controller Coordinator Placement (CRRP) problem in future RANs

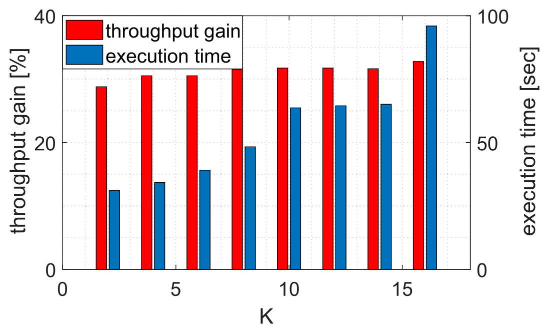


Fig. 9. CoMP throughput gain and execution time for increasing value of parameter K in Alg. 1 (Milan topology, *RCC opt. (CS&JT)*, $t_{RCC} = 0$).

where Coordinated Scheduling and Joint Transmission CoMP techniques are adopted to increase network throughput. To solve this problem, we provide a two-step heuristic approach consisting of 1) a first phase for for cells clustering and controller placement, aiming at maximizing overall network throughput, and 2) a second phase for routing of S1 and X2 traffic over a hierarchical metro-access network. We evaluated our approach and quantified the gains achievable with an optimized placement of RCCs at different nodes in both synthetic and realistic network topologies, analyzing the impact of several parameters affecting the overal throughput, i.e., CoMP technique, network size and topology, cell sites density and metro-access network technology (i.e., backhaul vs fronthaul).

We found that, i) for the considered scenarios, an optimized clustering and RCC placement provides up to 20% throughput enhancement, especially when C-RAN (i.e., fronthaul) architecture is adopted, ii) in urban and dense-urban scenarios the optimized RCC placement provides up to 10% gain increase with respect to fixed RCC placement in Main COs and Core COs, iii) when considering sparse and large networks, especially rural environments, the choice of proper CoMP technique is not trivial, as latency plays a crucial role, and substantially affects the opportunity of improving network throughput via CoMP, iv) adopting C-RAN in a dense-urban scenario enables higher gain with respect to traditional distributed (i.e., backhaul) RANs, however it requires high network capacity utilization due to fronthaul traffic.

As future work, we plan to consider also the effect of intermediate solutions between backhaul and fronthaul architectures, i.e., the so-called *RAN splits*, for which different coordination latency and transport traffic values should be considered, thus producing a different impact on the coordination gain and on the required network capacity.

Acknowledgment

The work leading to these results has been supported by the European Community under grant agreement no. 761727 Metro-Haul project.

References

[1] F. Musumeci, C. Bellanzon, M. Tornatore, A. Pattavina, and J. T. Gijon, "Enhancing RAN throughput by optimized controller placement in optical metro networks," in *IEEE International Conference on Communications (ICC) 2017*, May 2017, pp. 1–6.

[2] Next generation Mobile Network (NGMN) Alliance, "5G White Paper - Executive Version," Dec. 2014.

[3] M. M. Do and H. J. Son, "CoMP (1): CoMP types - CS, CB, JT and DPS," Netmanias Tech-Blog, available online at http://www.netmanias.com, Aug. 2014.

[4] "Open Big Data," available online at https://dandelion.eu/datamine/open-big-data/, accessed Mar. 2018.

[5] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for Mobile Networks: A Technology Overview," *IEEE Communications Surveys Tutorials*, vol. 17, no. 1, pp. 405–426, Firstquarter 2015.

[6] A. Pizzinat, P. Chanclou, F. Saliou, and T. Diallo, "Things You Should Know About Fronthaul," *IEEE/OSA Journal of Lightwave Technology*, vol. 33, no. 5, pp. 1077–1083, 2015.

[7] F. Musumeci, C. Bellanzon, N. Carapellese, M. Tornatore, A. Pattavina, and S. Gosselin, "Optimal BBU Placement for 5G C-RAN Deployment Over WDM Aggregation Networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1963–1970, Apr. 2016.

[8] M. Fiorani, B. Skubic, J. Mårtensson, L. Valcarenghi, P. Castoldi, L. Wosinska, and P. Monti, "On the design of 5G transport networks," *Photonic Network Comm.*, vol. 30, no. 3, pp. 403–415, Dec. 2015.

[9] L. Valcarenghi, K. Kondepu, F. Giannone, and P. Castoldi, "Requirements for 5G fronthaul," in *International Conference on Transparent Optical Networks (ICTON) 2016*, July 2016.

[10] J. Tang, W. P. Tay, T. Q. S. Quek, and B. Liang, "System cost minimization in cloud ran with limited fronthaul capacity," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3371–3384, May 2017.

[11] M. R. Raza, M. Fiorani, A. Rostami, P. Öhlen, L. Wosinska, and P. Monti, "Demonstration of dynamic resource sharing benefits in an optical C-RAN," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 8, no. 8, pp. 621–632, Aug 2016.

[12] X. Wang, C. Cavdar, L. Wang, M. Tornatore, Y. Zhao, H. S. Chung, H. H. Lee, S. Park, and B. Mukherjee, "Joint Allocation of Radio and Optical Resources in Virtualized Cloud RAN with CoMP," in *IEEE International Global Telecommunications Conference (GLOBECOM) 2016*, Washington D.C., USA, Dec. 2016.

[13] A. Davydov, G. Morozov, I. Bolotin, and A. Papathanassiou, "Evaluation of Joint Transmission CoMP in C-RAN based LTE-A HetNets with large coordination areas," in *IEEE Globecom Workshops 2013*, Dec. 2013, pp. 801–806.

[14] J. Liu, M. Sheng, T. Q. S. Quek, and J. Li, "Comp Transmission in Cloud Radio Access Networks," in *IEEE Globecom Workshops 2015*, Dec. 2015.

[15] ——, "D2D Enhanced Co-Ordinated Multipoint in Cloud Radio Access Networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4248–4262, June 2016.

[16] "Coordinated Multi-Point operation for LTE physical layer aspects (release 11)," 3GPP Technical report 36.874, 2013.

[17] Next generation Mobile Network (NGMN) Alliance, "RAN Evolution Project - CoMP Evaluation And Enhancement," Mar. 2015.

[18] A. Marotta, K. Kondepu, F. Giannone, D. Cassioli, C. Antonelli, L. Valcarenghi, and P. Castoldi, "Impact of CoMP VNF placement on 5G Coordinated Scheduling performance," in *European Conference on Networks and Communications (EuCNC) 2017*, June 2017, pp. 1–6.

[19] L. Ferdouse, A. Alnoman, A. Bulzacki, and A. Anpalagan, "Energy Efficient Multiple Association in CoMP Based 5G Cloud-RAN Systems," in *IEEE Vehicular Technology Conference (VTC-Fall 2017)*, Sept 2017, pp. 1–5.

[20] J. Zhang, Y. Ji, H. Yu, X. Huang, and H. Li, "Experimental demonstration of fronthaul flexibility for enhanced CoMP service in 5G radio and optical access networks," *Optics Express*, vol. 25, no. 18, pp. 21 247–21 258, Sep. 2017. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-25-18-21247

[21] Y. Qi, M. Z. Shakir, M. A. Imran, K. A. Qaraqe, A. Quddus, and R. Tafazolli, "Fronthaul data compression for Uplink CoMP in cloud radio access network (C-RAN)," *Transactions on Emerging Telecommunications Technologies*, vol. 27, no. 10, pp. 1409–1425, Aug. 2016.

[22] "Coordinated multi-point operation for LTE with non-ideal backhaul (release 12)," 3GPP Technical report 36.819, 2013.

[23] http://www.net2plan.com.

[24] *Analysis of transport network architectures for structural convergence*, European Project "COMBO", Deliverable D3.3, available at http://www.ict-combo.eu, 2015.

[25] N. Carapellese, M. Tornatore, and A. Pattavina, "Energy-Efficient Baseband Unit Placement in a Fixed/Mobile Converged WDM Aggregation Network," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 8, pp. 1542–1551, Aug 2014.