



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *ICC*.

Citation for the original published paper:

Alabbasi, A. (2017)

Interplay of Energy and Bandwidth Consumption in CRAN with Optimal Function Split.

In:

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-211472>

Interplay of Energy and Bandwidth Consumption in CRAN with Optimal Function Split

[†]Xinbo Wang, [‡]Abdulrahman Alabbasi, [‡]Cicek Cavdar

[†]Computer Science Department, University of California Davis, USA

[‡]Communication Systems Department, KTH Royal Institute of Technology, Sweden

Email: [†]xbwang@ucdavis.edu, [‡]{alabbasi, cavdar}@kth.se

Abstract—Cloud radio access network (CRAN) has been proposed as a potential energy saving architecture and a scalable solution to increase the capacity and performance of radio networks. The original CRAN decouples the digital unit (DU) from radio unit (RU) and centralizes the DUs. However, stringent delay and bandwidth constraints are incurred by fronthaul in CRAN, i.e. the network segment connecting RUs and DUs. In this study, we propose a modified CRAN architecture, namely hybrid cloud RAN (H-CRAN), where a DU’s functionalities can be virtualized and split at several conceivable points. Each split option results in two-level deployment of the processing functions, i.e., central cloud level and edge cloud level, connected by a transport layer called “midhaul”. We study the interplay of energy efficiency and midhaul bandwidth consumption when baseband functions are centralized at the edge cloud vs central cloud. We jointly minimize the power and midhaul bandwidth consumption in H-CRAN, while satisfying the network constraints. The addressed problem with the associated constraints are modeled as a mixed integer constraint optimization problem. Numerical results show the compromise between energy and bandwidth consumption, with the optimal placement of baseband processing functions in H-CRAN architecture.

I. INTRODUCTION

5G networks are envisioned to support 1000-fold more traffic and 10-fold lower latency. However, the associated system’s energy efficiency must be high [1]. Researchers have addressed the improvement of system’s energy efficiency, while considering several aspects [2]. In recent years, cloud radio access network (CRAN) has been proposed as a solution to reduce the energy consumption and cost [3], [4]. In CRAN, DU are decoupled from RUs and centralized at a central site (CS), called DU hotel. A DU hotel is evolved to a DU cloud if general purpose servers are employed in the DUs, hence their functions can be virtualized, which leads to virtualized CRAN. In this paper we refer to the DU cloud, central cloud. The centralization of DUs enables the sharings of not only computational resources, but also the infrastructure in central cloud. Our previous work has studied how to save power consumption in a virtualized CRAN architecture [5]. However, if all the baseband processing functions are centralized in the central cloud, bandwidth and transmission delay requirement becomes very challenging to satisfy. This high-bandwidth and low-latency link is referred to as “fronthaul” where I/Q samples generated by RUs must be transmitted to the central cloud. The 3ms delay constraint limits distance between a central cloud and RU to 20~40 km. Also, huge amount of bandwidth is required in fronthaul. For example, a single RU, with a 20 MHz carrier and 2*2 MIMO scheme, will generate 2.5 Gbps I/Q samples in downstream. In a densified radio

access network (RAN) envisioned for 5G, such a rigid delay and bandwidth requirement leaves dedicated fiber or (active) optical transport network (OTN) as the only viable solution, which can counteract the cost saving of CRAN.

To relax fronthaul constrains in CRAN, the functional split technique has been recently proposed to split the baseband processing chain at several conceivable points. This technique divides the cloud in two hierarchical layers: central cloud and edge cloud [6]–[9], hence, dual-site processing is incurred. Now, DU functions are split into two sites located in the cloud, the part toward RU is placed in the remote site (RS) at the edge cloud for partial baseband processing, whereas the other part is placed in the CS at the central cloud for the remaining processing. The transportation link between RS and CS is referred to as “midhaul” [10]. The motivation of dual-site processing is twofold: a) by conducting partial baseband processing at edge cloud, bandwidth requirement can be significantly relaxed for midhaul; b) by equipping RSs with general-purpose processors, to support computational and content-caching capabilities, traffic load can be terminated at edge cloud to relieve traffic load in core networks. In this study, we call CRAN with dual-site processing and function virtualization as hybrid cloud RAN (H-CRAN), where functions can be centralized either at the edge or at the central cloud. In the context of designing realizable fronthaul, several work have been conducted. In [11], the authors have studied the flexible RAN centralization techniques while considering several connectivity technologies, such as, free space optic, DSL, milli-meter wave, microwave, fiber (with different access), etc. Authors of [12] have considered a graph based framework to reduce the cost of the fronthaul via optimally placing the DU in the network. Their solution to this problem adopted a genetic algorithm approach. The authors of [13] have analyzed several design requirements for the fronthaul, considering different aspects: providing traffic load, supporting different logical topologies, guaranteeing low latency.

Deciding the optimal functional split is still an open problem, which profoundly depends on the objective to be optimized. Intuitively, if more functions are centralized at CS, as in CRAN, higher power consumption saving can be achieved, whereas, midhaul bandwidth consumption will increase. On the contrary, placing more processing functions at the edge cloud may lead to a higher power consumption but lower midhaul bandwidth consumption. Hence, a trade-off between placing the functions at the central cloud (CS) or at the edge cloud (RSs) should be investigated

In this study, our contributions are summarized as follows.

First, we present the architecture of H-CRAN with midhaul and dual-site processing. In H-CRAN, DUs are deployed at both RSs and CS for baseband processing, whereas, each RS is connected to a group of RUs, see Sec. II. Second, we model the baseband processing chain as a sequence of functions that can be split between any two functions, and these functions can be deployed either in DUs at RSs or in DUs at CS. Third, we develop a multi-objective mixed integer program using constraint programming to decide the optimal functional splits, while aiming to jointly minimize the system's power consumption and the bandwidth consumption in midhaul.

The organization of the remaining sections is as follows. Section II addresses the architecture of H-CRAN, and presents the functional split model. Section III presents the problem formulation. Finally, Sec. IV presents the simulation results.

II. NETWORK ARCHITECTURE

A. Proposed H-CRAN Architecture

We present a hybrid architecture that employs dual-site processing in virtualized CRAN, where DUs are deployed at both CS in central cloud and RS in edge cloud, so that baseband processing can be flexibly provisioned by a chain of virtualized functions for a RU or even for a user equipment (UE), while traffic are transported through the network. We call this architecture as hybrid cloud RAN (H-CRAN), as shown in Fig. 1.

H-CRAN is a three-layer architecture, which consists of cell layer (the coverage of a RU is referred to as a "cell"), RS layer at edge cloud, and CS layer at central cloud. Cell layer consists of cells that are being densified, each serving several UEs. A group of cells are connected to a RS as an aggregation point. The fronthaul between a cell and a RS can be implemented using a short fiber (as in conventional settings), or wireless links, e.g. mmWave links [14] or free-space optical links [15]¹. The RSs can be connected to CS via midhaul using various technologies, from expensive dark fiber or TON solutions, to cost-efficient PON families or other Ethernet-based technologies. In this work, we study the system's power consumption under the constraint of bandwidth capacity per midhaul link (for a RS), which ranges from 1 Gbps to 20 Gbps. The midhaul technology considered in this study is time-wavelength division multiplexing PON (TWDM-PON) [16], and each midhaul link is a wavelength channel, which needs an optical network unit (ONU) at RS and a Line-Card (LC) at CS as transceivers.

Edge cloud layer and central cloud layer are deployed with DUs, which are containers for virtualized functions, because their computational resources can be virtualized and shared by any connected RUs (if implemented in general purpose servers). For example, in upstream, traffic from cells can be partially processed at edge cloud so that bandwidth requirement can be relaxed for midhaul, then remaining processing will be conducted at central cloud. However, RS is usually less energy-efficient than CS, because the number of DUs, associated with RUs, at the CS is larger than that in

each RS. Hence, sharing infrastructure equipment results in higher energy saving at CS. Our trade-off becomes whether to centralize functions at RS in edge cloud (to save midhaul bandwidth), or to centralize more functions at CS in central cloud layer (to save power).

B. Reference Architectures

We define two extreme cases with no functional split as reference architectures for the performance analysis. (1) Edge-CRAN where all the baseband functions are centralized at the edge cloud and the connection to the central cloud is provided by a backhaul. In this case, DUs are stacked at a nearby cabinet within the RS. DUs, and infrastructure in cabinet, are dedicatedly serving RU of the base station (BS), and cannot be shared by other BSs, which leads to low energy efficiency. But since baseband processing is fully conducted at RS, the conventional backhaul requires a small amount of bandwidth as perceived by UEs. (2) Central-CRAN where all the baseband functions are centralized at the central cloud. In this case, sharing infrastructure for the required baseband processing results in reducing the power consumption [17]. However fronthaul must be considerably prolonged by using dedicated fibers, optical transport network, or passive optical network (PON) [10] because DUs are located at CS. Since no baseband processing is conducted at RS, huge amount of bandwidth is needed in the prolonged fronthaul.

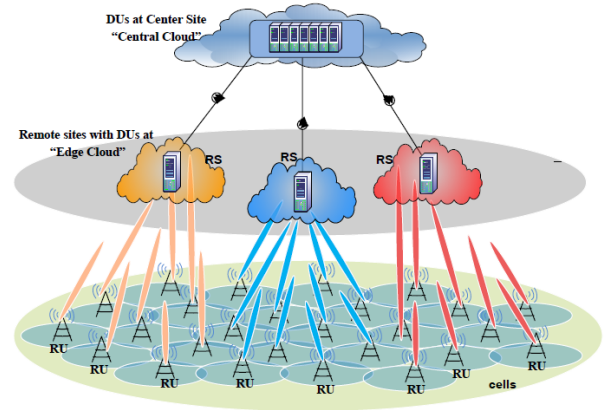


Fig. 1. Hybrid virtualized RAN architecture.

C. Functions Split Model

To study the function distribution against function centralization, we model the functional split of baseband processing chain for a cell, as shown in Fig. 2. First, baseband processing for a cell is modeled as a chain of functions, which includes m Cell-Processing (CP) functions and n User-Processing (UP) functions². CPs are a sequence of functions in physical layer that are dedicated for processing signals from a cell, when signals of UEs are multiplexed. For example, in upstream, CPs includes serial-to-parallel conversion (or common public radio interface (CPRI) encoding), removing cyclic prefix, fast fourier transform, and finally resource demapping etc. The per-cell processing will be terminated at CP_m , and now signals

¹Our architecture is not hard-wired to any specific fronthaul technology within RS, as they are not the focus of our study.

²In this study, $m = 3$ and $n = 3$, as considered in [7]–[9].

from a cell will be de-multiplexed as multiple signal streams, each belonging to a UE. Then, UPs is a sequence of functions that will continue to process the signal streams on a per-UE basis³, including equalization, inverse discrete Fourier transform, quadrature amplitude modulation, antenna demapping, multi-antenna processing, forward error correction, turbo decoding, and other Layer2 and Layer3 functions.

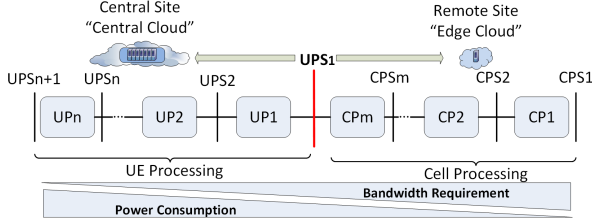


Fig. 2. Function split model.

As shown in Fig. 2, functional split can happen before CP_1 , after UP_n , or between any two functions. Note that CP split 1 (CPS_1) is the initial attempt to implement CRAN, which is based on full baseband centralization. UP split $n+1$ (UPS_{n+1}) is implemented by distributed RAN (DRAN), characterized by a fully distributed deployment.

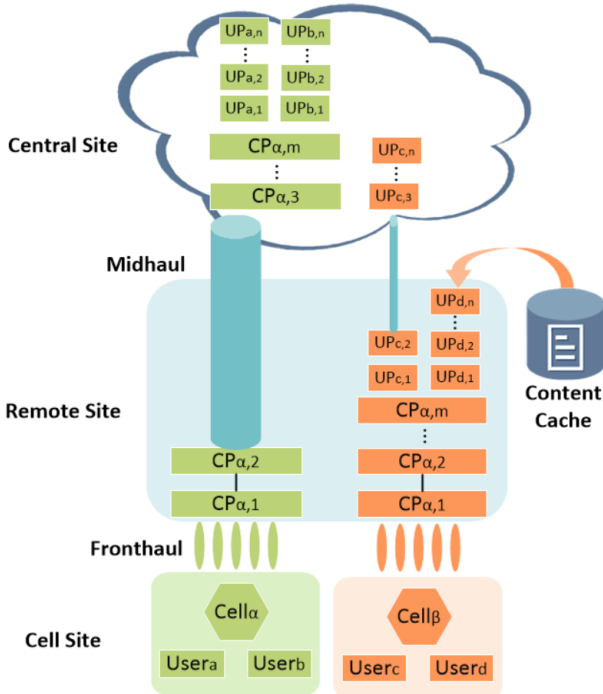


Fig. 3. H-CRAN including function split architecture.

In Fig. 3, we further illustrate our assumptions on the implementation of functional split in H-CRAN. First, the processing chain can be cut at most once and thus split into two parts, and the lower part will be deployed at RS and the upper part will be placed at CS. Second, each functional split will incur a bandwidth requirement, which can be calculated by formulas provided in [7], [18]. Third, once a functional split is decided, we can decide the number of CPs and UPs

³Note that we allocate fixed number of resource blocks (RBs) for each UE such that at full load cell the assigned 20 MHz per cell is enough to serve all users per cell.

that need to be deployed at RS and CS, respectively. For example, in left part of Fig. 3, for $cell_\alpha$, split happens in CP sequence, so large amount of bandwidth is required to transmit the partially-processed signals, and all CPs below the split point must be placed at the same DU in RS, and all CPs above the split point and all UPs must be placed at the same DU in CS. However, for $cell_\beta$, split happens in UP sequence, as shown in the right part of Fig. 3, cell's signals have been processed to a larger extent, so less bandwidth will be required in midhaul. In addition, since cell's signals has been de-multiplexed, different functional splits can be conducted and customized for different UE. UE based service differentiated cloud resource provisioning is left as a future work in this paper.

III. FUNCTIONAL SPLIT OPTIMIZATION PROBLEM

In this section, we aim at jointly minimizing the system's power consumption, and bandwidth consumption of midhaul for H-CRAN. Therefore, we model the interplay of power and bandwidth consumptions as a functional split optimization problem, which is formulated as a mixed integer program using constraint programming [19].

A. Given

- **Topology:** one CS connected to multiple RSs. Each RS is connected to an exclusive set of cells, and each cell exclusively covers a set of UEs
- I_x : set of UEs. When $x = 0$, it means all UEs in H-CRAN, $x = c$ means set of UEs in cell c , $x = r$ means set of all UEs in cells belonging to RS r .
- C_x : set of cells. When $x = 0$, it means all cells in H-CRAN, $x = r$ means set of cells belonging to RS r .
- D_x : set of DUs. When $x = 0$, it means all DUs in H-CRAN, $x = CS$ means set of DUs in the CS, $x = r$ means set of DUs in RS r .
- R : set of all RSs.
- W : set of wavelengths.
- F_x : set of functional split options, where x represents UP split or CP split.
- $H_x^y(\cdot)$: pre-calculated mapping from a split option $x = (UP/CP)$ to the number of (UP/CP) functions at site $y = (CS/RS)$ (x represents UP or CP, and y represents CS or RS). For example, if the UP sequence of a UE is split at the middle point (indexed by m), $H_{UP}^{RS}(m) = H_{UP}^{CS}(m)$ equals to half of UPs.
- $J_i(\cdot)$: pre-calculated mapping from UP split of UE i to the required midhaul bandwidth, which is proportional to the number of RBs allocated to the i 's UE [7], [18].
- $G_c(\cdot)$: pre-calculated mapping from CP split of cell c to the required midhaul bandwidth, which is proportional to the number of antennas and carrier bandwidth [7], [18].
- K : bandwidth capacity of a wavelength.
- L_x^y : capacity of a DU located at " y " side, in terms of the number of x functions that can be accommodated by this DU (x represents CP or UP, and y represents CS or RS). For example, L_{CP}^{CS} represents the maximum number of CPs that can be accommodated by a DU at CS.

- We also define an indicator function, $\mathbb{I}(\cdot)$, to test whether a constraint is satisfied. If the constraint, in the argument of the function, is satisfied, the indicator functions has value 1, otherwise, 0. The mathematical definition is expressed as follows,

$$\mathbb{I}(a = b) = \begin{cases} 1; & \text{if } a = b, \\ 0; & \text{if } a \neq b. \end{cases} \quad (1)$$

B. Integer Variables

- $p_i \in [0, |F_{UP}|]$: UP split of UE i . Note that if UP of UE i is not split, then $p_i = |F_{UP}|$, otherwise, $p_i \in [0, |F_{UP}|)$.
- $q_c \in [0, |F_{CP}|]$: CP split of cell c . If CP of cell c is not split, then $q_c = |F_{CP}|$, otherwise, $q_c \in [0, |F_{CP}|)$.
- $m_i \in D_r$: DU hosting UPs of UE i at RS r . Note that since the association between i and r is fixed, UE i can only choose a DU from a given set.
- $n_i \in D_{CS}$: DU hosting UPs of UE i at CS.
- $x_c \in D_r$: DU hosting CPs of cell c at RS r .
- $y_c \in D_{CS}$: DU hosting CPs of cell c at CS.
- w_r : wavelength used by RS r .
- e_r : number of active DUs at RS r .
- l : number of active DUs at CS.
- g : number of active wavelengths in the midhaul.

C. Objective

Our objective is to minimize a linearly weighted sum of the system's normalized power consumption plus the normalized total bandwidth consumption in midhaul, to study the trade-off of the individual/combined impact on the system. The multi-objective function is expressed as,

$$\min W_P \cdot \frac{p_T}{P_N} + W_B \cdot \frac{b_{MH}}{B_N} \quad (2)$$

where W_P and W_B are the weighting factor of the power consumption and the midhaul bandwidth consumption, respectively. We choose $W_P = 1 - W_B$, i.e., to highlight the complementary impact of the associated metrics. The parameters P_N and B_N are the normalization factor of each the power and bandwidth consumptions, respectively⁴. The notations p_T and b_{MH} denotes the total power consumption and midhaul bandwidth consumption, respectively. The total power consumption is expressed as,

$$p_T = g \cdot P_{LC} + (P_{CS} + l \cdot P_{CS}^{DU}) \cdot u_l + \sum_{r \in R} (P_{ONU} + P_{RS} \cdot u_{e_r} + e_r \cdot P_{RS}^{DU}) \quad (3)$$

where P_{LC} denotes the power consumption of a LC, and P_{ONU} is the power consumption of an ONU. $u_l = 1$ if $l \neq 0$ and $u_l = 0$ if $l = 0$, similarly, $u_{e_r} = 1$ if $e_r \neq 0$ and $u_{e_r} = 0$ if $e_r = 0$. P_{CS} and P_{RS} are the power consumption for housing DUs at CS and RS, respectively, where P_{CS}^{DU} and P_{RS}^{DU} are the power consumption of a DU in CS and RS, respectively. The power consumption of RUs and fronthaul part within RS are assumed to be the same for all three architectures, so their

⁴The normalization factors, P_N and B_N are the maximum consumed power and maximum consumed midhaul bandwidth, respectively.

values are constant. The midhaul bandwidth consumption is obtained, as follows,

$$b_{MH} = \sum_{w \in W} \sum_{r \in R} \mathbb{I}(w_r = w) \cdot \sum_{c \in C_r} \left(G_c(q_c) + \sum_{i \in I_c} J_i(p_i) \right), \quad (4)$$

by summing over all consumption induced by all remote sites.

D. Constraints

In this sub-section, we explain the constrains of the problem.

$$\mathbb{I}(p_i < |F_{UP}|) + \mathbb{I}(q_c < |F_{CP}|) = 1, \quad \forall c \in C_0, \forall i \in I_c \quad (5)$$

This ensures that function split can occur only once, either at CP or at UP.

$$(p_i < |F_{UP}|) \implies (m_i = x_c), \quad \forall c \in C_0, \forall i \in I_c \quad (6)$$

This ensures that if UP of UE i is split, then lower part UPs must be placed in the same DU with their CP at RS, as shown in Fig. 2, because otherwise complex inter-DU communication will be incurred.

$$(q_c < |F_{CP}|) \implies (n_i = y_c), \quad \forall c \in C_0, \forall i \in I_c \quad (7)$$

This ensures that if CP of cell c is split, upper part of CPs must be placed in the same DU with all UPs (of all UEs in cell c) at CS, as shown in Fig. 2.

$$\sum_{c \in C_r} H_{CP}^{RS}(q_c) \cdot \mathbb{I}(x_c = d) \leq L_{CP}^{RS}, \quad \forall r \in R, \forall d \in D_r \quad (8)$$

This ensures that the total number of CPs that are accommodated by a DU d at RS r cannot exceed this RS-DU's CP capacity. Note that L_{CP}^{RS} is less than L_{CP}^{CS} .

$$\sum_{c \in C_0} H_{CP}^{CS}(q_c) \cdot \mathbb{I}(y_c = d) \leq L_{CP}^{CS}, \quad \forall d \in D_{CS} \quad (9)$$

This ensures that the number of CPs that are accommodated by a DU d in CS cannot exceed this CS-DU's CP capacity.

$$\sum_{i \in I_r} H_{UP}^{RS}(p_i) \cdot \mathbb{I}(m_i = d) \leq L_{UP}^{RS}, \quad \forall r \in R, \forall d \in D_r \quad (10)$$

This ensures that the number of UPs that are accommodated by a DU d at RS r cannot exceed this RS-DU's UP capacity.

$$\sum_{i \in I_0} H_{UP}^{CS}(p_i) \cdot \mathbb{I}(n_i = d) \leq L_{UP}^{CS}, \quad \forall d \in D_{CS} \quad (11)$$

This ensures that the number of UPs that are accommodated by a DU d at CS cannot exceed this CS-DU's UP capacity.

$$\sum_{r \in R} \mathbb{I}(w_r = w) \cdot \sum_{c \in C_r} \left(G_c(q_c) + \sum_{i \in I_c} J_i(p_i) \right) \leq K, \quad \forall w \in W \quad (12)$$

This ensures that the total occupied bandwidth in a wavelength cannot exceed the wavelength's capacity. The occupied bandwidth in wavelength w is the sum of the bandwidth consumptions of all RSs that are using w . The bandwidth consumption of RS r is the sum of bandwidth consumptions of all cells belong to it ($c \in C_r$). The bandwidth consumption of cell c is either the bandwidth requirement incurred by CP split ($G_c(q_c)$), or the bandwidth requirement incurred by UP splits.

If it is the later, the bandwidth requirement of cell c is the sum of bandwidth requirements of all UEs ($J_i(p_i)$) belonging to c ($i \in I_c$).

$$e_r = \sum_{c \in C_r} \delta(x_c), \quad \forall r \in R \quad (13)$$

where $\delta(x_c) = 0$, if x_c is not active (no CP of cell c is placed at RS r), $\delta(x_c) = 1$, if x_c is active (at least one CP of c is placed at RS r). This counts the number of active DUs at RS r . The algorithmic model of this constraint is expressed as,

$$e_r = \text{countDiff}(\{x_c\}_{c \in C_r}) - \prod_{c \in C_r} (q_c = 0), \quad \forall r \in R \quad (14)$$

The `countDiff` constraint is a special operator that counts the number of distinct values taken by variables in array $\{x_c\}_{c \in C_r}$. When there exist active DUs in RS r , the number of active DUs is equal to the number of distinct values taken by $\{x_c\}_{c \in C_r}$. But in special case when there is no active DU in RS r , all cells of the RS r choose the rightmost split in Fig. 2, i.e. $q_c = 0, \forall c \in C_r$, so $\prod_{c \in C_r} (q_c = 0)$ equals to 1. To ensure that $e_r = 0$ in this case, all x_c are forced to choose the same value, i.e., `countDiff` ($\{x_c\}_{c \in C_r}$) equals to 1.

$$l = \text{countDiff}(\{n_i\}_{i \in I_0}) - \prod_{i \in I_0} (p_i = |F_{UP}| - 1) \quad (15)$$

This constraint counts the number of active DUs at CS. Explanation for (14) can apply to this constraint also, except that the special case that there is no active DU at CS happens when all UEs choose the leftmost split in Fig. 2, i.e. $p_i = |F_{UP}| - 1, \forall i \in I_0$.

$$g = \text{countDiff}(\{w_r\}_{r \in R}) \quad (16)$$

Equation (16) counts the number of midhaul's active wavelength.

IV. SIMULATION RESULTS

In this section, we evaluate the system performance of the multi-objective function split optimization and analyze the interplay between power and midhaul bandwidth consumption as more or less functions are centralized at the central cloud vs. edge cloud. As a reference, we consider two cases to compare the optimized network function placement: (1) Edge-CRAN: all the functions are placed at the edge cloud, i.e., at the remote site; and (2) Central-CRAN: all the functions are centralized at the central cloud. We utilize the IBM constraint programming tool to obtain the global optimal solutions for the function split. The system parameters used in the simulation are expressed in Table I.

In Figure 4, we study how function placement changes as more midhaul bandwidth is available, if we emphasize on minimizing the power consumption ($W_P = 0.99$ in (2)). We plot the percentages of functions that are placed at RS and CS,

⁵We consider a fully loaded scenario where each RU assign all physical resource blocks to 5 UEs.

⁶The configuration parameters are used to calculate the bandwidth consumption of a RU mapped from a functional split, using formula provided in [7], [18]

⁷Contains the power consumption of the connection link from RS to RUs and power radiated from RUs.

TABLE I
SIMULATION PARAMETERS.

Parameter Name	Value
Topology	1 CS, 6 RSs, 7 RUs per RS, and 5 UEs per RU ⁵ .
Configuration of RU ⁶	20 MHz, 2*2 MIMO, 64 QAM (DL)
Number of CPs & UPs per RU	3 & 3
Capacity of DU at RS	3 CPs and 15 UPs (one RU)
Capacity of DU at CS	27 CPs and 135 UPs (nine RUs)
Power of DU at RS	50 W
Power of infrastructure RS	250 W
Power of DU at CS	100 W
Power of infrastructure CS	500 W
LC Power+ ONU Power,	20 W + 5 W
Radio Access + Fronthaul link Power Consumption ⁷ .	20 W

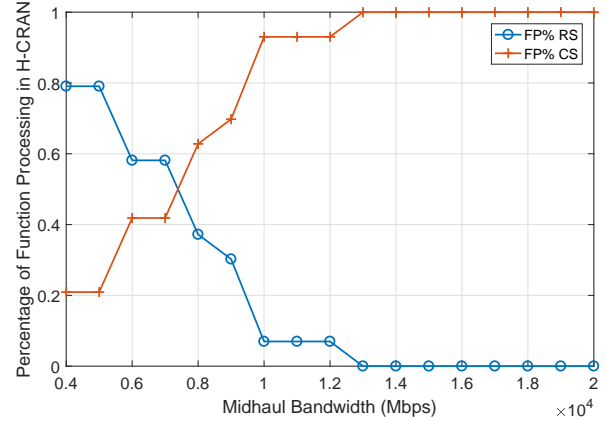


Fig. 4. Illustration of percentage of functions placed at RSs/CS, versus bandwidth capacity per midhaul link (under $W_P = 0.99$).

respectively, versus the bandwidth capacity per midhaul link (wavelength). As bandwidth capacity increases, more functions are centralized at central cloud CS, and less functions are placed at edge cloud RSs. Because when power consumption is minimized, and midhaul bandwidth is abundant, it is always preferable to centralize functions at an energy-efficient. In addition, there exists an upper bound for effective usage of wavelength channel, because when bandwidth is larger than 13 Gbps in Figure 4, all functions have already been centralized, and thus no more power can be saved. However, in practice, not all network operators can have abundant midhaul bandwidth, so the planning of function placement is mutually decided by the available midhaul solutions (with different bandwidth capacity per link).

Figure 5 compares the power consumption of the proposed H-CRAN with the power consumptions of central-CRAN and edge-CRAN, versus the bandwidth capacity per midhaul link, when power consumption is more valued ($W_P = 0.99$). Confirming the result in Figure 4, as more bandwidth is available in midhaul, less power will be consumed by hybrid-CRAN. When bandwidth capacity is small, hybrid-H-CRAN achieves the same power consumption as edge-CRAN, because all functions are centralized at the edge cloud layer divided between RSs. When bandwidth capacity is large, H-CRAN achieves the same power consumption of central-CRAN, because all functions are centralized at CS.

Figure 6 presents a comparison between different normalized system performances (i.e., the objective value, power consumption, midhaul bandwidth consumption, and percentages

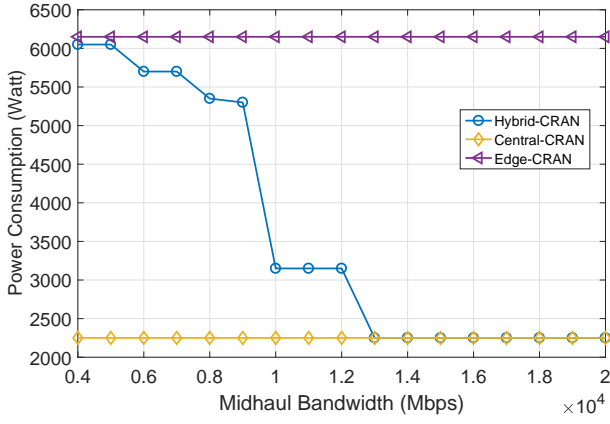


Fig. 5. Power consumption of hybrid-CRAN, central-CRAN, and edge CRAN

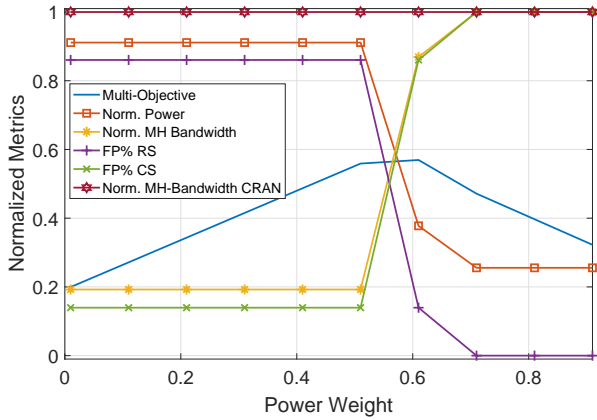


Fig. 6. System performances versus power and midhaul weights.

of functions placed at RSs and at CS, respectively), for different W_P values. When W_P increases, the power consumption decreases, and the bandwidth consumption increases, which indicates the trade-off between power and bandwidth consumptions. Moreover, the curves of power consumption and bandwidth consumption transpose when W_P ranges from 0.5 to 0.6, which indicates that their interplay is a drastic process. So, if we want to jointly minimize power and bandwidth consumptions, their weights in objective function must be carefully chosen, e.g. $W_P = 0.55$ where the two curves cross. It is also observed that as the bandwidth's weight decreases, power's weight increases, the bandwidth consumption of H-CRAN converges to that of central-CRAN.

V. CONCLUSION

In this paper, we consider a cloud radio access network architecture with dual-site processing, where baseband processing chain can be split, and virtualized functions can be provisioned at both RS and CS. We call this architecture as hybrid-cloud RAN (H-CRAN). We model the functional split in H-CRAN, as a sequence of cell processing functions and user processing functions. We propose an optimization framework to jointly minimize the power consumption and transport bandwidth consumption in H-CRAN, using a mixed integer program based on constraint programming. Numerical results showed that when power consumption is more valued, as more transport bandwidth is available, more functions are

placed at the central cloud CS to save power. Also, the interplay of power and bandwidth consumptions is drastic, and there exists a balanced point for joint minimization of them. Since the proposed optimization framework can decide the functional split and DU assignment for each user, future works can consider a dynamic scenario where users arrive/depart at/from the network, with different service requirements, e.g. bandwidth, latency etc., and functional split can consider different objectives, such as minimizing delay, or maximizing throughput for users.

REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [2] A. Alabbasi and B. Shihada, "Optimal cross-layer design for energy efficient d2d sharing systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 839–855, Feb 2017.
- [3] A. Checko, H. L. Christiansen, Y. Yan, L. Scolaro, 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.
- [4] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-Dense Networks: A Survey," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–1, 2016.
- [5] X. Wang, S. Thota, M. Tornatore, H. S. Chung, H. H. Lee, S. Park, and B. Mukherjee, "Energy-Efficient Virtual Base Station Formation in Optical-Access-Enabled Cloud-RAN," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 5, pp. 1130–1139, May 2016.
- [6] "Environmental Engineering (EE); Assessment of mobile network energy efficiency," ETSI ES 203 228 V1.1.6 (2016), ETSI, 2016.
- [7] "Functional splits and use cases for small cell virtualization." Release, Small Cell Forum, Jan. 2016.
- [8] "IEEE P1914.1 Meeting Materials. [online]:<http://sites.ieee.org/sagroups-1914/>," IEEE P1914.1 TF meeting materials, IEEE, August, 2016.
- [9] U. Ditsch, 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, June 2013.
- [10] T. Pfeiffer, "Next generation mobile fronthaul and midhaul architectures [invited]," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 7, no. 11, pp. B38–B45, November 2015.
- [11] A. Maeder, M. Lalam, A. D. Domenico, E. Pateromichelakis, D. Wbben, J. Bartelt, R. Fritzsche, and P. Rost, "Towards a flexible functional split for cloud-ran networks," in *Networks and Communications (EuCNC), 2014 European Conference on*, June 2014, pp. 1–5.
- [12] 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 *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 1958–1963.
- [13] J. Liu, S. Xu, S. Zhou, and Z. Niu, "Redesigning fronthaul for next-generation networks: beyond baseband samples and point-to-point links," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 90–97, October 2015.
- [14] M. Artuso, A. Marcano, and H. Christiansen, "Cloudification of mmwave-based and packet-based fronthaul for future heterogeneous mobile networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 76–82, October 2015.
- [15] M. A. Khalighi and M. Uysal, "Survey on Free Space Optical Communication: A Communication Theory Perspective," *IEEE Communications Surveys Tutorials*, vol. 16, no. 4, pp. 2231–2258, Fourthquarter 2014.
- [16] "40-Gigabit-capable passive optical networks (NG-PON2)," ITU-T G.989 series of Recommendations., ITU-T, March 2013.
- [17] J. Liu, S. Zhou, J. Gong, Z. Niu, and S. Xu, "Statistical multiplexing gain analysis of heterogeneous virtual base station pools in cloud radio access networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 8, pp. 5681–5694, Aug 2016.
- [18] X. Wang, "UC Davis Technical Report. [online]:<http://networks.cs.ucdavis.edu/xinbo/appendix-a-techno-economic-study-to-design-low-cost-edge-cloud-radio-access-network.pdf>," 2016.
- [19] "IBM ILOG CPLEX optimization studio: OPL language users manual." Version 12 Release 6, IBM, 2015.