Optimal Placement of Virtualized BBU Processing in Hybrid Cloud-Fog RAN over TWDM-PON

Rodrigo Izidoro Tinini*, Larissa C. M. Reis†, Daniel Macêdo Batista*, Gustavo Bittencourt Figueiredo†, Massimo Tornatore‡ and Biswanath Mukherjee§
*University of São Paulo, São Paulo, Brazil, Email: {rtinini, batista}@ime.usp.br
†Federal University of Bahia, Bahia - Brazil, Email: gustavo@dcc.ufba.br
‡Politecnico di Milano, Milan - Italy, Email: massimo.tornatore@polimi.it
§University of California, Davis - USA, Email: bmukherjee@ucdavis.edu

***Abstract*—In the context of future Cloud Radio Access Networks (CRAN), optical networks will play an important role to provide the required transport capacity between cell-sites and processing pools, especially for future 5G scenarios. For instance, using CPRI fronthaul technologies a single antenna element can generate data up to 24.3Gbps even with current configurations of radio transmissions, and it is expected to generate up to Tbps with the advance of technology. So, the transport segment of a 5G network needs to be accurately planned to accommodate all the generated traffic. In this work, we propose the use of a Passive Optical Network (PON) jointly with the emergent paradigms of Fog Computing and Network Function Virtualization (NFV) to energy-efficiently support the high traffic transported in emergent mobile networks in an hybrid architecture called Cloud/Fog RAN (CF-RAN) that allows local and remote baseband processing. We introduce an Integer Linear Programming (ILP) model to schedule the processing of CPRI demands among the processing nodes of the network and turn on or off processing functions on demand. Our approach is able to accommodate demands on the nodes of the network in the most energy efficient way. We compare our results with CRAN and distributed architectures (DRAN) and show that an energy efficient planning can achieve considerable gains in power consumption.**

I. Introduction

To increase network efficiency while limiting CAPEX and OPEX, mobile operators are increasingly adopting Cloud-Computing solutions to deliver their services. Cloud Radio Access Networks (CRAN), in particular, is a recent cloud-based paradigm already implemented in some LTE networks, and it is considered one promising option for emerging 5G networks. CRAN is capable of reducing the overall power consumption of the network by moving the BaseBand Units (BBU) from the cell sites to a BBU pool on the cloud, while leaving low-energy Remote Radio-Heads (RRH) at the cell sites [1]. In this architecture, the data communication between RRHs and the BBU pool is done through a transport network called the fronthaul.

Although CRAN can reduce power consumption, the Common Public Radio Interface (CPRI) protocol used in the fronthaul imposes high traffic demands and strict latency limits. Thus, optical networks are commonly proposed to support the traffic on the fronthaul. The CPRI line rates vary from 614.4Mbps to 24.3Gbps depending, e.g., on the Multiple Input - Multiple Output (MIMO) configuration, and on the type of aggregation of the supported antennas. In the latter case, each individual RRH can transmit a CPRI flow per antenna or one aggregated CPRI from different flows coming from different antennas [2]. Thus, the optical network infrastructure must be carefully planned to support such large traffic demands while achieving energy efficiency and optimally using the optical resources.

Passive Optical Networks (PON) are a good choice for fronthaul networks due to their low cost, but their transmission capacity is still a limiting factor. To address this limitation, Time and Wavelength-Division Multiplexing PON (TWDM-PON) are a promising solution to provide high bandwidth, low latency and cost efficiency. In TWDM-PON, Optical Network Units (ONU) are placed at the transmitter side to assign wavelengths to RRHs and an Optical Line Terminal (OLT) is placed at the receiver side to demultiplex the wavelengths received.

Several recent works investigated the operation of optical networks in support of CRAN. In [3], the authors studied the energy efficiency of BBU hotelling implemented over a Wavelength Division Multiplexing (WDM) optical network and the energy savings of BBU hotelling was evaluated. In [4] the authors considered the case of CRAN over PON to support coordinated multi-point techniques showing that the use of the PON achieved savings on the signaling time between cells. In [5], the use of PON for CRAN was able to reduce the number of handovers and improve the throughput by means of virtualization at the BBU pool and at the cells. To increase spectral and energy efficiency, authors in [6] propose an architecture called Heterogeneous Cloud Radio Access Network (H-CRAN) that combines the benefits of CRAN and the processing capabilities of conventional base stations. In this architecture the BBU pool is used for baseband processing whereas the local processing facility is used for user-oriented services.

In this paper, we extend the H-CRAN in an architecture called Cloud-Fog Radio Access Network (CF-RAN) by taking advantage of the emergence of the Fog Computing paradigm over a TWDM-PON. In CF-RAN, compared to H-CRAN, a wider range of services provided by the cloud can be placed closer to the user in dedicated servers. In addition, the CF-RAN adopts the concept of Network Function Virtualization

(NFV) to process the baseband signals from the RRHs. This processing is done in Virtual Machines (VM) that implement BBU processing functions instantiated on demand via NFV.

Thus, the processing nodes closer to the user, or fog nodes, of the CF-RAN can be activated or deactivated on demand according to traffic and processing demands to increase the energy efficiency and to save fronthaul utilization. Finally, the TWDM-PON is used to support the data transmission to the processing nodes.

However, enabling virtualized local-processing functions at the fog nodes via NFV may also increase OPEX, since cooling and processing functions at local sites need to be activated accordingly. As the local-processing functions at fog nodes are activated when the capacity of the fronthaul is exhausted by the data generated at the RRHs, this delineates a tradeoff between the fronthaul capacity and the total network power consumption. So, the optimal placement of virtualized local-processing function plays an important role to identify the most energy-efficient point of operation.

So, in this paper we propose a TWDM-PON architecture to support both the fronthaul network of the CF-RAN and the interconnection of cell-sites to the fog nodes.

For the operation of the CF-RAN architecture, we also propose a ILP model that decides the number of fog nodes necessary to accommodate all traffic demands and where to process them considering TWDM-PON capacity and the best overall network energy performance. To the best of our knowledge, there is no work that addresses the use of a TWDM-PON for H-CRAN and how to optimally place the processing functions in the network to achieve the best tradeoff between network capacity and power consumption.

The paper is organized as follows: Section II presents the proposed CF-RAN system architecture. Section III presents the proposed ILP model and the power consumption model used for the CF-RAN planning. Section IV presents the numerical results. Section V concludes the paper and discusses future research directions.

II. System architecture

In this section we present our proposed CF-RAN architecture. First, we will discuss the details of the interconnection network used in the fronthaul to support the traffic connecting each RRH to its BBU. Then we will present the architecture of the fog node and the operation of fog and cloud nodes.

A. CPRI Data Transmission

The transmission of the data aggregated at the RRHs towards the BBU is done under the CPRI protocol. CPRI establishes rules for the transmission of digitized radio signals between the radio equipment (RE), which is the RRH, and the radio equipment controller (REC), which is the BBU. In CPRI, each RRH can aggregate the traffic according to the MIMO configuration used, e.g., one RRH can serve a cell-site of 3 sectors using, e.g., one 2x2 MIMO group per each sector, which comprises radio receivers and transmitter elements. A RRH can serve more cell-site sectors by increasing its MIMO configuration, which will in consequence increase the data generated by the RRH and the fronthaul demand.

Therefore, a cell-site with multiple RRHs can generate hundreds of Gbps up to Tbps. To better use the transport network, CPRI also allows the aggregation of multiple RRHs into one single CPRI line using a tree topology. In our proposed architecture, the aggregation of radio signals into one CPRI line will allow the TWDM-PON to encapsulate multiple RRH signals into one single wavelength and, thus, save wavelengths of the optical links to be used by other transmissions.

B. The Fronthaul Interconnection Network

In a H-CRAN [6], all the baseband processing is done at the BBU pool located in the cloud, while the RRH local processing is used for user-oriented services. Moreover, the signaling used for operational purposes, like the implementation of Multipoint Coordinated (CoMP) techniques and broadcast of control messages, is processed in the cloud and delivered to the cells through the S1 and X2 [6] interfaces.

In our proposed architecture, shown in Figure 1, the interconnection of the cell-sites to the virtualized BBU pool is done by a point-to-multipoint TWDM-PON fronthaul. It connects the cell-sites to the cloud as well as to local fog nodes equipped with the same processing functions as the cloud. We assume that the signaling exchanged through the S1 and X2 interfaces are also transmitted through the same optical links used to transmit the CPRI traffic. Hence, all the processing related to the signaling of the S1 and X2 can be done either at the cloud or in the fog nodes.

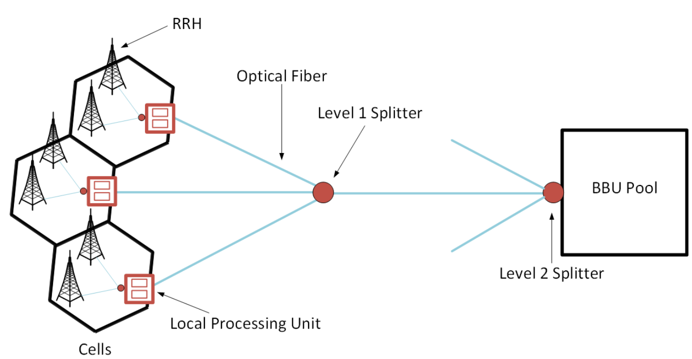

Fig. 1. Proposed CF-RAN architecture

Each RRH from a cell-site can be connected to an ONU or all RRHs from a cell-site can be all connected to a single ONU for all of them. In this second case, the ONU aggregates the traffic of the entire cell and transmits it through a single distribution fiber. Each ONU is equipped with a tunable laser responsible to assign the wavelengths to the transmissions, i.e., each ONU can tune its transmission to any available wavelength in its optical link.

The topology has three levels of multiplexing. In the first level of multiplexing, the optical signal transmitted by different RRHs is multiplexed towards the BBU pool located at the

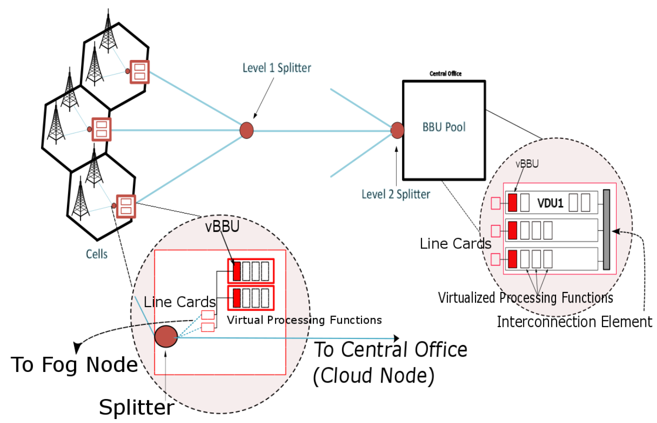

Fig. 2. Details of the processing nodes

cloud or towards the local processing functions. In the second level, a Level 1 optical splitter is used to multiplex the traffic transmitted in several distribution fibers into one feeder fiber. In the third level, close to the BBU pool, several feeder fibers are then multiplexed into a single Line Card (LC) that receives the optical signal for a particular wavelength.

The receiver node (located in a central office) is equipped with an OLT that demultiplexes and switches the traffic received on each wavelength to its corresponding Virtual Digital Unit (VDU), that comprises the virtualized environment for the baseband processing.

C. Operation of the Fog and Cloud Processing Nodes

Figure 2 shows the architecture of the Fog and Cloud nodes used to provide virtualized processing functions of the CPRI traffic or control messages.

Both fog and cloud nodes are equipped with a number of LCs to receive optical signals carrying the traffic to be processed. Each LC transmits the traffic to a specific VDU where it is received and processed.

Aligned to the NFV paradigm, the VDUs implement a set of Virtualized Processing Functions (VPF) that provide control functions or other additional network services. Among these functions is the Virtualized BaseBand Unit (vBBU), which is responsible for the baseband processing of the CPRI traffic. Since the functions are virtualized, they can be enabled or disabled dynamically according to network state.

Firstly, only the VPF of the cloud is used, so all the data aggregated at the cell sites is sent to the BBU pool in the cloud over the available wavelengths in the fronthaul. Only when the capacity of the fronthaul is exhausted, the local processing functions in the fog nodes are activated by the NFV capability on demand and accordingly to the availability of the wavelengths used for the transmission of data to the local BBU pool. After the wavelengths are assigned to the ONUs, they are transmitted to the local BBU pool via individual fibers, without the multiplexing of signals in a splitter. The fog nodes are activated one by one as the capacity of each fog node is exhausted.

D. Wavelength Assignment Schemes

To efficiently exploit the TWDM-PON capacity, we consider two policies of wavelength assignment. The first wavelength assignment scheme takes advantage on the RRH tree architecture allowed by CPRI, aggregating the data of all RRHs of a cell in one ONU/CPRI line rate and assigning only one wavelength to transport this traffic. The second scheme does not consider the aggregation of multiple RRHs traffic in one ONU/CPRI line rate so a unique ONU/CPRI line rate and a unique wavelength is assigned to each RRH of the cell.

The first assignment scheme allows wavelengths that would be uniquely assigned to one RRH to be assigned to multiple RRHs. Note that as the wavelength of the TWDM-PON has limited capacity, the second scheme can be used when the capacity of a wavelength can not meet the entire demand of multiple RRHs aggregated in one CPRI line rate, avoiding the blocking of the transmission.

III. Placement of Processing Functions in CF-RAN

The decision of where to place the processing functions considering a tradeoff between the required network capacity and the total power consumption of the network delineates a network planning problem. This network planning must decide when to activate local processing functions at the fog nodes to promote energy efficiency while serving the bandwidth demands of the cell-sites and considering the available wavelengths of the TWDM-PON.

As ILP is a relevant tool used to solve planning problems, we propose an ILP model to decide the optimal placement of the processing functions. In the next subsections we present the power consumption model used to model the power consumption of the network elements and the ILP formulation.

A. Power Consumption Model

We use a power consumption model based on the parameters introduced in [3] and [7] to model the consumption $i)$ of a stand-alone BBU placed with a RRH on each DRAN node, $ii)$ of a (cloud or fog) node that consolidates BBU processing and $iii)$ of each demand in terms of the line card/OLT ports used and the VM established for its processing.

The power consumed to maintain a stand alone BBU at the RRH is equal to 600 W, and this consumption is related to the BBU itself. The consumption of cloud and fog nodes nodes has a fixed value of 500 W and an additional consumption of 230 W per each hosted server. For each transmitted wavelength, we assume the power cost of 5 W of using one line card/OLT port. For each demand sent to a node, we also assume that it will consume 18 W per VM for the baseband processing. Finally, each processing node (cloud or fog) has an OLT that has a base consumption of 100 W [8].

Note that we only consider the power consumption related to the transmission and processing of demands. The power costs related to the radio network architecture and operation are not in the scope of this work as they do not depend by the placement of processing functions.

Table I summarizes all the power consumption values assumed in our model.

TABLE I
POWER CONSUMPTION PARAMETERS

Element	Cost (watts)
DRAN node	600 watts
Cloud and fog node base consuming	500 watts
Server fixed consuming	230 watts
Server variable consuming (VM)	18 watts
Line Card	5 watts
OLT	100 watts

B. ILP model

To define which fog nodes must be activated according to the traffic requests, we propose the following ILP optimization model. The formulation defines the best set of fog nodes to be activated to process a set of demands considering the availability of the TWDM-PON and the best overall power consumption of the network.

The proposed ILP model is a variation of the well known Bin-Packing (BP) problem, which is known to be NP-hard. In our formulation, the demands are the items and the optical links are the bins, with their total capacities divided over the available wavelengths. So, by solving the BP problem, we also determine the minimum number of fog nodes needed to meet the demand. Since deploying more fog nodes increases power consumption, the minimization of the number of fog nodes leads to an overall power-consumption minimization.

Input Parameters

R: set of RRH/Cell traffic demands i

N: set of all possible processing nodes n

W: set of wavelengths w at an optical link

B_i: bandwidth demand of RRH/Cell i

B_{nw}: capacity of wavelength w at node n

C_i: power cost of demand i

C_n: power cost of node n

Decision Variables

y^i_{nw}: = 1 if the traffic demand of RRH/Cell i is processed at node n being transmitted at wavelength w, 0 otherwise.

x_n: = 1 if node n is activated, i.e., works as a fog node, 0 otherwise.

Objective Function

$$Minimize \sum_{i=1}^{R} \sum_{n=1}^{N} \sum_{w=1}^{W} y^i_{nw} * C_i + \sum_{n=1}^{N} x_n * C_n$$

Constraints

(1) $\sum_{n=1}^{N} \sum_{w=1}^{W} y^i_{nw} * B_i \leq B_{nw} \mid \forall\ i \in R$

(2) $\sum_{n=1}^{N} \sum_{w=1}^{W} y^i_{nw} = 1 \mid \forall\ i \in R$

(3) $\sum_{i=1}^{R} y^i_{nw} \leq 1 \mid \forall\ n \in N, \forall\ w \in W$

(4) $y^i_{nw} = 1 \Rightarrow x_n = 1$

(5) y^i_{nw}, $x_n \in \{0, 1\}$

Constraint (1) ensures that the bandwidth demand of a RRH/Cell will never exceed the capacity of a wavelength. Constraint (2) ensures that the model only allocates one RRH/Cell demand to one wavelength and constraint (3) ensures that a single wavelength is assigned to one single demand. The decision variable x_n is used to calculate the power consumption of the active nodes at the network. Every time that a demand i is allocated to node n ($y_{nw} == 1$), constraint (4) ensures that the power consumption of that node will be computed. This constraint helps the model to allocate the demands in the most energy efficient way. Finally, constraint (5) assigns only binary values to the decision variables.

The activation of fog nodes is done when the capacity of the fronthaul is exhausted. So, when all optical resources of the fronthaul are used, i.e., all wavelengths are assigned to transmissions, a fog node is activated, and when the network capacity to transport CPRI data to this fog node is exhausted, another fog node is activated, and so on, until all demands are placed to a node. Note that the processing of the CPRI data of a cell-site can be divided between the cloud and in fog nodes. For instance, if a cell-site has 20 RRHs, and the fronthaul only support the transmission of the data from 15 of these RRHs, the remaining RRHs data will be transmitted to a fog node.

So, note that at the optimum solution the formulation will always exhaust the capacity of one node before start activating another one. It will check the bandwidth of each fiber and allocate the demands as there is enough capacity on a fiber. Only when the capacity of a fiber is met, the formulation will start allocating demands on the next available fiber, and so on. This behaviour is the necessary to schedule specific demand sets among the nodes of the CF-RAN considering the capacity of optical resources and do not require high costs of operation or high computing complexity, as we are considering demands with fixed size.

IV. PERFORMANCE EVALUATION

The performance of the CF-RAN is evaluated through executions of the ILP model using the CPLEX software. We considered a TWDM-PON with 10 wavelengths of 10Gbps. The total length of the links is set to 10km to respect the delay constraints imposed by the CPRI standard. We considered RANs formed by 5, 10 and 15 cell-sites. Each cell-site is equipped with 3, 4 or 5 RRHs. Each RRH covers one cell-site sector, thus generating 2.4Gbps. We evaluated the energy performance of the CF-RAN in comparison with the DRAN and the maximum cell-sites coverage that the CRAN and the CF-RAN can support. We also present the breakdowns of power costs during the operation of CF-RAN.

Figure 3 compares the power consumption of the CF-RAN and the DRAN. As it can be seen, for both wavelength assignment schemes, the CF-RAN has superior performance, being able to reduce the power consumption in 89.5%, 94.2% and 92.5% under the CA wavelength assignment scheme and 77.8%, 82.4% and 80.8% under the RRHA scheme in comparison to the DRAN in 5, 10 and 15 cell-sites with

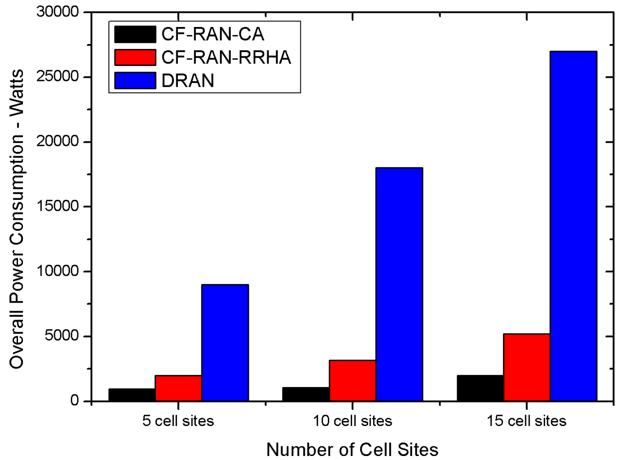

Fig. 3. Power consumption of CF-RAN and DRAN for different RAN sizes with 3 RRHs per cell-site

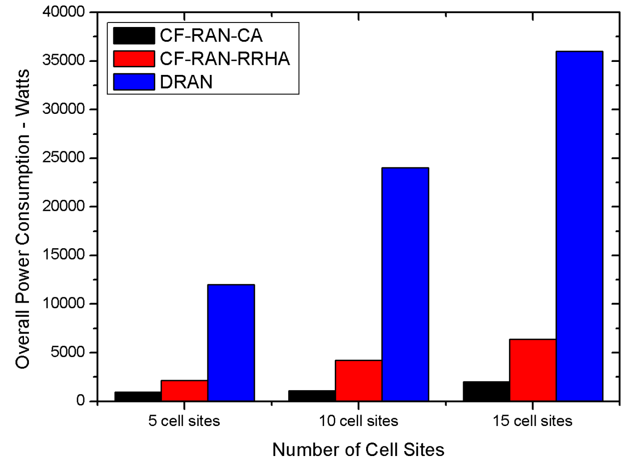

Fig. 4. Power consumption of CF-RAN and DRAN for different RAN sizes with 4 RRHs per cell-site

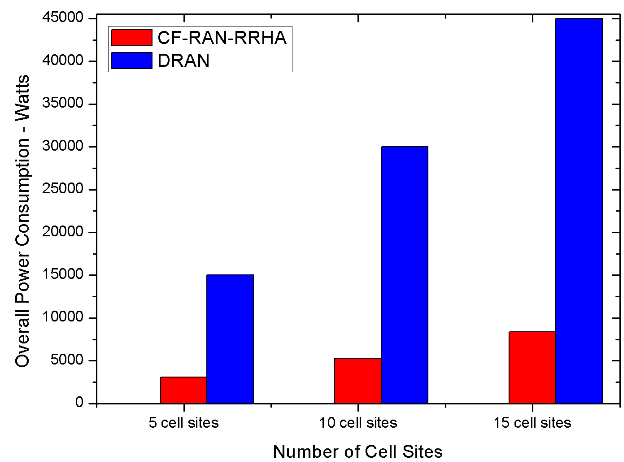

Fig. 5. Power consumption of CF-RAN and DRAN for different RAN sizes with 5 RRHs per cell-site

3 RRHs each one, respectively. This enormous difference of power consumption between DRAN and CF-RAN can be explained by the fact that in DRAN, all the processing functions and the infrastructure to support the computing usage are always active for each single RRH/DRAN node, while in CF-RAN, the infrastructure for supporting processings are activated on demand and many processing demands can be processed in a single node by virtualization, considerably reducing the cost in comparison to DRAN. Specifically, with the proposed TWDM-PON, each CF-RAN node can process 10 RRH demands per each active node, while in DRAN 10 demands would require 10 active DRAN nodes. The CA scheme has superior energy performance over the RRHA scheme because aggregating as many RRH data as possible in one CPRI line and one wavelength will demand less power consumption from the optical network and only demand the instantiation of a single vBBU per wavelength transmitted. So, in comparison with the RRHA scheme, the activation of processing functions will grow less in CA scheme, although the RRHA schemes also has superior energy performance than DRAN.

Similar energy performance behaviour can be observed in Figures 4 and 5, that show the power consumption of the network serving cell-sites with both 4 and 5 RRHs. In Figure 4, the CA scheme again shows best power consumption than the RRHA scheme serving 5, 10 and 15 cell-sites, with gains of 92.2%, 95.5% and 94.4%, respectively, compared to the DRAN, while the gains for the RRHA scheme over the DRAN was 82.4% for the three number of cell-sites. When the densification of RRHs in a cell-site grows, both the CF-RAN and DRAN power consumption will grow. Specifically, as can be seen in Figure 5, although the CA scheme has energy performance superior than the RRHA scheme in the previous figures, the CA scheme can not be used when the aggregation of multiple RRHs data into a CPRI line surpass the capacity of the wavelength, which in our executions was surpassed with the amount of 5 RRHs per cell-site (Figure 5). In this case, only the RRHA scheme can be used. The RRHA scheme has gains of 79.6%, 82.4% and 81.5% over the DRAN for 5, 10 and 15 cell-sites, respectively. The conclusion here is that for small groups of RRHs in each cell, the CA scheme is the best energy-efficient wavelength assignment scheme, but when the cell-sites grows, the RRHA scheme needs to be used to provide the coverage of more dense cell-sites to avoid the blocking of RRHs data transmissions.

In Figure 6, we compared the coverage capacity of CF-RAN and CRAN, for cell-sites with groups of 3, 4 and 5 RRHs, respectively, under the CA and RRHA schemes. Note that for 5 RRHs there is no result for the CA scheme because the capacity of the wavelength is surpassed when 5 RRHs traffic are aggregated into one CPRI line for RRHs that covers one cell-site sector (for RRHs with greater coverage, the number of supported aggregated RRHs into one wavelength will be less). It is clearly noted that in all cases the CF-RAN has greater coverage of the cell-sites. This is explained because CRAN is only using one processing node to support all the RRH traffic and it does not use additional processing nodes

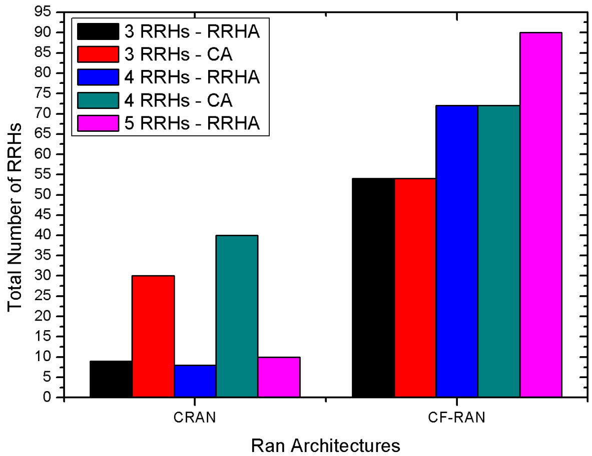

Fig. 6. Coverage of CF-RAN and CRAN for cell-sites with 3, 4 and 5 RRHs

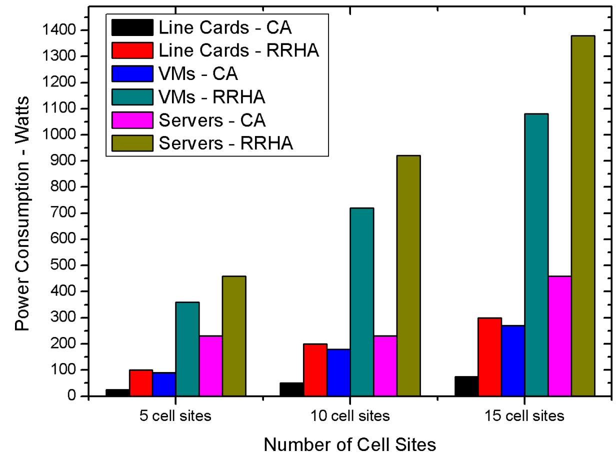

Fig. 7. Breakdowns of the power costs associated to CF-RAN operation

as CF-RAN does. For 3 and 4 RRHs, the coverage of CF-RAN under both CA and RRHA schemes are the same, but as discussed earlier, the power consumption of CA will be lower than RRHA, although RRHA is able to cover more densified cell-sites, as the operation of the CA scheme is limited to cell-sites of 4 RRHs with the considered MIMO configuration per RRH, as it can be seen by the absence of results to the CA scheme for 5 RRHs. On the other hand, it is important to note that in CRAN architecture, the use of the CA scheme increases the RRH coverage in comparison to the RRHA scheme, and must be considered for operators that uses CRAN with limited processing nodes to provide its services.

Finally, in Figure 7 we show the breakdowns of power costs for the processing of RRH demands. Only the 4 RRHs scenarios is presented due to space limitations. Note that the most consuming elements are the servers and the VMs, while the line cards used to transmit the traffic contribute with low power consumption, stating the energy-efficiency of the TWDM-PON and that it does not account much power costs to the overall power consumption of CF-RAN operation. Again, the CA scheme has lower power consumption due to the aggregation of multiple RRHs data into one transmission line.

V. Conclusion

In this paper we proposed the use of TWDM-PON to support the 5G traffic under an architecture called CF-RAN, combining the Fog Computing and NFV paradigms. We proposed an ILP formulation to plan the placement of processing functions in the CF-RAN and two wavelengths assignments schemes to transmit the traffic from the RRHs to the BBU pool. We stated that the CF-RAN can serve more cell-site sectors than a CRAN architecture and operate with much higher energy efficiency in comparison with the DRAN architecture. We also stated that the transmission of a cell-site traffic in one wavelength brings energy savings but at the cost of reducing the coverage of the network. On the other hand, the transmission of cell-site traffic in multiple wavelengths allows the network to cover more cell-site sectors. Finally, the breakdowns of power costs was presented and the low power cost contribution of the TWDM-PON was observed. In future works, we will expand our ILP model to deal with demands of different classes of services and compare the CF-RAN to H-CRAN and also study the use of elastic optical networks supporting 5G traffic.

Acknowledgment

This work is part of the INCT of the Future Internet for Smart Cities (CNPq 465446/2014-0, CAPES 88887.136422/2017-00 and FAPESP 2014/50937-1) and CNPq grant 420907/2016-5.

References

[1] J. Wu, Z. Zhang, Y. Hong, and Y. Wen, "Cloud radio access network (C-RAN): a primer," *IEEE Network*, vol. 29, no. 1, pp. 35–41, 2015.

[2] A. De la Oliva, J. A. Hernández, D. Larrabeiti, and A. Azcorra, "An overview of the CPRI specification and its application to C-RAN-based lte scenarios," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 152–159, 2016.

[3] 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, 2014.

[4] G. B. Figueiredo, X. Wang, C. C. Meixner, M. Tornatore, and B. Mukherjee, "Load balancing and latency reduction in multi-user CoMP over TWDM-VPONs," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[5] X. Wang, L. Wang, C. Cavdar, M. Tornatore, G. B. Figueiredo, H. S. Chung, H. H. Lee, S. Park, and B. Mukherjee, "Handover reduction in virtualized cloud radio access networks using TWDM-PON Fronthaul," *Journal of Optical Communications and Networking*, vol. 8, no. 12, pp. B124–B134, 2016.

[6] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 126–135, 2014.

[7] I. Waßmann, D. Versick, and D. Tavangarian, "Energy consumption estimation of virtual machines," in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*. ACM, 2013, pp. 1151–1156.

[8] J. Baliga, R. Ayre, W. V. Sorin, K. Hinton, and R. S. Tucker, "Energy consumption in access networks," in *OFC/NFOEC 2008 - 2008 Conference on Optical Fiber Communication/National Fiber Optic Engineers Conference*, Feb 2008, pp. 1–3.