Cell Deployment Optimization for Cloud Radio Access Networks using Teletraffic Theory

Andrijana Popovska Avramova, Henrik Lehrmann Christiansen and Villy Bæk Iversen

Department of Photonics Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark
Email: {apop, hlch, vbiv}@fotonik.dtu.dk

***Abstract*—Cloud Radio Access Network (C-RAN) is a new mobile radio access network design based on centralized and pooled processing. It offers potential cost savings by utilizing the so-called tidal effect due to user mobility in cellular networks. This paper provides a quantitative analysis of the performance (multiplexing) gain of such cellular networks. The used analytical model is based on a multi-dimensional loss system and can be applied to heterogeneous networks with various cell traffic profiles. Based on the analysis, the key parameters for cell deployment optimization are identified. The conditions for optimization are based on the aggregated traffic characteristics and baseband unit pool dimensioning. This paper considers cells with different traffic profiles and the optimal conditions for maximization of the pooling gain are determined. Furthermore, it is shown how the model can be applied to dynamically re-assign cells to a pool of baseband units. The re-assignment is based on the cell load and traffic characteristics such that effective utilization of the baseband resources is assured.**

Keywords–C-RAN, deployment optimization, multiplexing gain, baseband unit pool dimensioning, multi-dimensional loss system.

I. INTRODUCTION

The explosive increase in mobile traffic is a main driver for a spectrum, energy, and cost efficient design of the future radio access network (RAN). Network densification is a prevailing technique that addresses the challenge of 1000-fold traffic growth of mobile data. The full benefits of network densification can be realized if it is followed by complementary backhaul technology [1], such as Cloud RAN (C-RAN). C-RAN is a scalable and flexible RAN design where the baseband processing is virtualized, centralized and shared among base stations (BS). The centralization of the processing power enables high cooperation among distributed antennas. Virtualization on the other hand allows for processing aggregation and dynamic resource allocation. Thus, C-RAN reduces the operators capital and operating expenditures, provides high spectral and energy efficiency. C-RAN supports coexistence of multi-standard types of communication (device to device, full duplex), and multi-layer architectures. Additionally, C-RAN facilitates the deployment of services at the edge, opens new opportunities for services in the cloud, such as the ability to offer the radio access network as a service [2].

The C-RAN architecture consists of three main parts: remote radio heads (RRHs) that provide the wireless coverage, baseband unit pool (pool of virtual BSs) and a transport network (fronthaul) that connects the BBU pool with the RRHs. The up to date research confirms that the C-RAN design simplifies and reduces the cost of dense cell deployment [3]. Yet the conditions for optimal deployment under C-RAN remain an important area of research. The need for analysis, design and optimization of fronthaul and backhaul technologies for 5G is emphasized in a recent draft proposal of the pre-structuring model for the Horizon 2020 5G Infrastructure PPP [4]. In this work, traffic engineering approach is used in order to perform a quantitative study of C-RAN, and indicate the conditions for optimal multiplexing gain and dimensioning of the BBU pools. The model used in this paper is generalized and can be used for heterogeneous network deployments under various traffic models. The goal of this paper is to determine the key performance metrics that maximize the multiplexing gain. Furthermore, in our model, the optimal dimensioning of the BBU pool considers both the cost saving factor as well as the sensitivity to traffic variations. As a baseline, we consider a network consisting of two cell types that generate different traffic profiles. The work suggests the optimal ratio of the two types of cells for an energy efficient BBU pool, and how the architecture can adapt to the changes in the traffic conditions.

The remainder of the paper is organized as follows: Section II provides an overview of related works. Section III presents the model based on direct routing and how it is mapped to the C-RAN architecture. Section IV discusses the approach taken in this paper for evaluation of the multiplexing gain and dimensioning of the BBU pool. Section V presents and analyses the results for a specific case with respect to multiplexing gain and dimensioning, and elaborates how the model can be applied for dynamic mapping between RRH and BBU pools. Finally, the last section concludes the paper.

II. RELATED WORK

As indicated in [5] the main multiplexing gain in C-RAN comes from the fact that the cells have diverse traffic load during day hours depending on the area they serve. This is the so-called ”tidal effect” since the load in the mobile network moves according to the daily routine of the users. During the working hours more users are located in the office areas, hence the BSs associated to those cells are busiest. After working hours, the users move towards the entertainment and residential areas, increasing the traffic demand on the BSs associated to these cells. In case of traditional deployment, the residential cells during working hours and the office cells during evening hours will be underutilized. The benefit of dynamic assignment of baseband processing to RRHs (illustrated in Figure 1) has been analyzed in [6] through a system level simulation of a scenario where the generated traffic pattern follows the tidal effect. The paper shows that the multiplexing gain comes not only from the fact that the computational power can be shared among

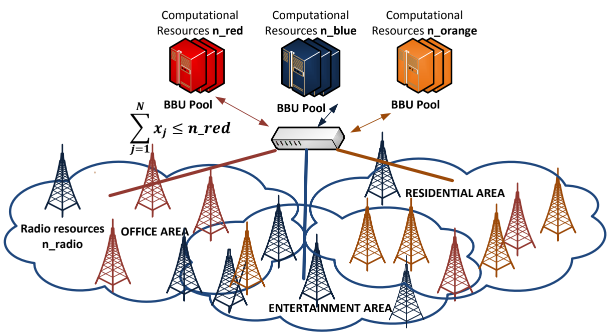

Figure 1. Dynamic allocation of RRHs to a BBU pool. The assignment is defined by different colors.

BSs but also from cost and energy efficiency perspectives. In [7] and [8], the need for dynamic RRH-BBU association is emphasized. Their work shows that the configuration in the network must be flexible in order to provide high performance and energy efficiency. Semi-static and dynamic RRH-BBU switching schemes have been proposed and analyzed with respect to efficiency in the BBU pool. The results show that a percentage of BBUs can be reduced, depending on the traffic load and the applied scheme for assignment.

In [9], the authors model the dynamics of the BBU pool with a multi-dimensional Markov model. The work shows that the system parameters such as pool size, QoS requirements at the radio part, and the traffic load have impact on the system design. In their analysis, all the cells that are associated in a common pool of BBUs, have the same characteristics: size (BS transmission power), type of traffic, QoS demand. Therefore, the proposed model cannot be directly used if heterogeneous deployments are analyzed. In this paper we present a model that can be used to estimate the performance metrics for a C-RAN architecture that can include cells with different size (as number of radio resources), cells with different traffic profiles (smooth, bursty and random), services that have different QoS requirements (as a minimum number of resources that need to be allocated), as well as multi-layer deployment (small cells overlapped with macro cell). The following performance metrics have been studied: blocking probabilities and carried traffic. Based on the desired resource utilization, we dimension the pool of BBUs using the Moe's principle for network dimensioning. We evaluate the dimensioning of the backhaul link based on the carried traffic characteristics. The network model considered in the numerical analysis is based on a mixture of residential and office cells. For the considered scenario, a method is proposed for determining the optimal ratio of the two cell types for multiplexing gain maximization.

III. Network Model

This section presents the mathematical model used to asses the benefit of placing baseband processing in a pool that can be shared among RRHs. First, the direct routing network model based on the multi-dimensional systems is described. Afterwards, the mapping of the model to the three different scenarios of network layout is explained.

A. Link model

In a multi-dimensional system, a single link with capacity of n basic units (BUs) is shared among N statistically independent (uncorrelated) flows of Binomial, Poisson, and Pascal (BPP) traffic. A stream is characterized by mean value A_j (offered traffic in number of BUs), standard deviation std_j, the required number of BUs for the entire connection d_j, and n_j is the maximum number of BUs that can be occupied by flow j. The system state at any time can be described by the vector $(x_1, x_2, ..., x_N)$ where $x_j = i_j \cdot d_j$ and i_j represent the number of connections of a flow j. Then the restrictions that lead to truncation of the state space can be formulated as:

$$0 \leq x_j \leq n_j, \quad \sum_{j=1}^{N} x_j \leq n, \quad where \quad \sum_{j=1}^{N} n_j \geq n \tag{1}$$

In the case where the last two restrictions are not valid (n is sufficiently large such that there is no global restriction), the system corresponds to N independent one-dimensional loss systems (classical BPP loss system), that are represented by state probabilities $p_j(x_j)$.

The system described above is reversible and has product form. Due to the product form, the algorithm based on convolution [10] can be applied to obtain the individual performance metrics of each stream. By successive convolution of one flow at a time, the state probabilities can be aggregated and a one-dimensional vector can be used to describe the system (* denotes the convolution operation):

$$p(x) = p_1(x_1) * p_2(x_2) * ... * p_N(x_N), \tag{2}$$

where $x = x_1 + x_2 + ... + x_N$. The convolution is done such that first two flows j and k are convolved with limitation $\min(n_j + n_k, n)$. Then the third flow is added to the previous convolution and so on. Due to the truncation, normalization at each step needs to be performed in order to get the true state probabilities. To calculate the time, call, and traffic congestion for a flow j, all flows except j need to be convolved into $p_{N/j}$. The derivation of the three types of congestion is given in [11], here only the calculation for the carried traffic (in number of BUs) is presented:

$$Y_j^n = \sum_{x=0}^{n} \sum_{x_j=0}^{x} x_j \cdot p_{N/j}(x - xj) \cdot p_j(x_j) \tag{3}$$

and $C_j^n = (A_j - Y_j^n)/A_j$ represents the traffic congestion. By applying the above method, the performance measures for each flow can be derived.

B. Network with direct routing

A network with direct routing [12] is characterized by routes R_j representing different traffic flows, links L_k and $d_{j,k}$ as the number of BUs a route j uses on a link k. Each link is represented with capacity l_k that defines the maximum number of basic unit that all flows can use on that link. The restriction on each link can be expressed as:

$$\sum_{j=1}^{N} x_{j,k} = \sum_{j=1}^{N} i_j \cdot d_{j,k} \leq l_k, \quad k = 1, 2, ...K \tag{4}$$

All the routes are independent, hence the convolution algorithm can be applied to aggregate the state probabilities of any two route to one route, until one route remains for which the performance metrics are calculated. Now, during convolution, each link has to be considered one at a time, as a restriction to the state space. Because each link can restrict one or more routes, the number of busy channels at each link, or the number of connections at each routes need to be tracked. The algorithm becomes more complex since multi-dimensional vectors need to be convolved, where the number of links defines the dimension. The state number increases to maximum $\prod_{k=1}^{K}(l_k + 1)$, which requires large memory for calculation.

C. Network layout mapping to a C-RAN deployment

Using the model presented, the following notation will be used throughout the paper to describe a C-RAN network. A BBU pool is associated with N RRHs, where a RRH j can use up to n_j radio resources. The number of baseband processing power (or computational resources) in a BBU pool is given by n, where $n \leq \sum_1^N n_j$. The traffic at RRH j is represented through the mean value of offered traffic A_j and standard deviation std_j. A call j requires d_j radio and computational resources for the entire duration of a connection. In the multi-dimensional Markov model there will be two types of truncations. The truncation due to the limited radio resources is referred to as blocking probability due to radio resources, while the truncation that is resulted from n is referred to as blocking probability due to computational resources (BBU pool limitation). Hence, for each traffic flow, the call blocking probability depends on the blocking probability due to radio resources and blocking probability due to computational resources.

Using the model with direct routing, the system can be represented through a matrix where the routes are identified as columns, and the links are defined by rows. We consider three different deployment scenarios in C-RAN in order to explain how the analysis can be performed. The reason for this is to show that this method is general and that the complexity of the algorithm can be highly reduced. The reduction can be done both in terms of dimensions of convolution vectors as well as in number of convolutions, due to reduced dependences on the links and generalizations on the cells characteristics.

D. Case study: proportion of office and home small cells

The direct routing equivalent for a network where the BBU pool aggregates a proportion of cells that serve office and residential area is presented in Table I. The number of office RRHs is O, where each RRH has n_o radio resources. The number of residential cells is $N - O$ where each RRH has n_r radio resources. The office cells are offered bursty traffic model with equal mean and standard deviation (Pascal distribution). The traffic at the office cell is modeled using smooth model (Engset distribution) and has equal characteristics among all residential cells. In this paper, this case study is considered as baseline for evaluation of the multiplexing gain in C-RAN. As it can be seen, the table consists of an identity matrix of dimension N. Hence, the complexity of the method described in Section III is highly reduced: the number of the convolutions required to get the performance metrics of one traffic stream is reduced to N. Since there are no dependencies among cells, except the last row, the aggregation of the streams can be done into one-dimensional vectors, and only the global state needs to be remembered. Thus, the number of the states and the required memory is of complexity $O\{n\}$.

E. Case study: mixture of traffic types

This case corresponds to the heterogeneous traffic characteristics in terms of BUs that a stream requires during the connection. Video services that require high bandwidth can be modeled with $d_j > 1$. Table II shows the equivalent direct routing model for C-RAN that aggregates RRH, that offer heterogeneous services in terms of bandwidth demand $d_1 \neq d_2$. The model is for a case of 4 non-overlapping cells, which can be easily extended to more cells. The two traffics types can have individual mean value and standard deviation, while the radio resource limitation could be the same or different. The limitations L_5 and L_6 could be left out, or used when QoS guarantee needs be implemented to make sure that the increase of one type of traffic does not block the other type of traffic. The complexity of the algorithm is again reduced due to the symmetry. The number of convolution for each individual traffic stream (in this case two) is equal to the number of cells, while the dimension of the each convolution vector is equal to the number of different traffic flows. Hence, in the considered example the number of states and the required memory is of complexity $O\{(n_{d1} + 1) * (n_{d2} + 1)\}$.

TABLE I. Direct routing equivalent to C-RAN that covers a mixture of office and home cells

Link	Routes R_1 R_2 ... R_O	R_{O+1} .. R_N	Capacity
L_1			n_o
...	Identity matrix of size O	Zero matrix of size $[R, O]$	...
L_O			n_o
L_{O+1}			n_r
...	Zero matrix of size $[O$ x $R]$	Identity matrix of size R	...
L_N			n_r
L_{N+1}	all ones vector of size $[1, N]$		n

TABLE II. Direct routing equivalent to C-RAN that covers a mixture of traffic types for each cell

	Routes								
	Cell1		Cell2		Cell3		Cell4		
Link	R_1	R_2	R_1	R_2	R_1	R_2	R_1	R_2	Capacity
L_1	d_1	d_2	0	0	0	0	0	0	n_r
L_2	0	0	d_1	d_2	0	0	0	0	n_r
L_3	0	0	0	0	d_1	d_2	0	0	n_r
L_4	0	0	0	0	0	0	d_1	d_2	n_r
L_5	d_1	0	d_1	0	d_1	0	d_1	0	n_{d1}
L_6	0	d_2	0	d_2	0	d_2	0	d_2	n_{d2}
L_7	d_1	d_2	d_1	d_2	d_1	d_2	d_1	d_2	n

F. Case study: Multi-layer deployments

Multi-layer heterogeneous deployments are considered as a way of increasing the throughput per area. A scenario where a BBU pool covers cells with different sizes, and traffic offloading exist among overlapping cells, should be considered. The analysis of such a case, should reveal the optimal number of small cells per sector of a macro cell, and could be used to indicate how to dimension BBU pool, depending of the traffic offloaded from the macro cells to the small cells. A direct routing equivalent for a three sector macro cell with two small cells per sector is shown in Table III. All small cells have the same characteristics for the offered traffic and size of a cell (n_m for macro cell and n_s for small cells). The traffic streams in the small cells can also use radio resources in the macro cells, with call rearrangements [12]. Regarding the complexity analysis, this is the most complex case compared to the previous case studies. Two sectors can be easily aggregated into one dimensional vector, so the number of one dimensional convolution is equal to double the number of small cells per sector ($small\_nr\_sector$). In order to find out the performance metrics for each traffic stream (one for macro cell and one for small cell), the algorithm requires one convolution vector of dimension equal to the number of small cells per sector. Then the number of states increases to order of $(n+1)\cdot(n_m+1)\cdot \prod_{small\_nr\_sector} (n_s+n_m+1)$.

TABLE III. Direct routing equivalent to C-RAN for multi-layer deployment

Links	Routes Sector1			Sector2			Sector3			Capacity
L_1	1	0	0	0	0	0	0	0	0	n_m
L_2	1	1	0	0	0	0	0	0	0	n_m+n_s
L_3	1	0	1	0	0	0	0	0	0	n_m+n_s
L_4	0	0	0	1	0	0	0	0	0	n_m
L_5	0	0	0	1	1	0	0	0	0	n_m+n_s
L_6	0	0	0	1	0	1	0	0	0	n_m+n_s
L_7	0	0	0	0	0	0	1	0	0	n_m
L_8	0	0	0	0	0	0	1	1	0	n_m+n_s
L_9	0	0	0	0	0	0	1	0	1	n_m+n_s
L_{10}	1	1	1	1	1	1	1	1	1	n

IV. Discussion on multiplexing gain and BBU pool dimensioning

This section outlines the approach considered for evaluation of the multiplexing gain and the conditions for optimal dimensioning and configuration of the pool. The rationals for the considered performance metrics are discussed as well.

A. *Multiplexing Gain*

In [9], it is demonstrated that, as more cells are aggregated to the BBU pool, the gain (defined as reduction of the number of BBU processing servers that are required to achieve a blocking probability lower than a certain threshold) is increasing. Furthermore, it is shown that as the pool size becomes large, the gain is increasing with a slow pace, such that at a very large pool size, the gain is approaching a limit. Still, the work is missing a discussion on the background for such trend of the gain. The increase in the multiplexing gain comes from the principle of group conservation [13]. In order to explain better, a comparison is made on the n number of resources (BUs) required to achieve a blocking probability of 1% in case of serving individual streams and an aggregation of the N streams. Figure 2 shows the comparison when $N=100$ traffic streams are considered, each with mean value of 10 (offered traffic is 10 erlang) and $std=\sqrt{\sigma^2}=\sqrt{10}$ (Poisson arrivals). The dashed line shows the normalized number of BUs (n/N) when the traffic streams are served independently, which is constant. The full line shows the normalized number of BUs required to serve the aggregated traffic that is decreasing as N is increased until a certain point after which it reaches a limit and becomes almost constant. The reason for this comes from the fact the way the summary statistics are derived for the aggregated stream. Since each stream is independent of the others, the mean and the standard deviation are calculated as:

$$A_{agg}=\sum_{1}^{N}(A_j),\quad std_{agg}=\sqrt{\sum_{j=1}^{N} {std_j}^2} \tag{5}$$

These equations indicate that the mean value of the total traffic is the same in case of individual streams and stream aggregation. The difference is in the standard deviation, or the

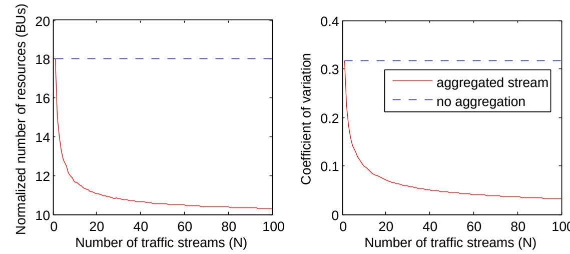

Figure 2. Analysis of multiplexing gain with aggregation.

coefficient of variation ($CV=\frac{std}{A}$) which is shown in Figure 2 to the right. The CV is reduced as the number of streams is increased, but already after $N=30$ the reduction is slow. Any additional increase of N will not lead to significant reduction of the number of required BUs. The channel utilization, defined as A/n will not be significantly improved at large pools, since any marginal increase of the offered traffic will lead to equal increase of number of BUs for each group, meaning $\frac{\Delta A}{\Delta n}$ becomes constant. This means that very large pools will not lead to significant increase of the gain compared to medium size pools. Due to high utilization, very large groups are even more sensitive to overload, and therefore large pools are not recommended. For that reason, the tradeoff between utilization and sensitivity should be considered when dimensioning.

Having in mind the discussion above, the multiplexing gain defined as in (6) is used as a performance metric to evaluate how much the coefficient of variation is reduced in case of the aggregating the individual streams.

$$\text{MultiplexingGain}=\frac{\sum_{j=1}^{N}(A_j+std_j)}{(A_{agg}^{carried}+std_{agg}^{carried})} \tag{6}$$

s.t. $A_{agg}^{carried}$ and $std_{agg}^{carried}$ are carried traffic characteristics.

B. *Dimensioning of computational resources.*

Two approaches of dimensioning can be considered: dimensioning with fixed blocking probability and dimensioning with fixed improvement function. With fixed blocking probability, the dimensioning of the BBU pools is done by restricting the time congestion to a threshold such that the number of calls that need to re-attempt the connection will be low. This type of dimensioning can easily lead to a system with high utilization (large pool size), but also very sensitive, since it does not consider the channel utilization.

On the other hand, the Moe's principle for dimensioning is based on the improvement function. The improvement function is defined as the increase in carried traffic when the number of channels (n) is increased by 1, $F_n(A)=Y^{n+1}(A)-Y^n(A)$, where $Y^n(A)=\sum_{j=1}^{N} Y_j^n(A_j)$. In this case the point where the $\frac{\Delta A}{\Delta n}$ becomes constant for all BBU pools indicates the dimension of the pools. The improvement function can be set to a fixed improvement value F_{target}, such that balance between high utilization and sensitivity is be ensured. A cost requirement can also be included in determining the optimal number of computational resources. Then the improvement value depends on the cost of the additional resource such as cost of fiber, BUU unit and alternatively the cost of adding a RRH to a BBU pool. The increase of the carried traffic should be included as well as income, such that $F_{target}=\frac{cost}{income}$.

V. C-RAN Optimization

A. Input Parameters

In this work, the study case where a BBU pool aggregates RRH that cover residential and office areas is considered. The chosen parameters for the analysis follow the examples given in [5] and [6]. The total number of cells is $N = 100$, while the percentage of office cells is varied between 1% and 99% with 1% as step. Each cell has $n_r = n_o = 28$ radio resources, which limits the maximum number of computational resources at the BBU pool at $N \cdot n_r = 2800$. The offered traffic, and standard deviation of the office and residential cells are summarized in Table IV. The traffic streams will result in very low radio resource blocking probability. The overall blocking probability will be mostly influenced due to the diagonal truncation which results from the limitation of the resources in the BBU pool. Two sub-cases have been considered as two different time snapshot. One is from daytime when the traffic of the office cell is higher than the traffic from the home cells. The other is in evening time, when the traffic of the residential cells is higher. By considering these two snapshots, the dynamic of the traffic during one day can be captured.

TABLE IV. Input parameters

	Daytime		Evening time	
Cell type	Office	Home	Office	Home
Load	30%	10 %	5%	15%
Traffic type	bursty (Pascal dist.)	smooth (Engset dist.)	smooth (Engset dist.)	bursty (Pascal dist.)
A	8.22	2.77	1.27	4.75
std	3.51	1.44	0.99	2.43

B. Multiplexing gain

The multiplexing gain, according to (6) for the considered case study is shown in Figure 3. During day time the multiplexing gain is reduced as the number of office cell is increased. This is because the mean value is increased but the difference in the standard deviation does not give influence in coefficient of variation of the aggregated and the sum of the individual streams. During night time the opposite trend is observed: the multiplexing gain is increasing as the number of office cells is increased. In this case, the mean value of the aggregation stream is decreasing with the increase of the office cells, and the deviation of the aggregation stream becomes smaller compared to the individual streams.

$$MG = \frac{\sum_{j=1}^{N} \max((A_j + std_j)^{day}, (A_j + std_j)^{night})}{(A_{agg}^{carried} + std_{agg}^{carried})} \quad (7)$$

By looking at the multiplexing gain of the sub-cases of daytime and night time, the optimal ratio of office cells and home cells cannot be deducted. In order to capture the traffic dynamics during one day, (6) has been modified to (7). This metric is also shown in Figure 3 and as it can be seen it peaks at 22% of office cells. Hence, for the this case, the largest gain is achieved when the number of office cell is 22 out of 100.

C. Dimensioning the BBU pool

For dimensioning the BBU pool in terms of computational resources, we use the Moe's principle. We do not focus in this paper on the cost, nor the income. We use improvement

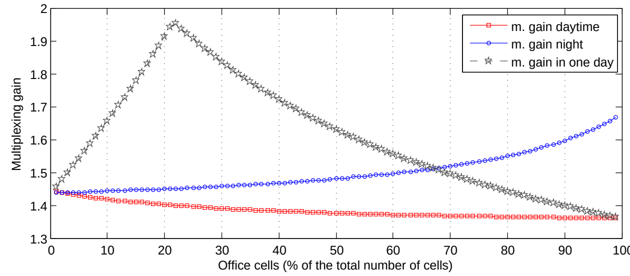

Figure 3. Multiplexing gain according to (6) and (7).

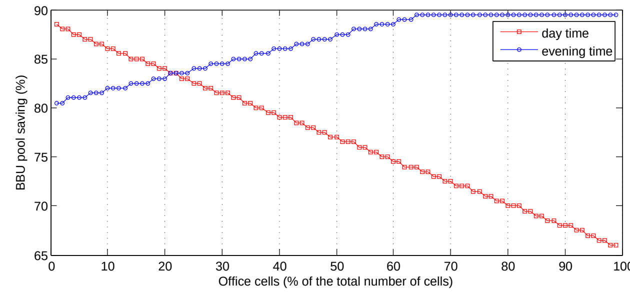

Figure 4. Optimal dimensioning of BBU pool.

value $F_{target} = 0.2$, such that $F_{n-1}(A) > F_{target} \geq F_n(A)$. The analysis has been done for the two considered sub-cases: daytime and night. Instead of giving the optimal number of computational resources, we indicate the percentage of the maximum number of computational resources that can be saved. Figure 4 shows the computational resources percentage that can be saved in case of multiplexing. In daytime analysis, the percentage of the saved computational resources is reduced with the increase of the number of the office cells. The reason for this is that the number of the computation resources scales with the mean value of aggregated traffic. As the mean value of the office cells traffic is larger than the mean value of the home cell traffic, by increasing the number of the office cells, the mean value of the aggregated traffic is increased. During evening time the opposite trend is observed: the increase of the percentage of the office cell reduces the mean value of the aggregated stream, and therefore less computational resources are required. From the figure, it can be noticed that the two lines cross at 22% of the office cells, meaning that with this ratio of office and residential cells, the same savings can be achieved during day time and night time. Hence, the optimal ratio of the office and residential cell is 22 office and 78 residential cells, by which almost 85% of the maximum resources in the pool can be saved. The analysis based on multiplexing gain and dimensioning on the BBU pool has shown the same results. Furthermore, the conclusion is comparable with the simulation based analysis in [6], which confirms the correctness of the described model.

D. RRH-BBU pool dynamic mapping

The optimal percentage of office cells for different mean values of the traffic streams for office and residential cells during day time and during night time is summarized in Figure 5. Additionally for each optimal deployment it shows the potential savings by dimensioning the size of the pool using the Moe's principle. The results show that in case of a change of the traffic characteristics, the model can be used for

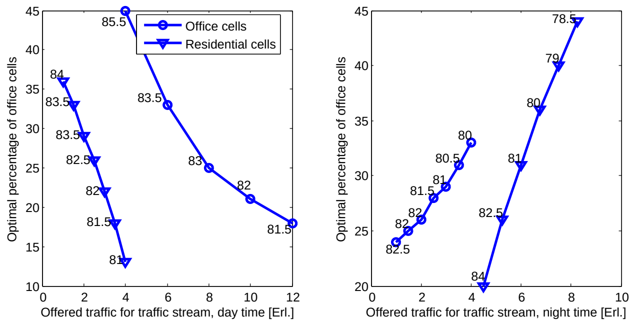

Figure 5. Optimal deployment for variable load during day time.

flexible and dynamic re-assignment of RRH to BBU pools. For example, if the mean value of the traffic stream for residential cells during night time is increased, the number of office cells per BBU pool need to be increased. On the other hand, if the mean value of the residential traffic stream during day time is increased, then the number of office cells need to be reduced.

The radio resourcce blocking probability is low as the load of the cell is not high (Table IV) and the overall blocking probability is influenced from the blocking probability due to computational resources. This is important as the model complexity is further reduced, as only the global state needs to be remembered, which can be described with one dimensional vector of length n. This simple analysis allows for adoption to the dynamic changes in the configuration. If a certain cell needs to be added to the BBU pool, a convolution needs to be performed in order to aggregate the new cell traffic. If one cell needs to be removed, deconvolution needs to be done.

The challenge of the fronthaul design is not just limited to high capacity requirement, but also to the ability to provide flexible and adaptive deployments with respect to RRH-BBU pool assignment. Fiber solutions are capable of supporting high data rates, but are lacking the ability for flexible re-assignment due to the need of manual configurations or very costly optical switches. Adopting any other transport solutions (ex. packet based: wired or wireless) is challenged with strict jitter and synchronization requirements but are capable of flexible reconfigurations. As C-RAN already integrates the concepts of network function virtualization and network virtualizaion ([14], [15]), adoption of software defined networking (SDN) can further optimize and simplify network design and operation. The proposed model can be implemented at an SDN controller. The SDN controller will be responsible for RRH to BBU pool re-assignment due to traffic distribution change and/or addition of new cells in the network. Thus, the SDN controller can instruct and manage all virtual network components in order to maximize the multiplexing gain and dimension the BBU pools optimally. Figure 1 illustrates the dynamic assignment of RRH to BBU pools, where not only the location, but the traffic load and type determine the assignment.

VI. Conclusion

This paper concludes the optimal conditions for dense cell deployments under which the multiplexing gain is maximized. In the presented study case, this is defined as the optimal ratio of the two types of cells: serving office and residential areas. The model has been compared with simulation based analysis, which confirms the correctness of the model. Additionally, we demonstrate that the model indicates the optimal ratio of the cell types depending on the individual traffic loads.

Furthermore, the analysis shows that not only cost, but sensitivity to traffic variations need to be considered when dimensioning the pool of baseband units. For the given ratio of the cell types, the indicated dimension is proven to be optimal.

The model used in the analysis is generalized, and various case studies have been identified. These studies include heterogeneous deployments and different traffic profiles. Due to its simplicity and low level of complexity, we show that the model can be adopted for dynamic re-assignment of RRH to BBU pool. In the future, additional cases are going to be studied, as well as further analysis will be conducted to investigate the implications of new radio technologies such as coordinated multipoint and carrier aggregation.

VII. Acknowledgment

This work was partially sponsored by the 7th Framework Programme for Research of the European Commission HARP project, under grant number HARP-318489.

References

[1] N. Bhushan, et al., “Network Densification: The Dominant Theme For Wireless Evolution into 5G,” Communications Magazine, IEEE, vol. 52, no. 2, February 2014, pp. 82–89.

[2] S. Ferreira, et al., “An architecture to offer cloud-based radio access network as a service,” in Networks and Communications (EuCNC), 2014 European Conference on, June 2014, pp. 1–5.

[3] A. Checko, et al., “Cloud RAN for Mobile Networks - A Technology Overview,” Communications Surveys Tutorials, IEEE, vol. 17, no. 1, Firstquarter 2015, pp. 405–426.

[4] “EC H2020 5G Infrastructure PPP Pre-structuring Model RTD and INNO Strands,” 2014, URL: http://5g-ppp.eu/wp-content/uploads/2014/03/March-2014-_5G-Infra-PPP_Pre-structuringModel_v1-0.pdf [accessed: 2014-12-01].

[5] “C-RAN The Road Towards Green RAN,” China Mobile Research Institute, Tech. Rep., 2011.

[6] A. Checko, A. Checko, H. Holm, and H. Christiansen, “Optimizing Small Cell Deployment by the Use of C-RANs,” in Proceedings of 20th European Wireless Conference, May 2014, pp. 1–6.

[7] C. Liu, K. Sundaresan, M. Jiang, S. Rangarajan, and G.-K. Chang, “The case for re-configurable backhaul in cloud-RAN based small cell networks,” in IEEE INFOCOM, April 2013, pp. 1124–1132.

[8] S. Namba, T. Warabino, and S. Kaneko, “BBU-RRH Switching Schemes for Centralized RAN,” in 7th International ICST Conference on Communications and Networking in China, Aug 2012, pp. 762–766.

[9] J. Liu, S. Zhou, J. Gong, Z. Niu, and S. Xu, “On the Statistical Multiplexing Gain of Virtual Base Station Pools,” in Global Communications Conference (GLOBECOM), 2014 IEEE, Dec 2014, pp. 2283–2288.

[10] V. B. Iversen, “The Exact Evaluation of Multi-Service Loss Systems with Access Control,” Teleteknik, English ed., vol. 31, no. 1, Firstquarter 1987, pp. 56–61.

[11] V. B. Iversen, Teletraffic Engineering. Chapter 7: Multi-dimensional loss systems. Technical University of Denmark, 2013.

[12] V. B. Iversen, V. Benetis, and P. D. Hansen, “Performance of Hierarchical Cellular Networks with Overlapping Cells,” in Proc. EuroNGI Workshop, 2004, pp. 7–19.

[13] M. Stasiak, M. Glabowski, A. Wisniewski, and P. Zwierzykowski, Modelling and Dimensioning of Mobile Wireless Networks: From GSM to LTE, 1st ed. Wiley Publishing, 2011.

[14] R. Wang, H. Hu, and X. Yang, “Potentials and Challenges of C-RAN Supporting Multi-RATs Toward 5G Mobile Networks,” Access, IEEE, vol. 2, 2014, pp. 1187–1195.

[15] C.-L. I, J. Huang, R. Duan, C. Cui, J. Jiang, and L. Li, “Recent Progress on C-RAN Centralization and Cloudification,” IEEE Access, vol. 2, 2014, pp. 1030–1039.

 ISBN: 978-1-61208-411-4