Characterizing the Relation between Processing Power and Distance between BBU and RRH in a Cloud RAN

Marcelo Antonio Marotta*, Hamed Ahmadi†, Juergen Rochol*, Luiz DaSilva‡, Cristiano Bonato Both⊤
*Federal University of Rio Grande do Sul, Brazil,
†The School of Electrical and Electronic Engineering, University College Dublin, Ireland,
‡University of Dublin, Trinity College, Ireland,
⊤Federal University of Health Sciences of Porto Alegre, Brazil
{mamarotta, juergen}@inf.ufrgs.br*, hamed.ahmadi@ieee.org†, dasilval@tcd.ie‡, cbboth@ufcspa.edu.br⊤

***Abstract*—In Cloud Radio Access Networks (C-RAN), base stations are replaced by Remote Radio Heads (RRHs) having their signals processed in Base-Band Unit (BBU). A BBU is limited to process the signal of RRHs within a maximum distance determined according to delay constraints. This delay is affected by three factors: (*i*) the distance between BBU and RRH; (*ii*) channel conditions; and (*iii*) processing power. This letter characterizes the relationship between these factors. Our results show that the processing power must be increased significantly for RRHs experiencing low Signal-to-Noise ratio (SNR) and/or at a longer distance from the BBU. The good news is that RRHs experiencing high SNR can have their processing take place in the cloud, at significant distances, and still meet tight Bit Error Rate (BER) and latency constraints. These two observations open the door to the possibility of dynamic load balancing mechanisms that change the locus of processing in a C-RAN according to channel conditions and availability of processing resources.**

***Index Terms*—Cloud Radio Access Network, Distance, Processing, Baseband Unit, Remote Radio Head**

I. Introduction

IN Cloud-Radio Access Networks (C-RANs), Base Stations (BSs) are replaced by Remote Radio Heads (RRHs): signal samples are digitized, transmitted through an optical infrastructure, and remotely processed in cloud data centers, which house Baseband Units (BBUs) [1]. Remote processing enables C-RANs to exploit the processing capacity available in the cloud, as well as to achieve load balancing and reuse of processing resources [2].

Remote processing incurs round-trip delay between BBU and RRH, which comprises the sum of (*i*) transmission, (*ii*) queuing, (*iii*) processing, and (*iv*) propagation delay components. Transmission delay at each intermediate node between BBU and RRH can be expressed as the ratio of the number of bits sent to the capacity of the outgoing link. Queuing delay is due to buffering the data at intermediate nodes. The BBU performs processing for signal demodulation, radio resource demapping, and precoding: the largest component of the processing delay is due to Forward Error Correction (FEC) [3]. Finally, propagation delay is given by the ratio of the distance between BBU and RRH to the speed of the signal transmitted in the link. This total delay must meet stringent latency requirements imposed by wireless communications mechanisms such as the Hybrid Automatic Repeat reQuest (HARQ) adopted in Long Term Evolution (LTE). In HARQ, the round-trip delay cannot surpass a fixed delay budget, regardless of whether local or remote processing is used. In the case of LTE, this delay budget is around 3 ms [4].

Delay constraints, in practice, dictate the maximum distance between an RRH and the BBU that processes its signals, limiting the area that a BBU can serve [2]. This distance is commonly considered to range between 20 and 40 km [1], the underlying assumption being that processing requirements and computing power (and, thus, processing delay) are fixed [5]. In fact, processing requirements change all the time according to the FEC processing that is most suitable for current channel conditions. In this letter, we show that it is possible to take advantage of these fluctuations to exploit unused processing power in the cloud, occasionally utilizing BBUs that are farther than 40 km away from the RRH.

This letter characterizes the relationship between channel conditions and the maximum distance between BBU and RRH, which has implications on processing load balancing and architectural decisions regarding the placement of the data centers that house the BBU pool. Our results include an analysis of the BBU processing power required to serve an RRH, under different channel conditions and target bit error rates. Our results demonstrate that channel conditions dictate the minimum processing power and the maximum distance supported between BBU and RRH and that it is possible to support RRHs that are far away in the network by slightly increasing the processing power allocated in the cloud for the RRH.

II. Round-Trip Delay Components

The round-trip delay components are depicted in the sequence diagram in Figure 1. A User Equipment (UE) transmitting to an RRH has its signal sampled, digitized, and forwarded for further processing by a BBU. The workload, comprising a set of in-phase and quadrature sample components, potentially traverses multiple hops between the RRH and the BBU. At each hop, transmission and queuing delay may be incurred to forward the sampled signal across the network. In the BBU, the decoding function of the FEC dominates the processing delay. Afterward, the BBU sends a reply back to the UE.

On the way from RRH to BBU, the workload traverses optical links, incurring propagation delay. The propagation delay for a fiber optic link is equal to $\frac{3d}{2c}$, where d stands for the total distance in meters between BBU and RRH, and c represents the speed of light.

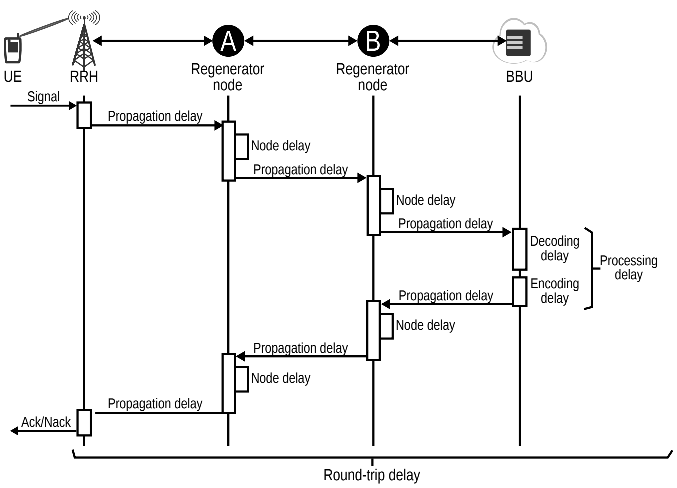

Figure 1: Round-trip delay components in a C-RAN.

Further delay is added when an intermediate node has to transmit/receive and enqueue the radio workload towards the BBU or RRH. We considered the best case scenario for the C-RAN connectivity, namely that the fronthaul consists only of dedicated fiber links to interconnect BBU and RRH, with no buffering required at intermediate nodes. Those only act as signal regenerators that must be inserted after a certain distance threshold to avoid optical signal degradation, *e.g.,* G.652 fiber requires a new regenerator after each 50 km. The ratio between the distance d over a threshold S, in km, indicates the number of intermediate nodes (signal regenerators) that must be inserted according to the optical technology in use: $\lceil \frac{d}{S} \rceil$. We assume a fixed per-node delay A, primarily to account for transmission delay.

The processing delay is the time consumed to process the radio signal, *e.g.,* the demodulation, coding, and radio resource demapping. In this processing, the FEC decoding is the most time-consuming function [3]. The decoding computation has its performance directly related to the number of recursions performed by the FEC, and the processing delay can be expressed as $\frac{kL_mF}{p_mO}$. The BBU executes k_m recursions of the FEC algorithm per code block m. L_m stands for the m^{th} code block length in bits, which can vary according to, *e.g.,* the technology in use, the coding rate, and the puncturing rate adjustment algorithm [6]. Each bit of the code block is usually processed through two identical constituent decoders of combined complexity F, expressed in operations per bit. The clock rate of the processor allocated to the processing of code block m is denoted as p_m (in Hz). The allocated processor has a processor efficiency O in operations per cycle.

Combining all the delay components discussed above, the round-trip delay between the BBU and the RRH can therefore be expressed as:

$$R = 2\left(\frac{3d}{2c} + A\left\lceil\frac{d}{S}\right\rceil\right) + \frac{k_mL_mF}{p_mO}. \qquad (1)$$

Among the factors we can control through design, we highlight the distance d in the first term: we can assign the workload of an RRH to a BBU that is nearby, in the fog, or farther away, in the cloud [7]. We also explore k_m and p_m in the second term: the former can be adapted according to current channel conditions and target bit error rate, and the latter reflects the processing power assigned in the cloud to an RRH. Given a round-trip delay budget Φ, to increase d, k_m must be decreased or p_m must be increased.

We validated equation (1) by comparing its outcome against experimental work reported in [3]. We consider the same scenario of [3], where a BBU has to decode a code block from an RRH and, setting R equal to a delay budget Φ, we find the maximum distance d between BBU and RRH.

The validation scenario parameters are as follows: A code block m to be decoded presents $L_m = 6144$ bits, the maximum code block size for LTE [8]. The decoder has a complexity of $F = 200$ operations per bit [9]. Considering a single core processor, we set $O = 1$ operation per cycle, providing a direct relation between the processor clock rate p_m and the total number of operations required to decode code block m. Based on a typical Brand A – G.652 fiber as the optical medium, we set $S = 50$ km as the maximum distance that can be covered without a regenerator. We consider an average node delay of $A = 20$ μs, as in [2]. As in [3], we set $k_m = 7$ recursions (fixed) and $p_m = 3.47$ GHz. Finally, we adopt a delay budget of $\Phi = 2.7$ ms, accounting for code block reply encoding and processing for the upper layers in the protocol stack [1], [3]. The result we obtain from equation 1 indicates a maximum distance $d = 20.86$ km between BBU and RRH, within 5% of the distance of 20 km condsidered in [3].

We now proceed to quantify the relationship between channel conditions, processing capabilities in the cloud, and the allowable distance between the RRH and the BBU.

III. Analyzing the Distance between BBU and RRH in a C-RAN

Load balancing in a C-RAN can be performed by allocating a BBU's processing power p_m at runtime according to demand, taking advantage of what is sometimes referred to as vertical elasticity [10]. In general, greater computational resources are available farther from the RRH (*i.e.,* in the cloud), so the allocation of those resources is affected by the maximum distance allowed between BBU and RRH, which, as we discussed in the previous section, is in turn dictated by delay considerations.

Considering that the decoding of a single code block is recursively computed without parallelism, we can characterize the relationship between the round-trip delay, the distance between the BBU and the RRH, and the processing power allocated in the fog or cloud through the number of decoder recursions performed by the FEC algorithm. In equation 2, we isolate k_m as a function of the round-trip delay budget Φ, the processing power p_m allocated in the cloud, and the distance d between BBU and RRH:

$$k_m = \left\lceil \frac{p_m O(\Phi - 3d/c - 2A\lceil d/S \rceil)}{L_m F} \right\rceil . \quad (2)$$

The decoder performance, in terms of BER, for the m^{th} code block can then be expressed as a function of the number of recursions of the FEC scheme and the SNR, expressed as $(\frac{Eb_m}{No})$, experienced during the transmission of the code block:

$$ber\left(\left\lceil \frac{p_m O(\Phi - 3d/c - 2A\lceil d/S \rceil)}{L_m F} \right\rceil, \frac{Eb_m}{No}\right). \quad (3)$$

The $ber(*)$ function can be empirically obtained from experiments or simulations with the desired recursive decoder. For a given target decoding performance and available processing power p_m, the round-trip delay R can be set equal to the delay budget Φ to obtain the maximum distance d between BBU and RRH in the following optimization problem:

$$\max_{d} \quad d \quad (4)$$

$$s.t.$$

$$ber\left(\left\lceil \frac{[p_m O(\Phi - 3d/c - 2A\lceil d/S \rceil)]}{L_m F} \right\rceil, \frac{Eb_m}{No}\right) \leq b \quad (5)$$

$$d \leq \frac{cS(\Phi p_m O - L_m F)}{p_m O(3S + 2Ac)} \quad (6)$$

$$d \geq 0 \quad (7)$$

In (4), we seek to maximize the distance d between BBU and RRH, subject to three constraints. The first constraint (5) sets the maximimum acceptable BER b. Constraint (6) sets an upper bound for d, corresponding to a single recursion of the FEC algorithm. The final constraint (7) guarantees that d is non-negative. The optimization problem is non-linear, due to the non-linearity of the function $ber(*)$ [6].

We solved the optimization problem in (4), characterizing the effects of the SNR and the available processing power on the maximum distance between BBU and RRH. We considered the same parameter set up described in the previous section, with small changes described as follows. A BBU decodes M code blocks, where each block m experiences different channel conditions, expressed in $\frac{Eb_m}{No}$ in the range $\{0.125, 0.25, 0.5, 1, 2, 4\}$ dB. Furthermore, we relied on the results of [6], which characterizes the FEC decoding performance in terms of BER achieved for a given $\frac{Eb_m}{No}$ and k recursions. In particular, we performed a linear interpolation on the graph in [6, Fig 10], yielding the function $ber(*)$. The interpolation considers the target BER b in the range of $[10^{-5}, 10^{0}]$ and sets the maximum number of recursions to 14, as in [6]. Finally, we used an exhaustive search algorithm to solve the maximization problem considering that d can vary in the set $D = \{1, .., 250\}$ km.

The results are depicted in Figure 2 for a target BER $b = 10^{-3}$. In the x-axis, p_m is presented in Hz, whereas the y-axis represents d in km. The curves correspond to different levels of SNR in dB. The available processing power p_m has a significant effect on the maximum distance, for an RRH experiencing low SNR ($\frac{Eb_m}{No} \leq 0.5$ dB). In this case, either significant resources must be dedicated to processing the signal or the BBU must be located nearer the RRH, in the fog rather than in the cloud. For high SNR ($\frac{Eb_m}{No} > 0.5$ dB), few iterations of the FEC algorithm suffice to achieve the target BER, and the effect of the processing power p_m is less signifcant. In both cases, the maximum distance eventually reaches a plateau, and the total latency is dominated by propagation delay.

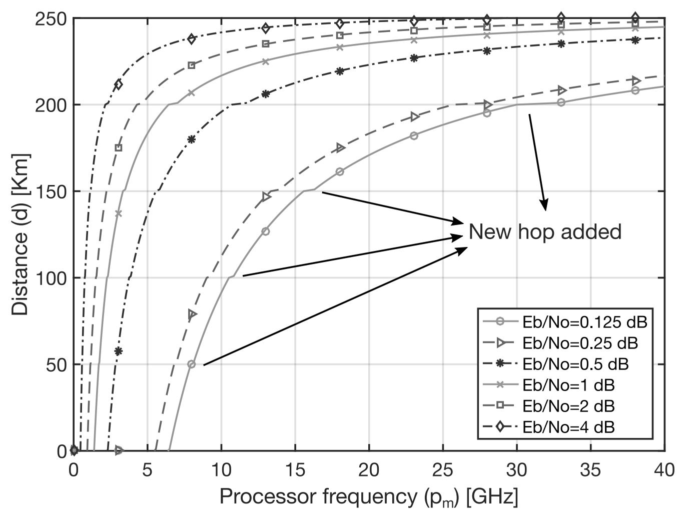

Figure 2: Maximum distance between an RRH and the BBU responsible for prrocessing its signals, as a function of available processing capabilities available at the BBU and the SNR experienced on the wireless channel.

It is also important to note that, in practice, the amount of resources that a BBU can allocate to processing signals coming from a given RRH may vary. For instance, a BBU can choose to decrease the processing allocated to RRHs experiencing high SNR, or located at shorter distances, in favor of those farther away, and/or experiencing low SNR. This decision can be made dynamically, fully exploiting the available processing power at the BBU.

We performed a second analysis considering different values for the target BER b and its influence when characterizing the relationship among the distance between BBU and RRH, the processing power p_m, and round-trip delay budget Φ. In this case, we considered the same set up of the previous analysis, except that the allocated processing power is fixed $p_m = 5$ GHz. The distance d is maximized for a target BER b in the range of $[10^{-5}, 10^0]$.

The results of this analysis are depicted in Figure 3. The y-axis shows the maximum distance d between BBU and RRH, and the x-axis presents the target BER b. From left to right, the curves in Figure 3 correspond to decreasing SNR observed at the RRH. As the target BER b increases (i.e., the error performance degrades), each step in the curves represents a decrease in the number of FEC recursions k_m required to process a code block m. The distance between BBU and RRH can reach approximately 92 km, in the case of high SNR ($\frac{Eb_m}{No} \geq 0.5\ dB$), even for a relatively ambitious target BER $b \leq 10^{-5}$, whereas an RRH facing low SNR ($\frac{Eb_m}{No} < 0.5\ dB$) cannot meet a challenging BER target ($b \leq 10^{-2.9}$).

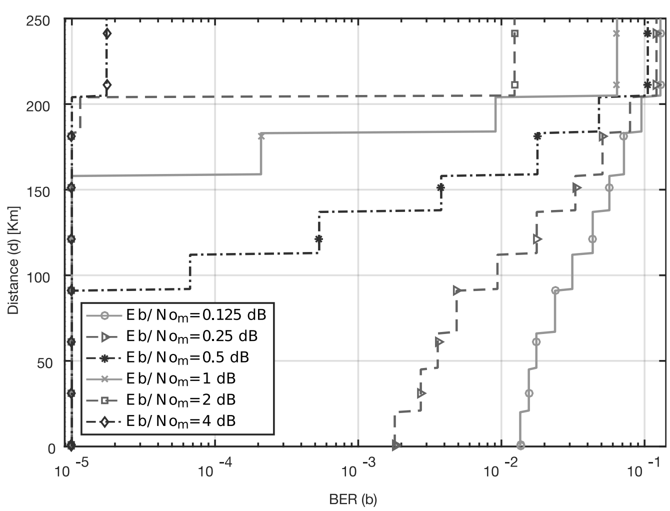

Figure 3: Maximum distance between an RRH and the BBU responsible for processing its signals, as a function of the target BER and the SNR experienced on the wireless channel.

The optimization problem we describe can be used to enhance the planning and allocation of processing resources in a C-RAN. During the planning of a C-RAN, the maximum distance between BBU and RRH and the required processing power can be determined together, considering a delay budget and target BER. These results can inform the dynamic allocation of processing resources in the network, and to determine which BBU, either in the fog or the cloud, is eligible to process the workload of different RRHs. These decisions can also account for load balancing considerations.

IV. Conclusions & Future Work

In this letter, we characterize the relationship between the distancce between an RRH and the BBU responsible for processing its signals and the round-trip delay in a C-RAN, which involves considerations of available processing power and channel conditions. We modeld the relationship among these factors as an optimization problem, considering the FEC decoding function as the most computationally intensive operation required at the BBU. Our results show that the processing power must be increased significantly for RRHs experiencing low SNR and/or at long distances from the BBU, to sustain a target bit error rate performance and to meet a round-trip delay budget. However, RRHs experiencing higher SNR can have their processing taking place in the cloud and still meet tight BER and latency constraints.

As the next step for this work, we will apply the tradeoff between processing power and distance between BBU and RRH to enable better load balancing in a C-RAN. One of the distributed BBUs in the pool can receive workload from multiple RRHs to be processed simultaneously, increasing the sharing of processing power as long as the round-trip delay budget is not violated.

Acknowledgments

This work is part of FUTEBOL, which has received funding from the European Union's Horizon 2020 for research, technological development, and demonstration under grant agreement no. 688941 (FUTEBOL), as well from the Brazilian Ministry of Science, Technology and Innovation (MCTI) through RNP and CTIC.

References

[1] China Mobile Research Institute, "C-RAN: The Road Towards Green RAN," *White Paper*, Version 2.5, 2011.

[2] F. Musumeci, C. Bellanzon, N. Carapellese, M. Tornatore, A. Pattavina, and S. Gosselin, "Optimal BBU Placement for 5G C-RAN Deployment Over WDM Aggregation Networks," *Journal of Lightwave Technology*, 2016.

[3] S. Bhaumik, S. P. Chandrabose, M. K. Jataprolu, G. Kumar, A. Muralidhar, P. Polakos, V. Srinivasan, and T. Woo, "CloudIQ: A Framework for Processing Base Stations in a Data Center," in *ACM Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2012.

[4] I. Alyafawi, E. Schiller, T. Braun, D. Dimitrova, A. Gomes, and N. Nikaein, "Critical Issues of Centralized and Cloudified LTE-FDD Radio Access Networks," in *IEEE International Conference on Communications (ICC)*, 2015.

[5] 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*, 2014.

[6] M. F. Brejza, L. Li, R. G. Maunder, B. M. Al-Hashimi, C. Berrou, and L. Hanzo, "20 Years of Turbo Coding and Energy-Aware Design Guidelines for Energy-Constrained Wireless Applications," *IEEE Communications Surveys & Tutorials*, 2016.

[7] Y.-j. Ku, D.-y. Lin, C.-f. Lee, P.-j. Hsieh, H.-y. Wei, C.-t. Chou, and A.-c. Pang, "5G Radio Access Network Design with the Fog Paradigm: Confluence of Communications and Computing," *IEEE Communications Magazine*, 2017.

[8] 3rd Generation Partnership Project (3GPP), "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and Channel Coding, (3GPP TS 36.2 .212 version 12.6.0 Release 12)," *Technical Specification*, 2015.

[9] H. Holma and A. Toskala, *LTE for UMTS-OFDMA and SC-FDMA based Radio Access*. John Wiley & Sons, 2009.

[10] R. d. R. Righi, V. F. Rodrigues, C. A. da Costa, G. Galante, L. C. E. de Bona, and T. Ferreto, "AutoElastic: Automatic Resource Elasticity for High Performance Applications in the Cloud," *IEEE Transactions on Cloud Computing*, 2016.