Performance Analysis of Fog-Based Radio Access Networks
Heba M. Abdel-Atty, Raad S. Alhumaima, Sherif M. Abuelenin, Eman A. Anowr

Abstract—For the new generation of cellular networks, new architectures are being examined to improve performance. Fog communication is a new paradigm that has been presented to exploit the power of edge computing in regards to throughput and latency. In this paper, we introduce a mathematical model to examine the performance of such proposals. Specifically, the delay, power consumption (PC) and energy efficiency (EE) of fog radio access networks (F-RANs) are tested in comparison with those of traditional cloud radio access networks (C-RANs). Although the literature has shown that F-RANs provide enhanced delay performance, this paper shows that they also consume a large amount of power, degrading their EE in comparison with their traditional cloud counterparts. However, the level of degradation depends on the number of deployed fog devices, which directly influence the PC. This paper also shows that improving the delay performance by using a fog architecture is not a straightforward process but rather requires particular care in terms of choosing the appropriate mode when placing/installing fog functions in fog devices.

Index Terms—Fog communications, Baseband unit pool, Cloud radio access networks, Energy efficiency, Power consumption

I. INTRODUCTION

It is expected that by 2020, the number of mobile base stations will have increased by millions, and the number of user equipment devices (UEs) will have increased by billions. Such an increase will surely lead to enormous consumption of power, high data rate demands, and high network costs [1]. Fifth-generation (5G) wireless communication systems are intended to serve an increasing number of network users (UEs) while consuming less power, achieving higher data rates, and incurring less delay and cost, that is, increasing the energy efficiency (EE) by a factor of 10, increasing the capacity by a factor of 1000, and decreasing the latency by approximately a factor of 5 [2] in comparison with 4G networks. Hence, 5G technology is trending towards providing the required data to UEs while using fewer network resources. The EE can be increased either by using optimisation techniques [3], resource allocation techniques [4], pre-coding techniques [5] or by using power reduction models [6]. Meanwhile, the network capacity can be increased by increasing the number of cells via a complex, heterogeneous network structure with small cells [7] or by using techniques such as the multiple-input multiple-output (MIMO) approach, in which numerous antennas simultaneously serve a number of UEs using the same time-frequency resources [8]. However, the power consumption (PC) problem has not yet been solved; for this reason, the cloud radio access network (C-RAN) structure has evolved. C-RANs represent a computing infrastructure that enables ubiquitous access to shared pools of configurable resources (e.g., computer networks, servers, storage, applications, and services) [9]. C-RAN architectures have been suggested by both operators (e.g., NTT, KT, France-Telecom/Orange, Telefonica, SoftBank/Sprint and China Mobile) and equipment vendors (e.g., Alcatel-Lucent’s lightRadio and Nokia-Siemens’ Liquid Radio). The base station (BS) is separated into a radio unit, called the remote radio head (RRH), and a processing unit, called the baseband unit (BBU), and the link between them is called the fronthaul (optical or wireless connection). The RRH provides the interface to the fibre and performs digital processing, digital-to-analogue conversion, analogue-to-digital conversion, power amplification, and filtering. On the other hand, the BBU is responsible for processing baseband signals and optimising radio resource allocation. Based on the aggregation of a number of BBUs in a centralised data processing centre called a BBU pool, an innovative and stand-alone architecture has evolved to reduce the cooling PC and, in turn, the total PC. C-RAN architectures also offer reduced capital and operational costs and reduced site rents and can serve as the basis for shared platforms, which are necessary for launching new services and applications in a timely fashion [10]. Nevertheless, the C-RAN approach also has some limitations, such as the heavy burden that is placed on the fronthauls due to combining the processing units into a unified structure. Consequently, the latency is increased in comparison to that in traditional architectures. However, decreasing the network delay is not a straightforward process, as the network delay is constrained by the physical link delay. Hence, alleviating this burden will require a new paradigm that takes advantage of edge device techniques, such as edge computing, also called the fog radio access network (F-RAN) paradigm [11] [12]. In an F-RAN architecture, some of the responsibilities for caching, storage, data processing, application execution, and computing and some of the functions of the BBU pool, such as collaboration radio signal processing (CRSP) and cooperative radio resource management (CRRM), are shifted to the end devices (RRHs and end UEs). In a fog network, a traditional RRH is called a fog device (FD), and a traditional UE is called a fog computing user equipment device (F-UE). However, we are interested only in the network side, i.e., the BBU pool and FDs. Although this new paradigm has been well studied in the literature in regards to improvements in network delay, its various advantages and disadvantages have not been fully evaluated, especially in terms of PC. However, the PC of an F-RAN architecture will strongly affect the network EE, and hence, the related parameters require investigation. By modelling the PC, EE and delay, a general platform can be developed to evaluate the network techniques proposed.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.
by other researchers. Furthermore, such a platform would offer a proof of concept of this new network paradigm in a simplified way. Therefore, in an attempt to find a solution to the previous basic limitations encountered by traditional networks and C-RANs, this paper introduces a new power model (PM) for computing the PC and EE for F-RANs, evaluating their latency, and comparing the results with those of C-RANs.

A. Related Work

Several studies have offered methods of evaluating the PC of an entire network, such as [13], which presents a PM that can be applied to long-term evolution (LTE) technology to evaluate all types of BSs. This work completely breaks down the structures of simplified traditional BSs. In [14], a PM is introduced to evaluate the fundamental blocks of a BS site and is used as the basis of a high-level PC model for various types of LTE BSs. However, this model is mostly designed to enable the exploration of different state-of-the-art networks and parameters, such as the installation of more antennas, the trade-offs between macro and small cells, the impact of load and the role of duty cycling. A set of state-of-the-art (SoA) models [15] has also been proposed for C-RAN PC calculations. This research provides useful PMs for various components that are parametrised with respect to different operational parameters and varying vendor configurations. These models are simplified and can be applied to any new cloud communication system for comparison with traditional BSs. This work also enables PC comparisons with traditional networks. In a later paper [16], a PM is proposed for software-defined networking-based C-RANs and is compared to the C-RAN model. Another PM has been proposed in [17] that includes PC calculations for network function virtualisation within the BBU pool, for comparison with the non-virtualised C-RAN model. In [18], a PM is introduced for calculating the PC of a heterogeneous C-RAN (H-CRAN) by adding the modifications to the network architecture to the PM equations. In addition to the PC, several works have discussed the latency of fog networks. For example, the authors of [19] have analysed the delay of fog networks, considering multiple fog nodes performing different computing tasks. However, the intuitive gain in latency is compromised by the increase in the PC incurred when the fog functions are abstracted. A similar work can be found in [20]. The overall performance of fog networks was investigated in [21] with regard to spectral efficiency and EE. Unfortunately, no mathematical model was presented to describe how the optimisation parameters influence each other. Finally, an offloading technique has been proposed in [22], in which the computation power is split into fog and cloud components to form a hybrid-oriented fog network design.

B. Main Contributions

Most researchers have neglected to calculate the PC. However, such calculations are very important for determining which parameters represent the largest contributions to the PC of a network and, accordingly, can provide guidance for research attempting to reduce such consumption without affecting network performance. The contributions of this paper are as follows:

1) Proposing the PM components that are necessary for calculating the PC associated with each unit of an F-RAN system. Because deploying a new paradigm is expensive, this work permits network vendors and operators to evaluate their gains in terms of PC and latency compared with traditional C-RANs prior to implementation.

2) Wrapping the complexity of the PM components into a simplified parametrised model, in which the sophistication of the representation of each unit’s PC is considerably reduced. Because only a few parameters in this parametrised model vary, while the others remain static, this PM can be used to measure the power savings achieved with an F-RAN architecture in a simplified way, from perspectives such as the effect of the number of antennas on the PC as well as the effects of reducing the transmission power, deactivating the antennas, modifying the bandwidth and implementing a sleep mode.

3) Comparing the PC, EE and latency performance of a C-RAN architecture with that of an F-RAN architecture. We consider various network characteristics in terms of the numbers of BBUs, RRHs, FDs, resource blocks (RBs) and antennas. Thus, we develop a large-scale, long-term model for holistic system analysis and evaluation.

II. FOG RADIO ACCESS NETWORK (F-RAN) POWER MODEL (PM)

There are two main contributors to the PM: the BBU pool and the RRHs (called the FDs in the fog paradigm). Fig. 1 shows the changes in structural design from the C-RAN architecture to the F-RAN architecture. However, the differences between these architectures are not limited to their structural design but also include the functions that each architecture can perform. Some functions that are processed in the BBU pool in the C-RAN architecture, such as caching, CRRM, and CRSP, will be hosted by the FDs in the F-RAN architecture, and the smart edge devices in an F-RAN system also support some new technologies, such as device-to-device communication, software-defined networking, network function virtualisation, and Internet of Things functionalities.
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2930666, IEEE Access

cooling systems. In this section, we will discuss the PC for each system separately, along with the total PC.

A. Base band unit pool (BBU pool)

Mainly, the BBU pool consists of a number of BBUs in the form of a cluster. The total PC of the BBU pool is the sum of the PCs of all units, as follows:

1) Base band unit: Each BBU performs the necessary digital signal processing before passing signals on to an RF transceiver and then to a PA. The digital computations and processing of a BBU can be measured in giga operations per second (GOPS) and translated into power figures. This can be achieved by multiplying the GOPS value by the corresponding technology scaling factor to obtain the number of operations performed per second per watt. A set of BBU functions, such as forward error correction, time- and frequency-domain processing, and central processing, can be linked to the GOPS value. A PM for complex components was described in [14]. However, this model is very complex and time consuming to calculate. Hence, we have developed a simpler model to describe the PC of a BBU based on the number of processed RBs. We assume that the PC of the BBU ($P_{BBU}$) increases linearly from a certain initial value ($P_{intBBU}$) as the number of RBs ($N_{RB}$) increases. This assumption translates into the following expression:

$$\frac{dP_{BBU}}{dN_{RB}} = \gamma P_{BBU}$$

(1)

The increase in $P_{BBU}$ is linearly proportional to $N_{RB}$, with a constant factor ($\gamma$). Solving this equation yields:

$$P_{BBU}(N_{RB}) = P_{intBBU} e^{\gamma N_{RB}}$$

(2)

2) DC-DC conversion for the BBU pool: Each BBU in the pool requires a certain amount of voltage to operate. Hence, there is a need for a DC-DC converter. Since the efficiency of this converter is less than unity [13], it can be expressed as a loss. This loss scales linearly with the PC of each component that requires DC-DC conversion. Accordingly, the DC-DC converter’s PC ($P_{DC,p}$) is given as follows:

$$P_{DC,p} = \sum_{b=1}^{B} \sum_{f=1}^{F} l_{DC,p}(\eta_{DC,p})[P_{bbub,f} + P_{opt,p,f}]$$

(3)

where $l_{DC,p}$ is the loss caused by DC-DC conversion as a function of the DC conversion efficiency ($\eta_{DC,p}$), $b \in \{1, ..., B\}$ is an index representing the active BBUs in the BBU pool, $f \in \{1, ..., F\}$ is an index representing the active FDs, and $P_{opt,p,f}$ denotes the PC for the optical device in the pool that connects the $b$-th BBU to the $f$-th FD.

3) AC-DC conversion: Before power is passed to the DC-DC converter, an AC-DC voltage converter is required to convert the AC mains voltage to a DC voltage. The efficiency of this conversion can also be expressed as a loss. Hence, the corresponding PC ($P_{AC,p}$) scales linearly with the PC of each unit that requires AC-DC conversion, as follows:

$$P_{AC,p} = P_{DC,p} + \sum_{b=1}^{B} \sum_{f=1}^{F} l_{AC,p}(\eta_{AC,p})[P_{bbub,f} + P_{opt,p,f}]$$

(4)

where $l_{AC,p}$ is the measured loss of the AC power converter as a function of the AC conversion efficiency ($\eta_{AC,p}$).

4) Cooling PC: The cooling system is a very important subcomponent, which, along with the PAs, consumes most of the power in the network. The cooling PC ($P_{cool}$) is modelled in the same way as the PCs for AC-DC and DC-DC conversion; it scales linearly with the number of devices or components that require cooling, as follows:

$$P_{cool} = l_{cool}(P_{AC,p} + P_{DC,p} + \sum_{b=1}^{B} \sum_{f=1}^{F} P_{bbub,f} + P_{opt,p,f})$$

(5)

B. Fog Devices (FDs)

An FD has the same basic components as a traditional RRH, i.e., an RF transceiver, a PA, antennas and an optical component, but no cooling system. As mentioned in [13], components with a PC of less than 500 W do not require a cooling system. Thus, the PC of an FD is a combination of the PCs of the following subcomponents:

1) Power amplifier PA: The PA is usually used to amplify the received signals from the RF transceiver before the antennas transmit these signals to the UEs, and vice versa. It is known that the PA consumes a large amount of power due to its low efficiency because its best operation point is located near the saturation level. However, it is forced to behave more linearly because of the high fluctuations existing in orthogonal frequency division multiplexing signals. Hence, a higher back-off, resulting in a weaker efficiency ($\eta_{PA}$), leads to an increase in the PC. The PC of the PA is denoted by ($P_{PA}$)in the PM. There are three main parameters that participate in its calculation: the output transmission power ($P_{out}$) of the antennas, the output power of the PA ($P_{TX}$) and the bandwidth ($BW$). The full PC description is given in [15] as $P_{PA} = \frac{P_{TX}}{\eta_{PA}}$, where $P_{TX} = P_{out}BW$.

2) Radio frequency (RF) transceiver: This unit is responsible for several functions in the FD, such as modulation, demodulation and analogue-to-digital conversion. Each antenna requires one RF unit and one PA unit to transmit signals. A complex model has been presented in [14], in which the PC of an RF transceiver depends on the number of antennas and the bandwidth. However, we require a simplified model that depends on a single, well-known parameter. Hence, we express the PC of an RF transceiver in terms of RB processing because the RF transceiver PC ($P_{RF}$) varies linearly with the number of processed RBs, as follows:

$$\frac{dP_{RF}}{dN_{RB}} = \xi P_{RF}$$

(6)
where \( \xi \) is a constant that describes the increase in power. Solving this equation yields:

\[
P_{RF}(N_{RB}) = P_{intRF} e^{\xi N_{RB}}
\]

where \( P_{intRF} \) denotes the initial PC of the RF unit.

3) Power Conversion: As discussed earlier, FDs also consume power for DC-DC (\( P_{DC,F} \)) and AC-DC (\( P_{AC,F} \)) CRRM function, respectively. The corresponding PC for DC-DC conversion varies linearly with the PC of other units in the FD, as follows:

\[
P_{DC,F} = l_{DC,F}(\eta_{DC,F})[P_{opt,F} + \sum_{b=1}^{B} \sum_{f=1}^{F} \sum_{t=1}^{T} (P_{PA})^f_t] + P_{RF}^t_f
\]

where, \( P_{RF}^t_f \) and \( P_{PA}^t_f \) denote the PCs (in W) of the \( t \)-th RF and PA, respectively, in the \( f \)-th FD connected to the \( b \)-th BBU and \( P_{opt,F} \) (in W) is the PC of the optical device in the FD. The AC-DC’s PC is given as:

\[
P_{AC,F} = l_{AC,F}(\eta_{AC,F})[P_{DC,F} + P_{opt,F} + \sum_{b=1}^{B} \sum_{f=1}^{F} \sum_{t=1}^{T} (P_{RF} + P_{PA})^f_t]
\]

where \( l_{AC,F} \) refers to the AC losses as a function of conversion efficiency (\( \eta_{AC,F} \)) and \( T \) is the total number of antennas.

C. Collaboration Radio Signal Processing (CRSP) and Cooperative Radio Resource Management (CRRM) Power Consumption

F-RAN PC modelling additionally requires modelling of the PC that is inherited from the functions shifted from the BBU pool to be hosted by the FDs. As a result, the FDs will incur some additional PC due to the installation of the new functions, i.e., CRSP and CRRM, which are necessary for the UEs’ assignment and authentication processes. On the other hand, the PCs for these functions will be deduced from the PC of the BBUs. However, the effects of these functions on the PC are not equal on both sides of the network, as the number of active FDs is usually larger than the number of active BBUs.

To model this situation, two practically implementable methods of shifting the functions to the FDs should be considered. First, these functions can be virtualised within the RF unit; second, the functions can be installed on a separate electronic chip, which can then be added to the FD. If the first scenario is assumed, the virtualised fog functions might increase the PC of the RF transceiver either linearly or exponentially. The increase in the CRSP PC with an increasing number of RBs is given by \( dP_{crsp}/dN_{RB} = \kappa P_{crsp} \); solving this equation yields \( P_{crsp} = P_{intcrsp} e^{\kappa N_{RB}} \), where \( P_{intcrsp} \) is the initial/static PC for CRSP when no RBs are processed. The same approach can be used to model the CRRM function: \( P_{crrm} = P_{intcrrm} e^{\lambda N_{RB}} \), where \( P_{intcrrm} \) denotes the initial PC for CRRM function, \( \kappa \) and \( \lambda \) are the constants describing the increases in the PC for CRSP and CRRM, respectively. The reason why an exponential expression is used in this modelling approach is that as the constant(\( \kappa \) or \( \lambda \)) approaches 0, the model tends towards linearity; hence, this expression simultaneously captures both linear and exponential behaviour. In addition, this expression permits the inclusion of a static level of PC, represented by \( (P_{intcrsp} + P_{intcrrm}) \) to describe the usage of power when no load is being processed. Hence, these PC values will be added to the PC of the FD, and thus, the PM of an FD in the first scenario is generally given as follows:

\[
P_{FD}^1 = P_{RF} + P_{crsp} + P_{crrm} + P_{PA} + P_{opt,F}.
\]

The modelling of an FD in the second scenario is achieved by incorporating the additional PC represented by the electronic chip. This additional PC is denoted by \( (P_{add}) \) and depends on the type of chip used; for example, an FPGA consumes approximately 30 W. However, this consumption depends on traffic, performance, and operating frequency. System-on-a-chip boards consume much less power but are much more expensive than FPGA boards. Nevertheless, we can generally assume that for any type of chip, the PC is initially \( P_{add} \) and is driven up to the maximum with an increase in the number of processed RBs. Hence, the PC can be expressed as \( P_{add} = P_{intadd} e^{\xi N_{RB}} \), where \( \xi \) is the constant describing the power increase. Finally, the PC of an FD of the second type can be generally formulated as

\[
P_{FD}^2 = (P_{RF} + P_{add} + P_{PA} + P_{opt,F}).
\]

On the other side of the network, the PC for these functions is deduced from the PC of the BBU pool:

\[
P_{BBU,f} = P_{BBU} - (P_{crsp} + P_{crrm})
\]

Subsequently, the total PC of the fog network is calculated by summing the PCs of both the BBU pool and the FDs, as follows:

\[
P_{fog}^{1,2} = P_{pool} + P_{FD}^{1,2}
\]

where

\[
P_{pool} = P_{cool} + P_{AC,P} + P_{DC,P} + \sum_{b=1}^{B} \sum_{f=1}^{F} \sum_{t=1}^{T} (P_{BBU,b,f,t} - P_{crsp,b,f,t} - P_{crrm,b,f,t} + P_{opt,bb,f})
\]

and first scenario of FD PC is given by:

\[
P_{FD,t}^1 = P_{AC,F} + P_{DC,F} + \sum_{b=1}^{B} \sum_{f=1}^{F} \sum_{t=1}^{T} ((P_{PA} + P_{RF})^b_t + P_{crsp,b,f,t} + P_{crrm,b,f,t} + P_{opt,b,f})
\]

where the second scenario of PC is given by:

\[
P_{FD,t}^2 = P_{AC,F} + P_{DC,F} + \sum_{b=1}^{B} \sum_{f=1}^{F} \sum_{t=1}^{T} ((P_{PA} + P_{RF})^b_t + P_{add} + P_{opt,b,f})
\]
where $P_{PA}^b, P_{RF}^b$ (in W) denote the PC of the PA and RF, respectively, of $f$-th antenna served by $b$-th BBU. $P_{BBU,f,t}^b$ (in W) denotes the PC of $b$-th BBU connected to $f$-th FD mounted on the $t$-th antenna.

The above PC description can be further simplified by representing the overhead consumption in the form of loss factors, i.e., $\sigma_{DC}^F, \sigma_{AC}^F$, and $\sigma_{cool}^f$ representing the DC-DC, AC-DC and cooling loss factors, respectively, corresponding to PCs that are linearly proportional to the PC values of other components. These factors will replace the previous PC representations for cooling, AC-DC conversion and DC-DC conversion. Hence, the total PC of the fog network in the first scenario can be expressed as:

$$P_{fog}^1 = P_{pool} + P_{FD,t}^1 \left( P_{BBU,f}^b + P_{opt,p} \right) \frac{1}{(1 - \sigma_{DC,F})} \frac{1}{(1 - \sigma_{AC,F})} \frac{1}{(1 - \sigma_{cool})} + \frac{1}{(1 - \sigma_{DC,F})} \frac{1}{(1 - \sigma_{AC,F})}$$

whereas the consumption of second scenario can be given as:

$$P_{fog}^2 = P_{pool} + P_{FD,t}^2 \left( P_{BBU,f}^b + P_{opt,p} \right) \frac{1}{(1 - \sigma_{DC,F})} \frac{1}{(1 - \sigma_{AC,F})} \frac{1}{(1 - \sigma_{cool})} + \frac{1}{(1 - \sigma_{DC,F})} \frac{1}{(1 - \sigma_{AC,F})}$$

III. FOG NETWORK EE ANALYSIS

The network EE is a more important measure than the bare capacity or spectral efficiency because the EE serves as a power indicator and thus provides an additional dimension to assess the network performance. To evaluate the EE, we consider a fog network that contains a total of $(F)$ FDs and a total of $(U)$ UEs. The small scale fading between FD ($f$) and the UE $u$ is denoted by $H_{f,u}$, and it is assumed Rayleigh fading. The power received by the UE $u$ from FD $f$ is expressed as $P_{r,f,u}^t = P_{t,f,u} H_{f,u} r_{f,u}$, where $P_{t,f,u}$ denotes the power transmitted from $f$-th FD to UE $u$, $r_{f,u} = d_{f,u}^{-\alpha}$ denotes the path loss between FD $f$ and UE $u$, with $\alpha$ being the path loss exponent; and $H_{f,u}$ represents the channel gain from the $f$-th FD to $u$-th UE. Furthermore, $d_{f,u}$ is the straight line distance between $f$-th FD and $u$-th UE, which is given by $d_{f,u} = \sqrt{(x_f - x_u)^2 + (y_f - y_u)^2}$, where $x_f, y_f, x_u, y_u$ indicate the Cartesian $x$ and $y$ coordinates of the FD and UE, respectively.

Subsequently, the EE of the fog network can be directly calculated as follows: $EE = \frac{BW \log_2(1 + P_{r,f,u} H_{f,u})}{P_{fog}^n}$. The fog network EE in the first scenario can thus be modelled as

$$EE^1 = \frac{BW \log_2(1 + P_{r,f,u} H_{f,u})}{\sum_b \sum_f P_{fog}^1}$$

while the EE in the second scenario is modelled as

$$EE^2 = \frac{BW \log_2(1 + P_{r,f,u} H_{f,u})}{\sum_b \sum_f P_{fog}^2}$$

where $\Gamma_{f,u} = b_{f,u} r_{f,u}$ denotes the signal to noise (SNR) ratio and $S$ is the total number of FDBs implemented in accordance with the second scenario. This formulation can be used to represent both cloud and fog implementations, as it includes parameters such as the number of traditional BBUs ($B$) and the number of RRHs or FDs ($F$) as well as the corresponding differences in their PC modelling.

IV. LATENCY ANALYSIS

It is assumed that the delay from an FD to the BBU pool is $\tau_{f,o} = d_{f,o}/c$, where $(o)$ denotes the origin of the pool’s geographical position and $c$ is the speed of light in the case of wireless links between the BBU pool and the FDs. In the case of optical fiber links, $\tau_{f,o} = d_{f,o}/c_{opt}$, where $(c_{opt} = c/ind)$ is the speed of light in an optical fiber with refractive index $(ind)$. In turn, the delay from the UE to an FD is expressed as $\tau_{f,u} = d_{f,u}/c$; such link can only be wireless. Clearly, if a UE is served by an FD rather than by the BBU pool, the delay gain will be equivalent to $(\tau_{f,o})$ because the UEs are no longer connected to the BBU pool at all times but rather are connected to the FDs. However, once the service type and authentication for a UE have been established, the UE can return to the BBU pool for persistent resource allocation. Nevertheless, when fog functions are shifted from the BBU pool to the FDs, an enhanced processing delay is incurred in the latter case due to virtualisation. In [23], it is mentioned that the execution/processing time is linearly proportional to the number of processed RBs and also depends on the modulation coding scheme (MCS) that is used to transmit these RBs. Due to virtualisation, this time can grow exponentially, as one virtual machine (VM) might take 5 times as many cycles to process a packet than a bare-metal device would [24]. Therefore, a model that can combine both traditional and virtualisation-based concepts is required. Let $(\tau_{device})$ denote the execution time of a device without virtualisation, where $\tau_{device} = \tau_{int} + (mod_{RB})$; here, $\tau_{int}$ represents the initial device delay due to functions other than MCS, whereas the MSC delay is represented by the constant factor $(mod)$. Due to virtualisation costs, $\tau_{device}$ is increased by a factor of $(e^{1/2N_{RB}})$ (presumed to be 1.7 in the case of 2 VMs), i.e., $\tau_{device} = \tau_{device}e^{1/2N_{RB}}$, where $\Omega$ is the constant describing the exponential increase. Hence, the total delay of all functions can be formulated as $\tau = \sum f \tau_{device,f}$, where $\tau_{device,f}$ denotes the execution time of the FD $f$. Notably, this model is valid only for FDs implemented following the first scenario, in which the functions are installed on the same board as the RF unit. If the fog functions are installed on a separate board/device, the formulation will instead be $\tau = \sum f \tau_{device,f} + \Omega_{brd}$, where $\Omega_{brd}$ denotes the delay of the separate board, which also includes a traffic-based delay (i.e., $\Omega_{brd} = \tau_{int} + e^{\omega N_{RB}}$), where $\omega$ is the constant describing the exponential increase in the execution time.

On the other hand, the decrease in the delay from the BBU pool, due to the shifting of the functions to the FDs, is represented by $(\tau_{drop} = q_{N_{RB}})$. Hence, the total delay in the first scenario is equal to $\tau^1 = \tau^1 - \tau_{drop}$. This logic also holds true for the second scenario, for which the total delay is given by $\tau^2 = \tau^2 - \tau_{drop}$.
V. RESULTS AND ANALYSIS

The proposed model depends on many variables, such as the numbers of antennas, FDs and BBUs. The values of these parameters that were used to produce the final results are shown in Table I. To ensure fair comparisons, the AC-DC and DC-DC loss factors are assumed to be identical in all cases.

A cooling PC comparison of the C-RAN and fog architectures for the first FD scenario is shown in Fig. 2 for different numbers of BBUs and different numbers of antennas. Between the C-RAN architecture and the fog architecture, the cooling PC is reduced in the latter due to the shifting of necessary functions from the BBU pool. In the second scenario, however, each FD consumes additional power due to the separate chip, causing the PC to increase relative to the first scenario, in which the fog functions are virtualised; in addition, there are AC-DC and DC-DC overheads for the separate chips that affect each FD individually. Hence, the cooling PC of the fog network in the second scenario is expected to be higher than that in the first scenario, as shown in Fig.3.

We assume that in the first scenario, the CRRM and CRSP functions increase the RF unit PC by 3 W for the processing of 100 RBs. Hence, the PC of the RF unit increases from 12.9 to approximately 15.9. The PC of the BBU pool, where these functions previously resided, is reduced by the same amount. By contrast, we assume that the PC in the second scenario is slightly greater than that in the first scenario, with the RF unit PC increasing from 12.9 to 17. Because the number of FDs is usually larger than the number of BBUs, the total amount of power consumed will increase. Fig. 4 shows the total PC comparison in the first scenario for different numbers of FDs, RRHs and BBUs. If the numbers of active BBUs and FDs are equal, the PC performance of the fog network will be identical to that of the corresponding C-RAN.

Due to the degradation in the PC performance, the EE performance will also be affected. Fig. 5 shows an EE comparison of C-RAN and fog architectures with 40 RRHs/FDs and approximately 350 UEs when processing 100 RBs.

TABLE I: Model Parameters.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Traditional (Generated)</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{DC}$</td>
<td>0.9250</td>
<td>-</td>
</tr>
<tr>
<td>$\sigma_{AC}$</td>
<td>0.910</td>
<td>-</td>
</tr>
<tr>
<td>$\sigma_{cool}$</td>
<td>0.90</td>
<td>-</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.002</td>
<td>-</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.002</td>
<td>-</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>0.001</td>
<td>-</td>
</tr>
<tr>
<td>$\mu_{BBU}$</td>
<td>0.014</td>
<td>-</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.01</td>
<td>-</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.0001</td>
<td>-</td>
</tr>
<tr>
<td>$\mu_{ind}$</td>
<td>1.3</td>
<td>-</td>
</tr>
<tr>
<td>$\Omega$</td>
<td>0.02</td>
<td>-</td>
</tr>
<tr>
<td>$\psi$</td>
<td>0.002</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>$t_{int}$</td>
<td>50 $\mu$sec</td>
<td>-</td>
</tr>
<tr>
<td>$P_{opt,P}$</td>
<td>20 $\mu$sec</td>
<td>-</td>
</tr>
<tr>
<td>$P_{RF}$</td>
<td>1 W</td>
<td>-</td>
</tr>
<tr>
<td>$P_{BBU}$</td>
<td>1 W</td>
<td>-</td>
</tr>
<tr>
<td>$P_{int}$</td>
<td>12.9 W</td>
<td>-</td>
</tr>
<tr>
<td>$P_{F}$</td>
<td>29.6528 W</td>
<td>-</td>
</tr>
<tr>
<td>$P_{ADD}$</td>
<td>29.4 W</td>
<td>-</td>
</tr>
<tr>
<td>$P_{add}$</td>
<td>8 W</td>
<td>-</td>
</tr>
</tbody>
</table>
total EE degradation is due to the increase in the PC in both scenarios when the numbers of BBUs and FDs are not equal. Each BBU is assumed to serve 2 FDs, which means that the number of BBUs is 20.

![Fig. 5: EE comparison of C-RAN and fog architectures.](image)

Moreover, to analyse the delay, the delay that is inherited from the MCS process is first calculated. It is assumed that in the first scenario, the virtualisation of the functions increases this delay by a factor of 1.7 by virtue of smart VM hypervisors (HVs). The latter are responsible for managing the fog functions/VMs in a timely manner and ensuring that the VMs do not disrupt each other while sharing the device’s resources. Hence, the more advanced the HV is, the less time each VM will spend processing its packets compared to the time required by a bare-metal device. An efficient HV can eliminate the additional time delay incurred by the device due to virtualisation, which degrades the network performance. Therefore, many studies, such as [25], have addressed the issue of reducing such costs. The RB-based FD delay starts at $\mu$sec and reaches approximately 200 $\mu$sec. The virtualisation process increases the initial FD delay by a factor of 1.7 when two VMs are running at the same time. However, these values can change depending on the different data acquired from different devices, and the model can accommodate such variations. In the second scenario, the initial delay for the separate chip is added to the initial RF unit delay to obtain the total delay within the same FD. Notably, the separate chip produces less delay when processing the RBs because the delay cost of virtualisation is higher than the delay incurred when the fog functions are installed on a bare chip. Fig. 6 shows the total delays generated by combining the channel delay and the abovementioned processing delays for various RRH/FD distances when processing 100 RBs. Notably, the link delay of a fog network consists of only the average delay between the UEs and the FDs, while in a traditional cloud network, the channel delay comprises both the delay between the UEs and the RRHs and the delay between the RRHs and the BBU pool. Thus, the fog architecture ultimately has an advantage against the legacy C-RAN architecture.

![Fig. 6: Total delay comparison of the first and second fog network scenarios with the C-RAN architecture for various RRHs/FDs distances.](image)

VI. CONCLUSION

This paper has discussed 5G fog networks and compared them with traditional C-RANs in terms of PC, EE and latency. A modelling scheme for the corresponding network parameters has been presented. Two scenarios are considered with respect to the fog architecture: in the first, the fog functions are virtualised within the RF units of the FDs, and in the second, the fog functions reside on a separate board within each FD. Although the presented model was developed based on specific data, it can be generalised to represent different devices with various manufacturing properties and specifications. This research has shown that the additional cost incurred by a fog network in terms of PC is noteworthy, yet ultimately insignificant. This PC cost does cause the EE of such fog-based designs to degrade, even as they achieve reduced delays. The latter is the paramount factor to consider when a network must provide real-time communications services. However, this research suggests that a fog network is not an appropriate choice for off-line applications because of the additional PC cost. Hence, an adaptable, hybrid design combining both the legacy cloud approach and the fog approach, in which a decision regarding which approach to implement can be made prior to serving a UE, is worthy of consideration. Such a decision can be made based on whether the UE is requesting on-line or off-line services. Accordingly, the UE’s packets will be directed to the BBU pool for an on-line request or to the FDs for an on-line request. In this case, some of the power cost will be eliminated, as the fog network can shift into the classical C-RAN mode, with a lower PC cost, during certain periods of time.
REFERENCES


DR. Heba M. Abdel-Atty received the Ph.D. degree in Electronics and Communications Engineering from the Faculty of Engineering-Port Said University, Egypt 2012. She is working as an associate professor at the Electronics and Communications Engineering Department from 2018 until now. She is founder and counselor of IEEE in Port-Said University student branch. She is a member of IoT Egypt Forum She published a lot of papers in international journals and conferences and she is a reviewer at a lot of international journals. Her current research interests are in wireless communication systems, multimedia processing, cognitive radio, security of wireless networks, artificial intelligence(AI), Internet of Things(IoT), Smart cities technologies, Sustainable development, and Field Programmable Gate Array.

Raad S. Alhumaima received his B.E. degree in Communications Engineering from the University of Diyala in 2003, his M.Sc. degree was in Laser/ Electronics and Communication from the Institute of Laser for Postgraduate Studies, University of Baghdad in 2011, and his Ph.D. in Communications and Electronics from Brunel University London-UK, in 2017. He has published several IET, IEEE and conference papers. He is currently a member stuff in the University of Diyala, College of Engineering, Department of Communications, Iraq. His current research interests cover next generation systems and networks, including SDN, C-RAN, NFV, and Fog-RAN. Furthermore, new research interests have been forwarded towards adapting quantum and classical communications.
Sherif M. Abuelenin, received the B.Sc. degree in electronics and communications from Suez Canal University, Egypt in 1999, the M.Sc. degree in Electrical Engineering from Tuskegee University in 2002, and the Ph.D. degree from Auburn University in 2005. He joined Tuskegee University as an assistant professor of Electrical Engineering in 2005. From 2007 to 2011 he served as an assistant professor in the faculty of Engineering Sciences, Sinai University, Egypt. In 2011, he joined the faculty of engineering, Port-Said University, where he currently is an Associate Professor of Electrical engineering. His research interests include signal processing for communications, vehicular communications, and fuzzy logic systems.

Eman A. Anowr received the B.S. degree in Electronics and Communications from Port-said University, Egypt in 2010. Her interests were about cisco academy (routing, switching and security of the network) and Microsoft professional and solution association (windows server). Her new interests cover the new generations of wireless communications, virtualizations, clouding and fog computing.