Energy Efficiency of 5G Mobile Networks in Hybrid Fog and Cloud Computing Environment

Stojan Kitanov
Mother Teresa University
School of Informatics
Skopje,
Republic of Macedonia
stojankitanov@hotmail.com

Toni Janevski
Ss Cyril and Methodius University
Faculty of Electrical Engineering
and Information Technologies,
Skopje, Republic of Macedonia
tonij@feit.ukim.edu.mk

Abstract

The new emerging applications in 5G network, in the context of the Internet of Everything (IoE), will introduce high mobility, high scalability, real-time, and low latency requirements that raise new challenges on the services being provided to the users. Fortunately, Fog Computing and Cloud Computing, with their service orchestration mechanisms offer virtually unlimited dynamic resources for computation, storage and service provision, that will effectively cope with the requirements of the forthcoming services. 5G will use the benefits of centralized high performance computing cloud centers, cloud and fog RANs and distributed peer-to-peer mobile cloud that will create opportunities for companies to deploy many new real-time services that cannot be delivered over current mobile and wireless networks. This paper evaluates a model for fog and cloud hybrid environment service orchestration mechanisms for 5G network in terms of energy efficiency per user for different payloads.

1 Introduction

Mobile and wireless networks have made tremendous growth in the last decade. This growth is due to the support of a wide range of applications and services by the smart mobile devices such as laptops, smartphones, tablets, phablets, etc. This resulted with an increased demand for mobile broadband services [Jan15].

Therefore, many global research and industrial initiatives are already working on the building blocks of the next fifth generation of mobile and wireless networks, also known as 5G [Jan14], [Jan09], [Tud11]. 5G will enable the future Internet of Services (IoSs) paradigms such as Anything as a Service (AaaS), where devices, terminals, machines, and also smart things and robots will become innovative tools that will produce and use applications, services and data.

5G will have to support huge mobile traffic volumes, 1000 times larger than those today in the order of multiples of gigabits per second [SKT14], [Dat13], [Tik15], [GSA15]. The new emerging applications in 5G network, in the context of the Internet of Everything (IoE) [Kal15], [Bre13], will introduce high mobility, high scalability, real-time, and low latency requirements that raise new challenges on the services being provided to the users.

Fortunately, Fog Computing [Bon12], [Lua15], [Vaq14], and Cloud Computing [Arm10], [Zha10], [Kit14] with their service orchestration mechanisms offer virtually unlimited dynamic resources for computation, storage and service provision, that will effectively cope with the requirements of the forthcoming services. Fog Computing extends cloud computing and services to the edge of the network. With its service orchestration mechanisms, it provides data, computing, storage, and application services to end-users that can be hosted at the network edge or even end devices such as set-top-boxes or access points. The main features of Fog are its proximity to end-users, its dense geographical distribution, and its support for mobility.

5G will use the benefits of the centralized cloud, distributed cloud and fog Radio Access Networks and the distributed peer-to-peer mobile cloud among the smart devices. This will create opportunities for companies to deploy many new real-time services that cannot be delivered over the existing mobile and wireless networks [Kit16]. The core idea is to take full advantages of local radio signal processing, cooperative radio resource management, and distributed storing capabilities in edge devices, which can decrease the heavy burden on front haul and avoid large-scale radio signal processing in the centralized baseband unit pool [Chi15].

This paper presents a further extension on the previous studies given in the conference papers [Kit16], [Kit14]. It proposes an architecture for the hybrid cloud and fog computing environment in 5G network. Then this environment is explored in terms of energy efficiency. The rest of the paper is organized as follows. Section 2 describes the 5G requirements from different perspectives. Section 3 describes the hybrid fog and cloud computing environment in 5G network. Section 4 evaluates this environment in terms of energy efficiency per user for different payloads. Finally, Section 5 concludes the paper and provides future work directions.

2 Service Requirements in 5G

5G will be a multi-layered heterogeneous network that will consist of existing 2G, 3G, LTE and future Radio Access Technologies (RATs). It may also converge many other radio technologies like mobile satellite system (MSS), digital video broadcasting (DVB), wireless local access network (WLAN), wireless personal access network (WPAN), etc., with multi-tiers coverage by macro, pico, femto, relay and other types of small cells [Jan14].

5G requirements should be defined in multiple dimensions such as technology perspective, user perspective, network operator perspective and traffic models [Dat13].

From the technology perspective, 5G will be the continuous enhancement and evolution of the present radio access technologies, and also the development of novel radio access technologies to meet the increasing user's demand of future.

From the user's perspective, 5G mobile system will enhance user's experience in many aspects such as: higher demand for data rate and capacity, good performances in terms of pervasive coverage, reliable QoS and battery life of the mobile device, easy to use, affordable price for subscription, safety and reliability, and personalization of the services. 5G should provide user-centric services, where the users can customize subscription of services and add/remove subscriptions at his/her own will at any time.

From the network operator's point of view 5G should provide sufficient bandwidth and capacity in order to support the high data traffic volume (1000 times greater than today in the order of multiple gigabits per second and at affordable cost. 5G should provide low cost, easy deployment, and simple, scalable and flexible operation in order to decrease CAPEX and OPEX. 5G network should provide a support for backward compatibility with current and legacy networks for investment protection. The future 5G system should support different types of services. The 5G mobile network will be an open service platform to bear all kinds of mobile internet applications and it will support more flexible model of operation that will enable both network operators and service providers to generate their own revenue.

Two key traffic models should be considered: high-speed video flow from the server to the subscriber and massive Machine-to-Machine (M2M), or Device-to-Device (D2D) communications [Tik15].

5G will support a wide range of applications in the context of Internet of Everything (IoE) [Kal15], [Bre13], and services to satisfy the requirements of the information society by the year 2020 and beyond. It will have user-centric approach, where telecom operators will invest in developing new applications that will provide ubiquitous, pervasive, seamless, continual and versatile mobile experience to the end-user [Jan09]. The applications will become more personalized, and more context-aware and will be able to recognize user identity, user location, and user preferences [Kit14]. The new emerging applications in 5G network, will introduce high mobility, high scalability, real-time, and low latency requirements that raise new challenges on the services being provided to the users.

3 Fog and Cloud Computing Environment in 5G network

In order to satisfy 5G Requirements it is necessary Full Network Function Virtualization (NFV) to take place in 5G. Network virtualization pools the underlying physical resources, or logical elements in a network, by using the current technologies such as cognitive and software defined radios in the 5G RAN for fog computing, and software defined networking for centralized cloud services in 5G core [Mar12]. 5G in the hybrid fog and cloud computing environment will use the benefits of the centralized cloud, cloud RAN and fog RAN and the distributed Peer-to-Peer mobile cloud among the devices which will create opportunities for companies to deploy many new real-time services that cannot be delivered over current mobile and wireless networks.

An overview of such 5G network architecture in a hybrid fog and cloud computing environment is given in Figure 1. The architecture consists of centralized cloud computing nodes in 5G core, and the fog computing nodes in the 5G RAN.

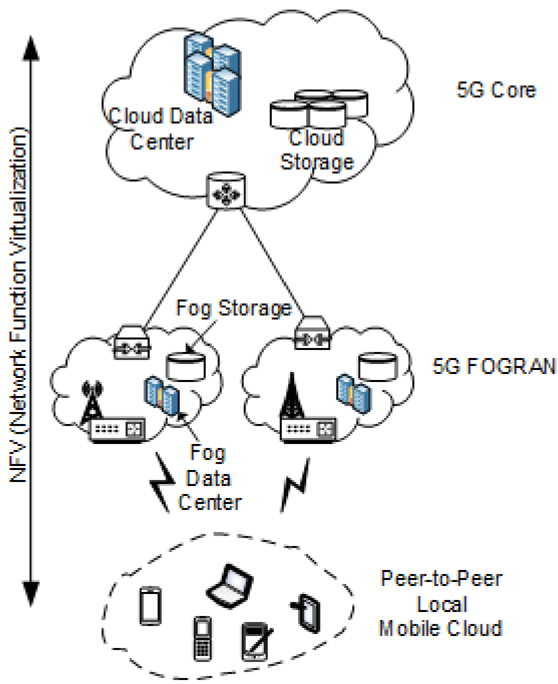

Figure 1: 5G Network Architecture in a Hybrid Fog and Cloud Computing Environment

The centralized cloud computing nodes are powerful, centralized and high performance computing platforms located in 5G core. They provide to the smart devices ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. Like that the limited data processing and storage capabilities of the mobile devices are solved by moving both the data storage and data processing away from the mobile device to the cloud computing nodes [Dih11], [Qur11], [Hua11].

Fog computing nodes (FCN) are typically located away from the main cloud data centers, at the edge of the network. They extend the cloud computing at the edge of the network. Cloud computing on fog nodes enables low and predictable latency. The main features of fog computing nodes are their proximity to end-users, and their dense geographical distribution

Table 1 : A Comparison between Fog and Cloud Computing Nodes [Lua15]

	Fog Computing Nodes	Cloud Computing Nodes
Target Type	Mobile users	General Internet users
Service Type	Limited localized information services related to specific deployment locations	Global information collected from worldwide
Hardware	Limited storage, compute power and wireless interface	Ample and scalable storage space and compute power
Distance to Users	In the physical proximity and communicate through single hop wireless connection	Faraway from users and communicate through IP networks
Working Environment	Outdoor (streets, parklands, etc.) or indoor (restaurants, shopping malls, etc.)	Warehouse-size building with air conditioning systems
Deployment	Centralized or distributed in regional areas by local business (local telco vendor, shopping mall retailer, etc.)	Centralized and maintained by Amazon, Google, etc.

[Bon12], [Lua15], [Vaq14]. They provide applications with awareness of device geographical location and device context. The fog nodes support the mobility of devices i.e. if a device moves far away from the current servicing FCN, the fog node can redirect the application on the mobile device to associate with a new application instance on a fog node that is now closer to the device. [ETS15]. A comparison between cloud computing nodes and the fog computing nodes is given in Table 1.

FCNs absorb the intensive mobile traffic using local fast-rate connections and relieves the long back and forth data transmissions among cloud and mobile devices. This significantly reduces the service latency and improves the service quality perceived by mobile users, and more importantly, greatly saves both the bandwidth cost and energy consumptions inside the

Internet backbone. Fog computing represents a scalable, sustainable and efficient solution to enable the convergence of cloud-based Internet and the mobile computing. Therefore, fog paradigm is well positioned for real time big data analytics, 5G network, and IoT.

In this environment the distributed Peer-to-Peer (P2P) mobile cloud approach among the smart devices can be used [Gup11], [Kav12]. Like that a group of mobile devices acts as a cloud and provides cloud services to other mobile devices with a guaranteed certain level of service agreements. The peers have strong capacities such as storage space, computational power, online time, and bandwidth. The workload of the application is managed in a distributed fashion without any point of centralization. The lack of centralization provides scalability, while exploitation of user resources reduces the service cost. P2P architectures have ability to adapt to network failures and dynamically changing network topology with a transient population of nodes/devices, while ensuring acceptable connectivity and performance. Thus, P2P systems exhibit a high degree of self-organization and fault tolerance.

4 Architecture Evaluation

The performances of the hybrid fog and cloud computing environment in 5G can be explored in many ways such as Round Trip Time (RTT) latency, throughput, product latency – throughput, energy efficiency and power consumption. The focus in this paper is the energy efficiency per user for different data payloads: 10 KB and 10 MB. The most significant impact in the energy efficiency will have the RAN type, while the 5G core impact on the energy efficiency can be treated as a constant, and therefore it can be neglected.

The following scenario will be used. There is a region that contains a group of *N* users uniformly distributed, which are simultaneously covered by several different RANs. Each RAN is connected to several clouds, which can be in the same or different region with the RANs. The smart user devices are assumed to be equally capable, and are located on a different distance from the RANs. They can be simultaneously served by the RANs and the clouds.

Table 2 : Energy per bit for different RAN types and different data file size

	RAN Type		
Parameter	3G	4G	5G
Energy per bit [μJ/bit] (Data File: 10 KB)	100	170	17
Energy per bit [μJ/bit] (Data File: 10 MB)	4	0.3	0.03

4.1 Energy Efficiency per User

The energy efficiency per user (*EE)*, that uses fog or cloud computing service is a product of the energy per bit which depends from the RAN type and the size of data file being transferred to the user:

$$EE = E_{ran}T \qquad \textbf{(1)}$$

where,

E_{ran} is the energy per bit that depends from the type of the RAN;

and T is the size of the payload 10 KB or 10 MB.

The values for the energy per bit for different types RAN networks is provided in [Hua12], and are summarized in Table 2. Here it is assumed that 5G RAN will have 90% improvement in energy per bit over 4G [Tik15].

4.2 Analysis of the Results

Our simulation scenario consists of the following: 10 cloud computing centers, three types of RANs are considered (3G, 4G, and 5G), and the number of the users is varied from 100 to 1000. For simplicity the impact of the distance between the smart user device and the RAN on energy efficiency was neglected. The simulation results are provided in Figure 2. The following can be noticed.

3G RAN wastes a lot of energy for the transfer of big data files. 4G RAN provides much better energy efficiency for large data files, compared to 3G RAN. On the other hand, 4G RAN wastes energy for the transfer of small data files, and 3G RAN demonstrates better performances. Finally, 5G RAN has the best energy efficiency for the transfer independently from the size of data files, that the user is requesting them from the fog or cloud.

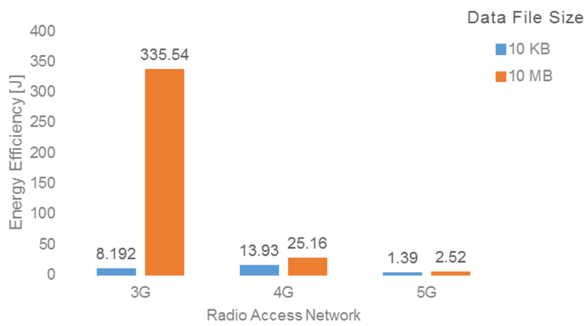

Figure 2: Energy Efficiency in 5G Network in the Hybrid Fog and Cloud Computing Environment

In 5G network where the user device will be served by different RANs, it has to make choice which RAN will be the most suitable for transferring data files. The choice should be make primarily of the size of data files being transferred, throughput, latency, energy efficiency, etc. The algorithm for such RAN selection will be our future work direction.

5 Conclusion

This paper evaluates energy efficiency per user in different payload and networks. The results show that 5G RAN has the best energy efficiency for the transfer independently from the size of data files, compared to 3G and 4G RAN.

5G network will act as a nervous system of the digital society, economy, and everyday people's life. The cloud in 5G networks will be diffused among the client devices often with mobility too, i.e. the cloud will become fog.

More and more virtual network functionality will be executed in a fog computing environment, and it will provide *mobiquitous* service to the users. This will enable new AaaS service paradigms, where devices, terminals, machines, and also smart things and robots will become innovative tools that will produce and use applications, services and data.

Finally, the choice of selecting the most suitable RAN, should be make primarily of the size of data files being transferred, throughput, latency, etc. The algorithm for such RAN selection will be our future work direction.

References

[Jan15] T. Janevski. *Internet Technologies for Fixed and Mobile Networks*. Artech House, USA, 2015.

[Jan14] T. Janevski. *NGN Architectures Protocols and Services*. John Wiley & Sons, UK, 2014.

[Jan09] T. Janevski. *5G Mobile Phone Concept*. Proceedings of 6th IEEE Consumer Communications and Networking Conference - CCNC 2009, pp. 1–2, Las Vegas, Nevada, USA, 2009.

[Tud11] A. Tudzarov and T. Janevski. *Functional Architecture for 5G Mobile Networks*. International Journal of Advanced Science and Technology (IJAST), vol. 32, pp. 65–78, July 2011.

[SKT14] SK Telecom Network Technology Research and Development Center 5G White Paper, *SK Telecom's View on 5G Vision, Architecture, Technology, and Service and Spectrum*, SK Telecom, October 2014.

[Dat13] Datang Mobile Wireless Innovation Center 5G White Paper, *Evolution, Convergence and Innovation*, Datang Telecom Technology and Industry Group, December 2013.

[Tik15] V. Tikhvinskiy and G. Bochechka. *Prospects and QoS Requirements in 5G Networks*. Journal of Telecommunications and Information Technologies, Vol. 1, No. 1, pp. 23 – 26, 2015.

[GSA15] *The Road to 5G: Drivers, Applications, Requirements and Technical Development*. A GSA (Global mobile Suppliers Association) Executive Report from Ericsson, Huawei and Qualcomm, 2015.

[Kal15] V. L. Kalyani, and D. Sharma. *IoT: Machine to Machine (M2M), Device to Device (D2D) Internet of Everything (IoE) and Human to Human (H2H): Future of Communication*. Journal of Management Engineering and Information Technology (JMEIT) Volume -2, Issue 6, pp. 17 – 23, Dec. 2015.

[Bre13] B. Brech, J. Jamison, L. Shao, G. Wightwick. *The Interconnecting of Everything*. IBM Redbook, 2013.

[Bon12] F. Bonomi, R. Milito, J. Zhu, S. Addepalli. *Fog Computing and its Role in the Internet of Things*. Proceedings of the First Edition of the ACM SIGCOMM Workshop on Mobile Cloud Computing (MCC 2012), pp. 13-16, Helsinki, Finland, 2012.

[Lua15] H. T. Luan, L. Gao, Z. Li, L. X. Y. Sun. *Fog Computing: Focusing on Mobile Users at the Edge*. arXiv:1502.01815[cs.NI], 2015.

[Vaq14] L. M. Vaquero, L. Rodero-Merino. *Finding your Way in the Fog: Towards a Comprehensive Definition of Fog Computing*. ACM SIGCOMM Computer Communication Review Newsletter, Vol. 44, No. 5, pp. 27-32, 2014.

[Arm10] M. Armbrust et al. *A view of cloud computing*. Communications of the ACM, vol. 53, No.4, pp.50–58, April 2010.

[Zha10] S. Zhang, S. Zhang, X. Chen, X. Huo. *Cloud Computing Research and Development Trend*. Proceedings of the Second IEEE International Conference on Future Networks (ICFN 2010), pp. 93-97 Sanya, Hainan, 2010.

[Kit14] S. Kitanov and T. Janevski. *State of the Art: Mobile Cloud Computing*. Proceedings of the Sixth IEEE International Conference on Computational Intelligence, Communication Systems and Networks 2014 (CICSYN 2014), pp. 153–158 Tetovo, Macedonia, 2014.

[Kit16] S. Kitanov, E. Monteiro, T. Janevski. *5G and the Fog – Survey of Related Technologies and Research Directions*. in Proceedings of the 18th Mediterranean IEEE Electrotechnical Conference MELECON 2016, pp. 1-6, Limassol, Cyprus, 2016.

[Chi15] M. Chiang: Fog Networking: An Overview on Research Opportunities, white paper, 2015.

[Mar12] J. Marinho and E. Monteiro. *Cognitive radio: Survey on Communication Protocols, Spectrum Decision Issues, and Future Fesearch Directions*. Wireless Networks, Vol. 18, Issue 2, pp. 147–164, Springer, USA, February 2012.

[Dih11] T. H. Dihn, C. Lee, D. Niyato, and P. Wang. *A Survey of Mobile Cloud Computing: Architecture, Applications, and Approaches*. Wireless Communications and Mobile Computing, Wiley, Vol. 13, Issue 18, pp. 1587–1611, 2011.

[Qur11] S. S. Qureshi, T. Ahmad, K. Rafique, and Shuja-ul-islam. *Mobile Cloud Computing as Future for Mobile Applications – Implementation Methods and Challenging Issues*. Proceedings of IEEE Conference on Cloud Computing and Intelligence Systems (CCIS), pp. 467–471, Beijing, China, 2011.

[Hua11] D. Huang et al. *Mobile cloud computing*. IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, Vol. 6, No. 10, pp. 27–31, 2011.

[ETS15] *Fog Computing and Mobile Edge Cloud Gain Momentum Open Fog Consortium*. ETSI MEC and Cloudlets, Version 1.1 Guenter I. Klas, November 22, 2015.

[Gup11] A. Gupta, and L. K. Awasthi. *Peer-to-Peer Networks and Computation: Current Trends and Future Perspectives*. Journal of Computing and Informatics, Vol. 30, pp. 559–594, 2011.

[Kav12] H. Kavalionak and A. Montresor. *P2P and Cloud: A Marriage of Convenience for Replica Management*. Proceedings of the 6th IFIP TC 6 International Conference on Self-Organizing Systems, pp. 60–71, Delft, Netherlands, 2012.

[Hua12] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, O. Spatscheck. *A Close Examination of Performance and Power Characteristics of 4G LTE Networks*. Proceedings of the 10th international conference on Mobile systems, applications, and services (Mobisys 2012), pp. 225-238, 2012.