Open Access

Open Journal of Internet of Things (OJIOT)
Volume 4, Issue 1, 2018

http://www.ronpub.com/ojiot
ISSN 2364-7108

Dynamic Allocation of Smart City Applications

Igor Miladinovic, Sigrid Schefer-Wenzl

Computer Science and Digital Communications, University of Applied Sciences Campus Vienna, Favoritenstrasse 226, Vienna, Austria, {igor.miladinovic, sigrid.schefer-wenzl}@fh-campuswien.ac.at

Abstract

Cities around the world are evaluating the potential of Internet of Things (IoT) to automate and optimize public services. Cities that implement this approach are commonly referred to as smart cities. A smart city IoT architecture needs to be layered and scalable in order to fulfill not only today's but also future needs of smart cities. Network Function Virtualization (NFV) provides the scale and flexibility necessary for smart city services by enabling the automated control, management and orchestration of network resources. In this paper we consider a scalable, layered, NFV based smart city architecture and discuss the optimal location of applications regarding cloud computing and mobile edge computing (MEC). Introducing a novel concept of dynamic application allocation we show how to fully benefit from MEC and present relevant decision criteria.

Type of Paper and Keywords

Visionary paper: *smart city, mobile edge computing, Internet of Things, 5G*

1 Introduction

The Internet of Things (IoT) refers to the interconnection of billions of devices. Predictions on the amount of IoT devices, for example by Nokia [18], Ericsson [7], Gartner [9] or Cisco [8], show all in the same direction—a massive growth of the number of IoT devices in the next years resulting in multiple tens of billions connected devices by 2020. The steadily increasing number of IoT devices with heterogeneous characteristics requires that future networks evolve providing a new architecture capable to manage the expected increase in data generation and to serve people and things optimally [23].

The aim of smart city IoT concepts is to improve the quality of public administration by continuous measurements of city data and adapting behaviour of people and things accordingly [24, 2]. The administration of public resources and services in the majority of cities does not reach the optimal level today. One of the reasons for that is a lack of transparency of the needs for and usage of these resources and services. Without this transparency, a targeted and tailored optimization of city administration is not possible. Data from various sources, such as street-sensors, cameras or vehicles, need to be collected, analyzed and evaluated in order to gain the current state of these resources and services, and identify improvement potential.

Network function virtualization (NFV) [13, 6] provides the scale and flexibility necessary for IoT services by enabling the automated control, management and orchestration of network resources. In our previous work, we have introduced an NFV based scalable IoT architecture [16]. We analyzed and discussed several challenging aspects by implementing NFV in an IoT architecture, including scalability, maintainability, security, interoperability, availability, network traffic, and latency. The outcome was that the NFV based architecture

This paper is accepted at the *International Workshop on Very Large Internet of Things (VLIoT 2018)* in conjunction with the VLDB 2018 Conference in Rio de Janeiro, Brazil. The proceedings of VLIoT@VLDB 2018 are published in the Open Journal of Internet of Things (OJIOT) as special issue.

provides benefits regarding the first five aspects, whereas network traffic and latency increase.

In this paper we address these two challenges by proposing processing capabilities not only in the cloud, but also near the end devices, called fog and edge computing [5]. Fog computing is an emerging technology which provides services closer to the IoT devices promising to address these challenges [20]. Similar to the cloud, it allows applications from different vendors to run in a virtualized environment, but more closely to the IoT devices to improve efficiency and performance. In the case that fog cannot meet time-critical requirements some applications may be running directly on IoT devices, which is known as edge computing. We analyze and evaluate characteristics of cloud, fog and edge architectures for IoT applications with particular focus on the smart city environment. Given that 5G is a promising technology for connecting a large number of smart city devices, we will show how our architecture can be applied on a 5G network.

The remainder of the paper is structured as follows. Section 2 presents our layered smart city architecture. In Section 3, we identify 5G as key access technology for smart cities and introduce Mobile Edge Computing (MEC). Subsequently, Section 4 outlines our approach for the dynamic allocation of smart city applications. In Section 5 we present related work and compare it with our approach. Finally, Section 6 concludes the paper.

2 Cloud, Fog and Edge Computing in a Smart City IoT Architecture

In a smart city IoT environment some devices are connected directly with the network and some over IoT gateways. The cloud approach enables IoT gateways to be very simple and to avoid any application logic—providing a kind of translation between layer 2 protocols and the IP protocol. IoT applications are running in the cloud data center on a standard hardware shared among different applications.

Besides several advantages, the cloud approach does not go without some drawbacks, as described in [16]. In general, the latency and the network traffic are increased, as the data has to be transferred to the cloud and after processing back to the IoT devices. To address this issue, fog and edge computing can be applied.

The main differences between cloud and fog computing are in the distance to the IoT devices and in the computing power. The fog has less computing power than the cloud, but it is placed near the IoT devices, ensuring lower latency and enabling the connection to the Internet to be less critical compared to the cloud [11]. Edge computing, in contrast, requires more complex IoT devices with some computation power and provides even lower latency without the need of permanent connectivity.

Figure 1 illustrates the position of cloud, fog and edge computing in the layered IoT architecture for smart cities we proposed in [17]. At the street layer, sensors and actuators are connected either directly or via an IoT gateway with the IP network. Examples are magnetic sensors, video cameras or lighting controllers. Data processing performed directly on the IoT devices corresponds to edge computing. The city layer consists of network routers and switches building up an IP network for the connectivity of IoT devices with the Internet. Some computing components can be placed in the city layer close to IoT devices, enabling fog computing for certain applications. Data processing on the IoT gateways also counts in this category. Finally, at the data center layer IoT data are processed and made available for the application layer. The key technology in the data center layer is the cloud providing an elastic, scalable, secure, and reliable data processing infrastructure.

3 The Role of 5G for Smart Cities

In a smart city architecture the connection between street and city layer is a challenging topic—there is a need for a new, converged access architecture, capable of serving people and things optimally [23]. 5G is a promising technology, expected to be able to address this challenge. Unlike other evolution steps in mobile networks, such as 3G and 4G, 5G goes beyond higher data rates for mobile Internet [10]. Development on 5G focuses on eight major requirements [1], all highly relevant for IoT. These requirements include:

(1) up to 10 Gbps data rate in real networks,

(2) round trip latency of 1 ms,

(3) high bandwidth in unit area,

(4) very large number of connected devices,

(5) five-nines availability (99.999%),

(6) almost full coverage,

(7) up to 90% reduced energy consumption, and

(8) high battery live of connected devices (which is related to the 7th requirement).

With these requirements 5G fits very well in a smart city architecture—connecting the growing number of IoT devices efficiently. Therefore, we presume 5G in our smart city architecture, as shown in Figure 1. To reduce the latency in a 5G network, an emerging technology has been introduced—Mobile-Edge Computing (MEC) [22],

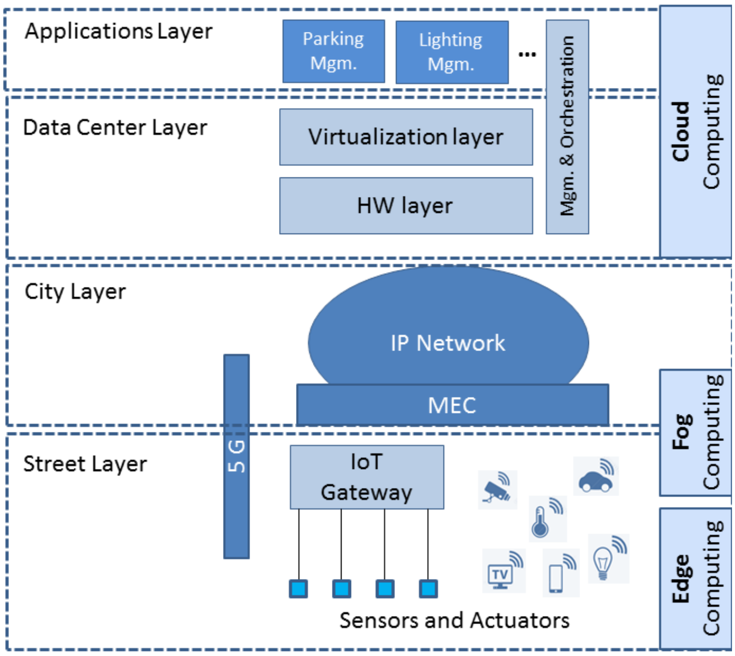

Figure 1: Layered IoT architecture

specified by ETSI MEC ISG [15]. In [19] Sabella et al. consider the MEC architecture with respect to IoT. The concept of MEC is to offer cloud computing functionalities within the Radio Access Network (RAN), very close to end customers. This reduces the latency for cloud applications, reduces the traffic towards the core network and increases availability—as the same application can be deployed several times on different RANs.

Using cloud, fog and edge terminology, MEC is comparable with fog computing with the exception that data processing on the IoT gateways, which counts as fog computing, is not in the scope of MEC. From the point of view of a 5G network, the IoT gateway has a role of an end device, similar as any other IoT devices with 5G connectivity. In this paper we will focus on a smart city IoT architecture with 5G and discuss the differences between cloud computing and MEC.

4 Dynamic Application Allocation

Smart city applications have very different requirements. Some of them are time critical (e.g. traffic incident detection), some of them generate a lot of data (e.g. video surveillance), so that the allocation of applications among cloud and mobile edge is a non-trivial task. While MEC reduces the network traffic and enables a network latency of 17ms [25], it is limited in processing power, storage capabilities and scalability. In general, non-time-critical applications which need a small amount of sensor data should be placed in the cloud. On the other hand, time-critical applications have to be placed in the edge, as close as possible to the IoT devices, in order to ensure the required functionality. The question is how to manage all the applications which do not belong to one of these two categories.

To address this question we need a classification of applications. We specify the following criteria for each application: priority, storage, and data. Each application is evaluated towards these criteria and for each criterion a rating is assigned, among H (high), M (medium) and L (low). Priority is the most critical criterion and all applications with priority H have to be placed at mobile edge to ensure their functionality. An example is car incident detection where all driving cars in the vicinity have to be alerted immediately in order to take appropriate actions. Applications with priority L in contrast are not time-critical and can be placed in the cloud without any impact on their functionality. The storage criterion specifies the storage need of the application, including the application itself and any application data stored locally. Finally, the data criterion indicates the amount of data this application receives from sensors and forwards to actuators. It is measured in bytes per hour and mapped to the H, M and L rating.

Using these criteria we determinate which applications should be placed in the mobile edge and which in the cloud. Applications with priorities M and L are rated considering other criteria and ranked. The first n applications from this ranking are then placed in the mobile edge, whereas n depends on the available capacity of the mobile edge. For example, an application with storage L and data H will be top rated. Allocating this application in the mobile edge does not cost much storage space and reduces the traffic towards the core networks, as the data can be processed by MEC directly.

However, these criteria, in particular the data, can vary depending on several factors, including time of day, day of week or season. Therefore, to ensure an optimal application allocation we need a dynamic allocation of applications, where the three criteria are periodically evaluated. Applications are accordingly moved from the cloud to the mobile edge and vice versa.

Consider for example two typical smart city scenarios, smart parking and smart lighting. Ineffective parking management is a constant challenge in cities around the world. It contributes to pollution, causes frustration and increases traffic incidents. Parking sensors including in-ground magnetic sensors, video-based sensors and radar sensors may be connected over IoT gateways indicating availability of parking spaces. An application shows the parking availability on smart phones where an action, for example booking, can be taken. Furthermore, the application can support drivers with disabilities to locate suitable parking spots. A parking management application would have a priority M for the mobile edge, storage L or M, depending on the implemented functionalities, and data between L, M or H, depending on the location, but also on time of day and possibly day of week. During the day the data generated for this application will be in general higher as in the night time.

Smart lighting controls and monitors street lights, and provides the current status of each light. A lighting management application performs automated scheduling for lights with automated dimming or brightening, as needed. It can also manage the light intensity, depending on environmental conditions, weather, season, time of day and location within the city. A lighting management application would have a priority M for the mobile edge, storage L, and data between L or M, depending on the mentioned factors. In contrast to the parking management application, the data generated for this application will be higher in the night time than during the day.

In our smart city architecture applications are dynamically allocated between the cloud and the mobile edge. Only if the capacity of mobile edge allows it, the two presented applications would be both placed at the mobile edge. Otherwise, only the application with higher ranking is placed in the mobile edge—which is likely to be the parking management application during the day and the lighting management application in the night time. Table 1 summarizes the values of our three criteria for these two applications in the day time and in the night time, as well as the allocation of both applications. As a side effect, the latency for the application used more frequently and hence placed in the mobile edge is also reduced, resulting in better user experience.

Table 1: Examples of dynamic application allocation

Application	Prio.	Strg.	Lat.	Allocation
Day time				
Smart Parking	M	L-M	M-H	MEC
Smart Lighting	M	L	L	Cloud
Night time				
Smart Parking	M	L-M	L	Cloud
Smart Lighting	M	L	M	MEC

Dynamic application allocation does not come without challenges. First, moving applications between cloud and edge generates additional network traffic. Therefore, these transitions need to be performed in the time when the network traffic is low, according to the historical data of the particular system. Second, permanently installing applications in the edge and in the cloud could cause a failure during the installation and reduce the availability of the service. To address this issue we propose that the application in the cloud is always installed and only deactivated if not needed. The resources in the cloud are not as limited as in the edge and enable this behavior. In the case that the installation in the edge fails, the application in the cloud can be activated, as a kind of cold standby redundancy. This also contributes to the first issue, as there is no need to move applications from the edge to the cloud anymore.

5 RELATED WORK

The comparison between edge and cloud computing has been a subject of multiple research works. The fact that it is not efficient to forward all raw sensor data in the cloud has been discussed in [12]. In that paper a smart gateway with edge computing is used for rule and event processing and to reduce the amount of raw data to be transmitted to the cloud. A comprehensive literature survey on MEC including its benefits and challenges is presented in [14], with a particular focus on resource management. Chaudhary et al. compared in [4] fog and cloud computing with focus on network service chaining in 5G networks with and without SDN and NFV technologies. They discussed issues on latency, data offloading and security. In contrast to our work, they analyzed particularly security aspects and did not consider dynamic reallocation of the services. The issue with limited resources in the edge is also addressed by Beraldi et al. in [3]. They proposed a solution based on load balancing and resource sharing between two edge data centers and stated that the quality of service can be improved by sharing resources between two data centers. Le Tan et al. proposed in [21] an algorithm for location aware load prediction and resource allocation in MEC, by using its historical load time data and data of neighbors' MEC data centers. The results show a more efficient resource allocation and cost savings.

All these works identify edge benefits, but also its limited resources. Several solutions have been proposed to overcome this limitation, in our paper we presented a novel one with dynamic application allocation depending on the presented criteria.

6 Conclusion

In this paper we present an innovative NFV based IoT architecture for smart cities with dynamic allocation of applications between the cloud and the mobile edge. Given that the number of smart city applications is expected to be permanently growing, the optimal location of applications is essential to fully benefit from the potential of MEC. The presented scenarios exemplarily demonstrate that the optimal location of applications is changing depending on several conditions. The proposed dynamic allocation using criteria based ranking ensures the optimal use of resources in the mobile edge. In our future work we are going to apply this concept on a smart railway station our university is working on and to implement and further evaluate several ranking algorithms.

References

[1] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5g wireless networks: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, thirdquarter 2016.

[2] E. Ahmed, I. Yaqoob, A. Gani, M. Imran, and M. Guizani, "Internet-of-things-based smart environments: state of the art, taxonomy, and open research challenges," *IEEE Wireless Communications*, vol. 23, no. 5, pp. 10–16, October 2016.

[3] R. Beraldi, A. Mtibaa, and H. Alnuweiri, "Cooperative load balancing scheme for edge computing resources," in *2017 Second International Conference on Fog and Mobile Edge Computing (FMEC)*, May 2017, pp. 94–100.

[4] R. Chaudhary, N. Kumar, and S. Zeadally, "Network service chaining in fog and cloud computing for the 5g environment: Data management and security challenges," *IEEE Communications Magazine*, vol. 55, no. 11, pp. 114–122, NOVEMBER 2017.

[5] M. Chiang and T. Zhang, "Fog and IoT: An Overview of Research Opportunities," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 854–864, Dec 2016.

[6] J. Costa-Requena, J. L. Santos, V. F. Guasch, K. Ahokas, G. Premsankar, S. Luukkainen, O. L. Pérez, M. U. Itzazelaia, I. Ahmad, M. Liyanage, M. Ylianttila, and E. M. de Oca, "SDN and NFV integration in generalized mobile network architecture," in *2015 European Conference on Networks and Communications (EuCNC)*, June 2015, pp. 154–158.

[7] Ericsson, "Ericsson mobility report, on the pulse of the networked socienty," Report, Jun 2016.

[8] D. Evans, "The internet of things, how the next evolution of the internet is changing everything," Whitepaper, Cisco Internet Business Solutions Group (IBSG), Apr 2011.

[9] Gartner Inc., "Gartner says 6.4 billion connected things will be in use in 2016, up 30 percent from 2015," http://www.gartner.com/newsroom/id/3165317, Nov. 2015.

[10] GSMA Intelligence, "Understanding 5G: Perspectives on future technological advancements in mobile," White Paper, Dec. 2015.

[11] D. Hanes, G. Salgueiro, P. Grossetete, R. Barton, and J. Henry, *IoT Fundamentals: Networking Technologies, Protocols, and Use Cases for the Internet of Things*, ser. Fundamentals. Pearson Education, 2017.

[12] Y. H. Lee and S. Nair, "A smart gateway framework for iot services," in *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Dec 2016, pp. 107–114.

[13] J. Li, E. Altman, and C. Touati, "A general sdn-based iot framework with nvf implementation," *ZTE Communications*, vol. 13, no. 3, pp. 42–45, Sep. 2015. [Online]. Available: https://hal.inria.fr/hal-01197042

[14] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Communications Surveys Tutorials*, vol. 19, no. 4, pp. 2322–2358, Fourthquarter 2017.

[15] MEC ETSI ISG, "Mobile edge computing (mec); framework and reference architecture," ETSI GS MEC 003 V1.1.1 (2016-03), 2016.

[16] I. Miladinovic and S. Schefer-Wenzl, "A highly scalable iot architecture through network function virtualization," *Open Journal of Internet Of Things (OJIOT)*, vol. 3, no. 1, pp. 127–135, 2017, Special Issue: Proceedings of the International Workshop on Very Large Internet of Things (VLIoT 2017) in conjunction with the VLDB 2017 Conference in Munich, Germany. [Online]. Available: http://nbn-resolving.de/urn:nbn:de:101:1-2017080613543

[17] I. Miladinovic and S. Schefer-Wenzl, "Urban mobility in the city of things," in *International Conference on Working Class Districts*, Vienna, Austria, September 2017.

[18] Nokia, "A buyer's and influencer's guide to connected device management, for all things connected in mobile, home and IoT," Strategic Whitepaper, Jul 2016.

[19] D. Sabella, A. Vaillant, P. Kuure, U. Rauschenbach, and F. Giust, "Mobile-edge computing architecture: The role of mec in the internet of things," *IEEE Consumer Electronics Magazine*, vol. 5, no. 4, pp. 84–91, Oct 2016.

[20] Y. Y. Shih, W. H. Chung, A. C. Pang, T. C. Chiu, and H. Y. Wei, "Enabling low-latency applications in fog-radio access networks," *IEEE Network*, vol. 31, no. 1, pp. 52–58, January 2017.

[21] C. N. L. Tan, C. Klein, and E. Elmroth, "Location-aware load prediction in edge data centers," in *2017 Second International Conference on Fog and Mobile Edge Computing (FMEC)*, May 2017, pp. 25–31.

[22] S. Wang, X. Zhang, Y. Zhang, L. Wang, J. Yang, and W. Wang, "A survey on mobile edge networks: Convergence of computing, caching and communications," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2017.

[23] M. K. Weldon, *The Future X Network, A Bell Labs Perspective*. CRC Press, 2016.

[24] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22–32, Feb 2014.

[25] J. Zhang, W. Xie, F. Yang, and Q. Bi, "Mobile edge computing and field trial results for 5g low latency scenario," *China Communications*, vol. 13, no. 2, pp. 174–182, Nov. 2016.

AUTHOR BIOGRAPHIES

Igor Miladinovic is the head of the degree program Information Technologies and Telecommunication at the University of Applied Science Campus Vienna. He received the Ph.D. degree (with honors) in electrical engineering from Vienna University of Technology in 2003. He worked for more than 10 years on leading positions at Alcatel-Lucent (later Nokia) in area of telecommunication software and in parallel as a lecturer at two universities. His research interests cover telecommunication networks, software engineering and IoT, with over 40 publications in international journals, conferences and as book chapters.

Sigrid Schefer-Wenzl is a senior researcher and lecturer at the University of Applied Sciences Campus Vienna, WU Vienna, and the University of Salzburg. She has worked as a software analyst and developer in several companies and received the Ph.D. degree (with honors) in Information Systems from WU Vienna. Her current research and teaching activities focus on the fields of software engineering and IT-security. Sigrid has published the results of her work in top ranked journals and presented her work at various international conferences.