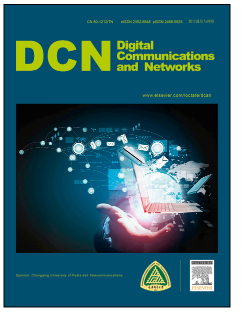
Accepted Manuscript



Edge cloud computing technologies for internet of things: A primer

Yuan Ai, Mugen Peng, Kecheng Zhang

PII: S2352-8648(17)30133-5
DOI: 10.1016/j.dcan.2017.07.001
Reference: DCAN 87

To appear in: *Digital Communications and Networks*

Received Date: 24 April 2017
Revised Date: 28 June 2017
Accepted Date: 2 July 2017

Please cite this article as: Y. Ai, M. Peng, K. Zhang, Edge cloud computing technologies for internet of things: A primer, *Digital Communications and Networks* (2017), doi: 10.1016/j.dcan.2017.07.001.

Title Page

Paper Title:

Edge Cloud Computing Technologies for Internet of Things: A Primer

Authors:

Yuan Ai, Mugen Peng, and Kecheng Zhang

Author Information:

Yuan Ai (e-mail: aiyuan@bupt.edu.cn), Mugen Peng (e-mail: pmg@bupt.edu.cn), and Kecheng Zhang (e-mail: buptzkc@163.com) are with the Key Laboratory of Universal Wireless Communications for Ministry of Education, Beijing University of Posts and Telecommunications, China.

Tel: 86-10-62286952

Edge Computing Technologies for Internet of Things: A Primer

Yuan Ai, Mugen Peng, and Kecheng Zhang

Abstract

With the rapid development of mobile internet and internet of things (IoT) applications, the traditional centralized cloud computing is encountering severe challenges, such as high latency, low spectral efficiency (SE), and non-adaptive machine type of communication. Motivated to solve these challenges, new technology is driving a trend which shifts the function of centralized cloud computing to the edge devices of the network. Several edge computing technologies originating from different backgrounds have been emerging to decrease latency, improve SE, and support the massive machine type of communication. This paper comprehensively presents a tutorial of three typical edge computing technologies, including the mobile edge computing, cloudlets, and fog computing. In particular, the standardization efforts, principles, architectures, and applications for these three technologies are summarized and compared. From the viewpoint of radio access network, the difference between mobile edge computing and fog computing are highlighted, and the characteristics of fog computing based radio access network are discussed. Finally, open issues and future research directions are identified as well.

Index Terms

internet of things (IoT), mobile edge computing, cloudlets, fog computing

This work was supported in part by the National Natural Science Foundation of China under Grant No. 61361166005, the National High Technology Research and Development Program of China under Grant No. 2014AA01A701, the National Basic Research Program of China (973 Program) under Grant No. 2013CB336600, and the State Major Science and Technology Special Projects (Grant No. 2016ZX03001020-006).

Yuan Ai (e-mail: aiyuan@bupt.edu.cn), Mugen Peng (e-mail: pmg@bupt.edu.cn) and Kecheng Zhang (e-mail: buptzkc@163.com) are with the Key Laboratory of Universal Wireless Communications for Ministry of Education, Beijing University of Posts and Telecommunications, China.

I. Introduction

Over the past decades, cloud computing has been greatly developed and applied due to high cost-efficiency and flexibility achieved through consolidation, in which computing, storage, and network management functions work in a centralized manner. With the rapid development of mobile internet and internet of things (IoT) applications, the existing centralized cloud computing architecture is encountering severe challenges. Mobile devices connected to the distant centralized cloud servers try to obtain sophisticated applications, which imposes additional load on both the radio access networks and the backhaul networks and introduces a high latency [1]. In addition, with the explosive growth in variety of access devices and end users demands, IoT is driving a digital transformation in all aspects of the current modern lives [2]. It is estimated by Cisco that the number of devices connected to the IoT will be added to 50 billion by 2020 [3]. The emerging IoT introduces new challenges that cannot be adequately addressed by the centralized cloud compute architecture, such as stringent latency, capacity constraints, resource-constrained devices, uninterrupted services with intermittent connectivity, and enhanced security [4]. An advanced cloud computing paradigm that breaks through the centralized architecture and alleviates the capacity and latency constraints is urgently required to cope with these challenges.

Internet of Things refers to the interaction and communication between billions of devices that produce and exchange data related to real-world objects (i.e. things) [5]. The IoT’s characteristics, including an ultra largescale network of things, device and network level heterogeneity, and large numbers of events generated by these things, will make development of the diverse applications and services a very challenging task [6]. These requirements are becoming hard to accomplish in the IoT+Cloud scenario. IoT applications generate enormous amounts of data by IoT sensors. The big data are subsequently analyzed to determine reactions to events or to extract analytics or statistics. However, sending all the data to the Cloud will require prohibitively high network bandwidth. Recent research efforts are investigating how to better exploit capabilities at the edge of the network to support the IoT and its needs [7]. In edge computing, the massive data generated by different kinds of IoT devices can be processed at the network edge instead of transmitting it to the centralized cloud infrastructure due to bandwidth and energy consumption concerns. Edge computing can provide services with faster response and greater quality, in comparison with cloud computing. Edge computing is more suitable to be integrated with IoT to provide efficient and secure services for a large number of end-users, and edge computing-based architecture can be considered for the future IoT infrastructure [8].

Recently, nascent technologies and applications are driving a trend in the computing and communication landscape which shifts the function of centralized cloud computing into the edge devices of the network [9]. Software defined networking (SDN) and the associated concept of network function virtualization (NFV) are proposed as emerging solutions for the future network [10]. In particular, NFV enables the edge devices to provide computing services and operate network functions by creating multiple virtual machines (VMs). Moreover, ultra-low latency is identified as one of the major requirements of the fifth generation (5G) radio access networks (RANs) [11]. To decrease the latency, mobile operators are prone to deploying the application and content at the edge of the network. Meanwhile, operators can open the edge devices of RANs to third-party partners, allowing them to rapidly deploy innovative applications and content towards mobile subscribers, enterprisers, and other vertical segments [12]. Although the computing capabilities of wearable watches, smart phones, and other IoT devices have been significantly improved, they are still constrained by the fundamental challenges, such as memory size, battery life, heat dissipation, etc. Mobile devices need to extend battery lifetime by offloading energy consuming computation of the applications to the edge of the network [13].

Motivated to efficiently provide the massive machine type of communication, ultra-reliable low latency communication, and high spectral efficiency (SE), the industry investment and research interest focus on the edge of the network have grown dramatically [14]. To support low-latency requirements for resource-intensive applications, a new architectural element called cloudlets, has been proposed by [15]. In order to accelerate the development of the ecosystem based on cloudlets, the open edge computing (OEC) initiative has been launched in June 2015 by Vodafone, Intel, and Huawei companies in partnership with Carnegie Mellon University (CMU). Similarly, Nokia Networks company introduced a computing platform in 2013, which is integrated with the base station [16]. The initial concept that applications and services are executed at the edge of the network has been formed gradually. In September 2014, a new Industry Specification Group (ISG) was proposed to be set up in European Telecommunications Standards Institute (ETSI) to allow the creation of industry specifications for mobile edge computing (MEC), which has been supported by Huawei, IBM, Intel, Nokia Networks, NTT DoCoMo, Vodafone, and etc. [17]. In MEC World Congress 2016, the MEC ISG has renamed *Mobile Edge Computing* as *Multi-access Edge Computing* in order to reflect the growing interests from non-cellular operators. It is anticipated that *Multi-access Edge Computing* will take effects starting from the end of 2017 [18]. Another similar technology for the edge computing is known as fog computing, which was initiated by Cisco in 2012 [19]. In order to accelerate the adoption of fog computing, the OpenFog Consortium has been founded by ARM, Cisco, Dell, Intel, Microsoft and Princeton University in November 2015.

As three typical edge computing solutions for IoT, it is urge to highlight the differences among them, and the development of a framework for edge computing technologies with reference to background, system architecture, and key techniques is challenging. An existing tutorial articles on edge computing is shown in Table I. The scope of these existing publications provided neither an in-depth discussion nor a comparison in terms of improving SE, decreasing latency, and supporting the massive machine type of communication from the perspectives of the overall system architecture and key techniques. It is critical, therefore, to show a tutorial in edge computing by presenting a comprehensive survey framework. Considering the important aspects and active research activities of edge computing, a tutorial of system architectures, key techniques, and application characteristics, is presented in this paper to promote the research and commercial success of edge computing. Additionally, given the extensiveness of the research area, more open issues and challenging work on both transforming well established infrastructure of fog computing for 5G RANs, which are necessary to make the elaborative investigations in the future, are introduced as well.

TABLE I

SUMMARY OF EXISTING SURVEY ARTICLES ON EDGE COMPUTING

Aspects	**Survey Papers**	**Contributions**
Cloudlet	[15]	A new cloudlet-based architecture for overcoming the technical obstacles in mobile computing.
	[24]	An open ecosystem based on the concept of cloudlets supporting many exciting mobile applications.
Mobile edge computing	[12]	An overview of MEC definition, architectural blueprint, the requirements and challenges for MEC as well as the objectives of the MEC initiative.
	[13]	A comprehensive survey of the state-of-the-art MEC research focusing on joint radio and computational resource management.
	[41]	A comprehensive survey of major use cases and reference scenarios, current advancement in standardization of the MEC, and the research on computation offloading.
Fog computing	[4]	A summary of the opportunities and challenges of fog computing focusing primarily on the networking context of IoT.
	[54]	An overview of Fog computing definition, reference architecture, use cases and challenges for fog computing as well as the future research and work.
Comprehensive Surveys	[9]	An overview of edge computing definition, origin and background, challenges, and applications. Discussions of the future research directions of edge computing.
	[14]	A comprehensive comparison of three approaches: fog computing, MEC and Cloudlet. Discussions of further work and research in order to get concepts like Fog, MEC and Cloudlets adopted by industry.
	This Article	A comprehensive tutorial of state-of-the-art three edge computing technologies, including MEC, cloudlets and fog computing. A comparison of the standardization efforts, principles, architectures, and applications for these three technologies. The difference between mobile edge computing and fog computing from the viewpoint of radio access network.

The rest of this paper is organized as follows: Section II surveys the principles and applications of cloudlets. The standardization efforts, applications, architecture and key technologies of mobile edge computing are presented in Section III. Section IV summarizes the standardization efforts, applications, system architecture of fog computing, and compares MEC and fog computing based radio access networks (Fog-RANs). The open issues and challenges are shown in Section V, prior to the conclusion in Section VI.

II. CLOUDLET: THE EDGE OF INTERNET

One of the critical challenges in cloud computing is the end-to-end responsiveness between the mobile device and associated cloud [20]. To address this challenge, the cloudlet is proposed, which is a mobility-enhanced small-scale cloud data center that is located at the edge of the internet. A cloudlet is a trusted, resource-rich computer or cluster of computers that is well-connected to the internet and available for use by nearby mobile devices [15].

The main purpose of the cloudlet is supporting resource-intensive and interactive mobile applications by providing powerful computing resources to mobile devices with lower latency. UEs can access the computing resources in the nearby cloudlet through a one-hop high-speed wireless local area network. Cloudlets represent the middle tier of the 3-tier hierarchy architecture (mobile device layer, cloudlet layer, and cloud layer) to achieve crisp response time. There are a few but important differentiators between cloud and cloudlet:

1) Compared to the cloud data center, a cloudlet needs to be much more agile in their provisioning because the association with mobile devices is highly dynamic with considerable churn due to user mobility;

2) To support user mobility, virtual machine (VM) handoff technology needs to be used to seamlessly migrate the offloaded services on the first cloudlet to the second cloudlet as a user moves away from the currently associated cloudlet;

3) Since cloudlets are small data centers distributed geographically, a mobile device first has to discover, select, and associate with the appropriate cloudlet among multiple candidates before it starts provisioning.

A. Principle of Cloudlet

Mobile users exploit VM technology to rapidly instantiate the customized software on a nearby cloudlet. Generally, the customization of a base system specialized for a certain application is small. If the base VM exists on the cloudlet, only its difference relative to the desired custom VM, called a VM overlay, needs to be transferred. The approach of using VM overlays to provisioning cloudlets is called VM synthesis [21]. Cloudlets use the approach of VM synthesis for the rapid provisioning and VM handoff. Fig. 1 shows the relevant steps of dynamic VM synthesis. A mobile device delivers a small VM overlay to the cloudlet that already possesses the base VM from which the overlay is derived (the delivery can be either from the cloud or from the storage on the mobile device). The cloudlet decompresses the overlay, applies the overlay to the base VM to derive the launched VM, and then creates a VM instance from it. The mobile device can begin performing offload operations on this instance.

At the end of the session, the instance is destroyed, but the launched VM image can be retained in a persistent cache for future sessions. To retain some training data for future offload sessions, the cloudlet generates a VM residue that can be sent back to the mobile device and incorporated into its overlay. Experimental results show that cloudlets can decrease 51% response time and reduce energy consumption by up to 42% in a mobile device compared to cloud offload [22]. VM handoff is responsible for the seamlessly transferring VM-encapsulated execution to an optimal offload site as users move. Validation experiments confirm that the resulting mechanism is a promising technique for enabling user mobility with low end-to-end latency applications [23].

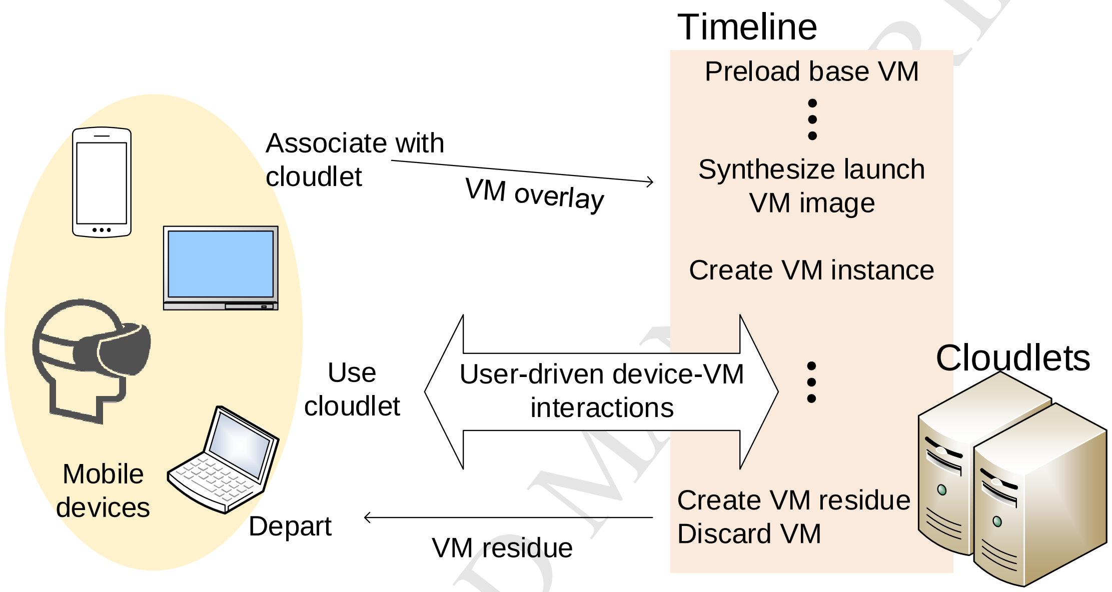


Fig. 1. Dynamic VM Synthesis

B. *Application of Cloudlet*

An open ecosystem based on the concept of cloudlets supports and enables many exciting mobile applications that are both compute-intensive and latency-sensitive [24]. For example, by leveraging a low end-to-end latency, the real-time interaction can be implemented on wearable cognitive assistance [25]. By real-time data analysis at the edge of internet, cloudlets can reduce ingress bandwidth into the cloud [26]. By serving as physically proximate representatives of the cloud that are unavailable due to failures or cyberattacks, cloudlets can improve robustness and availability in hostile environment [27]. Moreover, cloudlets can enable mobile access to the huge legacy world of Windows-based desktop applications. An VM encapsulating the personal desktop environment of a user is run on a cloudlet, and the user connects to it through a remote desktop protocol. Thus, users can use Windows-based desktop applications on a mobile device such as an Android tablet that is typically ARM-based.

To drive the development of software ecosystems surrounding cloudlets, the open edge computing (OEC) initiative has been launched, synchronizing work with other efforts by ETSI ISG MEC and OPNFV. In addition, Carnegie Mellon University has implemented an open source platform called OpenStack++ that is a derivative of the widely used OpenStack platform for cloud computing, which extends the functionality of Openstack to support cloudlets. Some key technology such as cloudlet discovery, just-in-time provisioning, and VM hand-off have been implemented to be available as open source [28].

III. Mobile Edge Computing: The Edge of Mobile Network

MEC is identified as a key enabler for IoT and mission-critical, vertical solutions, and is recognized as one of the key architectural concepts and technologies. The concept of MEC was defined by ETSI as a new technology that "*provides an IT service environment and cloud-computing capabilities at the edge of the mobile network, within the Radio Access Network (RAN) and in close proximity to mobile subscribers*" [12]. ETSI has published a white paper on MEC, and MEC has been considered a key emerging technology to be an important component of future generation networks [17].

A. MEC Standardization

ETSI has established an Industry Specification Group (ISG) on MEC to develop a standardized, open environment that will allow efficient and seamless integration of third-party applications across multi-vendor platforms in December 2014. Until January 2017, MEC ISG has released six specifications, one of which provides a glossary of terms related to the conceptual, architectural and functional elements of MEC [29]. The purpose of this specification is to enable the consistent use of terminology within ETSI MEC specifications and, beyond the ISG, more widely in industry. Another specification specifies the technical requirements enabling interoperability and deployment and describes example use cases and their technical benefits [30]. The other specification provides a framework and reference architecture to enable mobile edge applications to run efficiently and seamlessly in a mobile network [31]. Moreover, the forth specification in MEC ISG introduces a number of service scenarios that would benefit from the MEC technology [32]. The proof of Concept (PoC) framework specification defines a framework to coordinate and promote multi-vendor Proofs of Concept (PoC) projects illustrating key aspects of MEC technology [33]. PoC are essential to demonstrate MEC as feasible and valuable, to validate the specifications that are being developed, to demonstrate use cases, and ultimately to help develop a diverse and open MEC ecosystem. The last specification describes various metrics which can potentially be improved through deploying a service on a MEC platform, such as latency, energy efficiency, network throughput, system resource footprint and quality [34]. Furthermore, the last specification also describes the best practices for measuring such performance metrics. ETSI has announced six different Mobile Edge Computing Proofs of Concept (MEC PoCs) in Sep. 2016, which have been accepted in MEC World Congress in Munich and contribute to strengthen the strategic planning and decision-making of organizations, and help to identify which MEC solutions may be viable in the network. MEC ISG is now working on 9 new studies related to MEC APIs, management interfaces and essential platform functionality. In addition, the MEC in an NFV environment is emerging on an end-to-end mobility. The NFV platform may be dedicated to MEC or shared with other network functions or applications. MEC exploit the NFV management and orchestration entities and interfaces as much as possible.

B. Application of MEC

Due to the advanced characteristics, such as low latency, proximity, high bandwidth, and real-time insight into radio network information and location awareness, MEC enables a large number of new kinds of applications and services for multiple sectors, such as consumer, enterprise, health, etc. In particular, MEC is deemed to be a promising solution for handing video streaming services in the context of smart cities [35]. Video streams from the monitoring devices are locally processed and analyzed at a MEC server to extract meaning data from video streams. The valuable data can be transmitted to the application server to reduce core network traffic. Augmented Reality (AR) mobile applications have inherent collaborative properties in terms of data collection in the uplink, computing at the edge, and data delivery in the downlink [36]. Augmented reality data requires low latency and a high rate of data processing in order to provide the correct information depending on the location of user. The processing of data can be performed on a local MEC server rather than on a centralized server to provide a perfect user experience. In addition, a vehicular delay-tolerant network-based smart grid data management scheme that leverages the mobile edge computing paradigm was proposed in [37]. The IoT generates additional messaging on telecommunication networks, and requires gateways to aggregate the messages and ensure low latency and security. A new architecture by leveraging MEC to collect, classify, and analyze the IoT data streams was introduced in [38]. The MEC server is responsible for managing the various protocols, distribution of messages and for the processing of analytic. The MEC environment creates a new value chain and an energized ecosystem, which in turn creates new opportunities for mobile operators, application, and content providers.

C. System Architecture of MEC

As shown in Fig. 2, the MEC reference architecture, described by ETSI [31], enables the implementation of MEC applications as software-only entities that run on the MEC host. The mobile edge platform offers the essential environment and functionality required to run MEC application. MEC application are running as VM on top of the virtualization infrastructure, and can interact with the mobile edge platform to perform certain support procedures related to the life-cycle of the application. Furthermore, the virtualization infrastructure includes a data plane that executes the traffic rules received by the mobile edge platform, and routes the traffic among applications, local networks, and external networks. The MEC host level management comprises the mobile edge platform manager and the virtualization infrastructure manager. The former manages the life cycle of applications and the application rules and requirements including service authorizations, traffic rules, DNS configuration and resolving conflicts. The latter is responsible for allocating, managing and releasing visualized (compute, storage and networking) resources of the virtualization infrastructure.

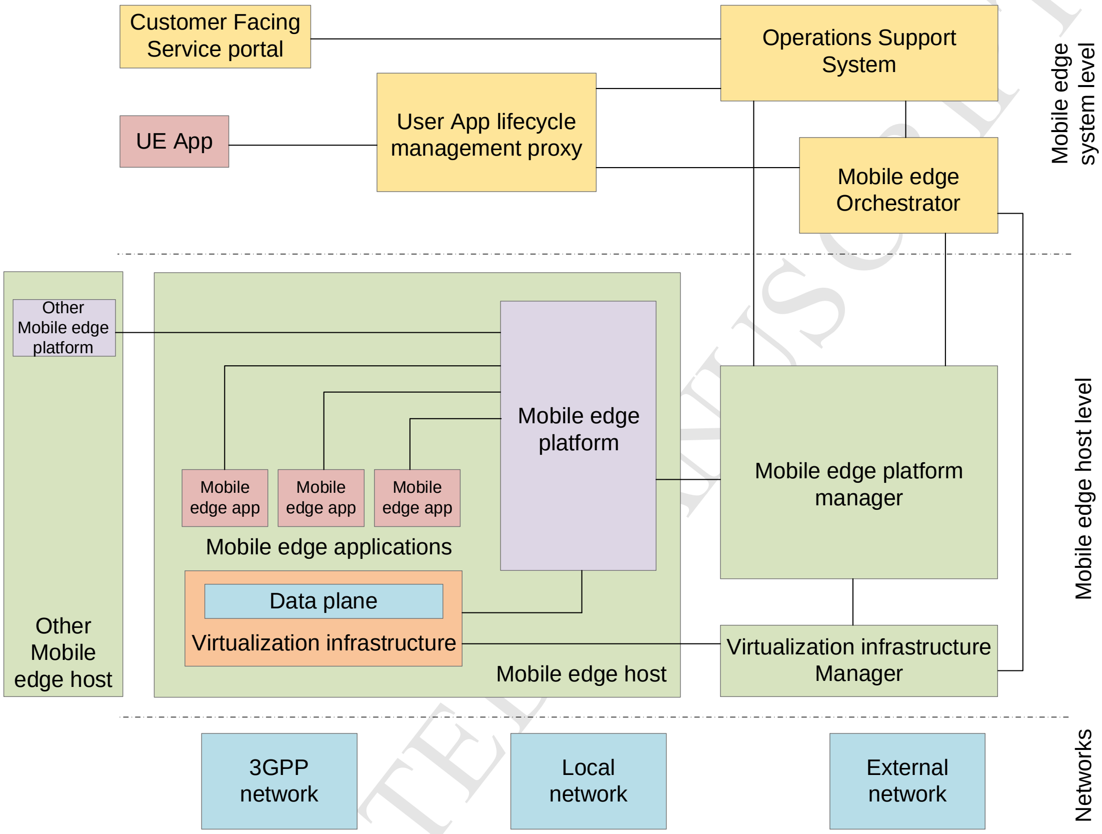


Fig. 2. Mobile edge system reference architecture

The operations support system receives request by a user application via a life-cycle management proxy, or by operators' third-party customers via the customer facing service portal. The operation support system decides whether the requests are granted or not. The granted requests are forwarded to the MEC orchestrator for further processing. The MEC orchestrator is the core functionality as it maintains an overall view based on the deployed MEC hosts, available resources, available MEC services, and topology. For reasons of performance, costs, scalability, operator preferred deployments, MEC supports different deployment scenarios [30], such as at the cellular macro base station (eNodeB) site, at the 3G radio network controller (RNC) site, at a multi-radio access technology cell aggregation site, and at an aggregation point (which may also be at the edge of the core network, e.g. in a distributed data center). An network planning problem determining where to optimally install the MEC servers among the available sites to find a tradeoff between installation costs and quality of service (QoS) has been explored in [39].

D. Key Technologies of MEC

The key technologies of MEC including computation offloading and mobility management are illustrated in Fig. 3. Computation offloading is a procedure that migrates resource-intensive computations from a mobile device to the resource-rich nearby infrastructure [40]. Although mobile devices are constrained by computing capabilities, battery life, and heat dissipation, MEC enables running new sophisticated applications at user equipment (UE) by offloading energy consuming computations of the applications to the MEC server. An important part regarding computation offloading is to decide whether to offload or not, whether full or partial offloading is applicable, what and how could be offloaded. The offloading decision depends on the application model that be classified according to three criteria [41]. The first criterion is whether or not the application contains non-offloadable parts that cannot be offloaded (e.g., user input, camera, or acquire position that need to be executed at UEs). Second, there is no way to estimate the amount of data to be processed for some continuous-execution applications. The third criterion is a mutual dependency of individual part to be processed. Generally, UE needs to be composed of a code profiler, system profilers, and decision engine to manage offloading process [42]. The code profiler is responsible for managing what to offload depending on application type and code/data partitioned. System profilers are in charge of monitoring multiple parameters, such as available bandwidth, data size to transmit, and energy to execute the code. These parameters influence when to offload. The decision engine determines whether to offload or not. Computation offloading decision algorithms have been investigated and compared comprehensively in [41]. The majority of algorithms aims to minimize the energy consumption at the mobile device while subject to the execution delay acceptable by the offloaded application or to find a optimal tradeoff between these two metrics. Numerical results demonstrate that MEC can improve energy efficiency by the computation offloading in heterogeneous networks [43]. A game theoretic approach was proposed for the computation offloading decision making problem among multiple mobile device users in [44]. Numerical results show that the proposed algorithm achieves excellent computation offloading performance and scales well as the user size increases. The energy-efficient resource allocation problems for computation offloading have been researched in [45]. In addition, some efforts have been focused on the joint optimization of radio and computational resources, aimed at minimizing energy consumption under latency and power budget constraints in [46].

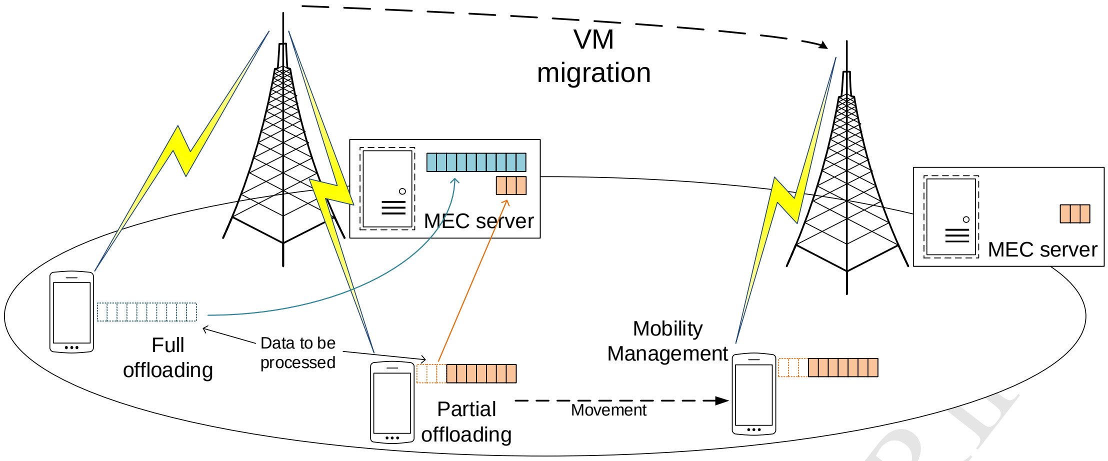


Fig. 3. Offloading and Mobility Management

When the UE performs a handover to another cell, it is important to guarantee the service continuity and QoS requirements [47]. Several MEC applications expect to continue serving the UE after a location change. If the UE forwards the computation to the MEC, VM migration needs to be initiated to provide the continuity of the service. VM migration is a service included in many hyper-visors to move VMs from one physical machine to another, commonly within a datacenter [48]. A crucial part is to decide whether VM migration depends on the service type and requirements, content size, and user class [49]. Note that an ongoing video service with strict QoS requirements may be migrated, and a delay-sensitive measurement task for an emergency warning machine type communication services always are migrated to the optimal MEC server.

On one hand, the decision has to be made indicating whether the service has to be fully or partially migrated, while considering the VM migration cost including the time required for the VM migration and backhaul resources consumption because the traffic exchanges among computing nodes. On the other hand, the benefit of the VM migration is to reduce the service delay. Meanwhile, backhaul resources do not have to be allocated for the transmission of computation results back to the UE. A profit maximization avatar placement strategy was proposed in order to optimize the tradeoff between the migration gain and the migration cost by selectively migrating the VM to their optimal locations [50]. In addition, the prediction technology that is used for the dynamic VM placement and to find the most suitable communication path according to expected users' movement has been explored in [51]. Comparing to state of the art approaches, the proposed algorithm reduces the offloading delay between 10% and 66%.

IV. Fog Computing: Edge Work with Cloud

The OpenFog Consortium was founded to drive industry and academic leadership in fog computing architecture, testbed development, and a variety of inter-operability and composability deliverables that seamlessly leverage cloud and edge architectures to enable end-to-end IoT scenarios [52]. OpenFog Consortium published a white paper on fog computing in February 2016, in which the OpenFog Consortiums approach to an open fog computing architecture (OpenFog architecture) has been outlined [53]. The OpenFog Consortium defines fog computing as "*Fog computing is a system-level horizontal architecture that distributes resources and services of computing, storage, control and networking anywhere along the continuum from Cloud to Things*". Fog computing different from edge computing provides tools for distributing, orchestrating, managing and securing resources and services across networks and between devices that reside at the edge. Edge architecture places servers, applications, and small clouds at the edge. Fog jointly works with the cloud, while edge is defined by the exclusion of cloud.

A. Standardization of Fog Computing

The fog computing standardization is mainly charged by the OpenFog Consortium, whose objective is to influence standard bodies to create standards so that IoT systems at the edge can inter-operate securely with other edge and cloud services in a friction-free environment. The OpenFog Consortium has set up six working groups, including architecture working group, communications working group, manageability working group, security working group, software infrastructure working group and testbed working group. These working groups evaluate, classify and recommend standards, practices and technologies that are appropriate for the OpenFog architecture to address corresponding challenges.

The OpenFog Consortium announces the release of the OpenFog Reference Architecture in Feb. 2017, which is a universal technical framework designed to enable the data-intensive requirements of the IoT, 5G and artificial intelligence (AI) applications [54]. This architecture is the baseline to develop an open architecture fog computing environment, which creates a roadmap and is the first step in creating standards for fog computing. The OpenFog Consortium will establish detailed guidance, interface with standards organizations such as IEEE on recommending standards and specifying APIs for key interfaces in the reference architecture.

The structural aspects and perspectives of the reference architecture are illustrated in Fig. 4, which is used as a common baseline for achieving a multi-vendor interoperable fog computing ecosystem. It is a composite of multiple views to address stake-holders in the fog computing value chain, such as software view, system view and node view. The node view is the lowest level view, which includes the protocol abstraction layer and sensors, actuators and control. The system view is composed of one or more node views coupled with other components to create a platform. Software view is the top three layers that sit on top of the platform hardware layer. The software is running on fog platforms to satisfy a use case requirement. Five cross-cutting perspectives are employed throughout fog computing implementations, including (1) performance and scale perspective, (2) security perspective, (3) manageability perspective, (4) data analytic and control perspective, and (5) IT business and cross-fog applications perspective.

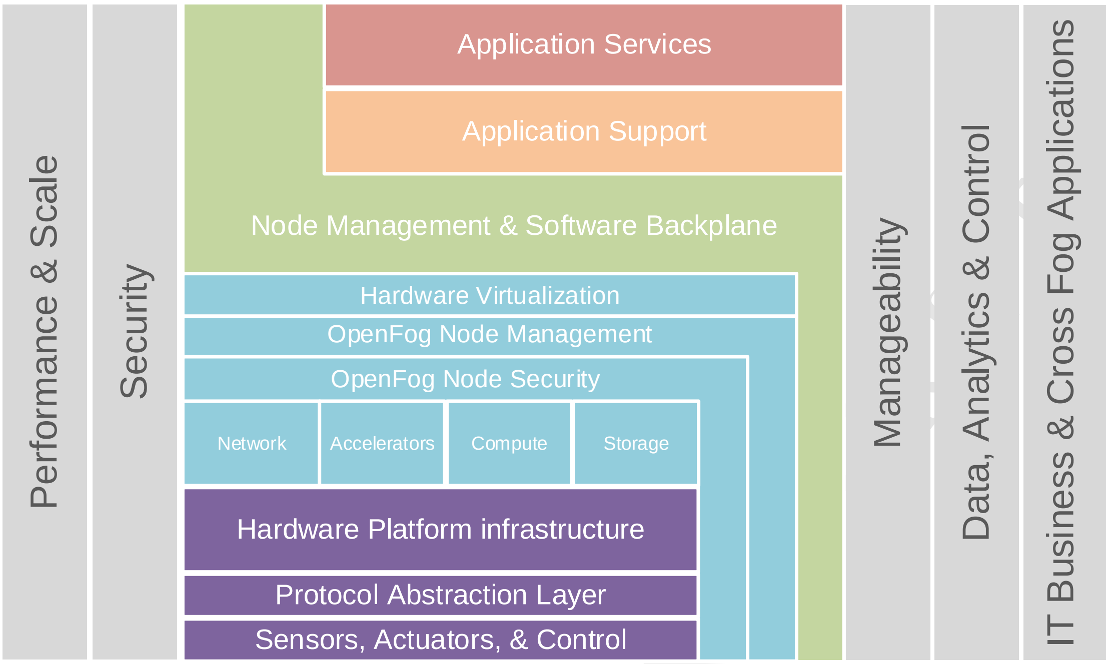


Fig. 4. The OpenFog Reference Architecture Description with Perspectives

B. Application of Fog Computing

An open architecture based on fog computing enable interoperability in IoT, 5G, AI, tactile internet, virtual reality, and other complex data and network intensive applications. IoT applications generate unprecedented amounts of data that can be useful in many ways [55]. Based on this situation, fog nodes can be used carry out data mining and data analysis on a large volume of multi-modal and heterogeneous data from various sensor devices and other IoT devices to achieve real time and fast processing for decision making [56]. A hierarchical fog computing architecture for big data analysis in smart cities was introduced in [57]. Meanwhile, a prototypical system for smart pipeline monitoring was constructed to experimentally evaluate the performance of this architecture. Experimental results demonstrate the feasibility of the system's city-wide implementation in the future smart cities scenario. A privacy-preserving protocol for enhancing security in vehicular crowdsensing based road surface condition monitoring system using fog computing was proposed in [58]. The fog computing based face identification and resolution framework has been explored to solve some security and privacy issues [59]. In addition, since the fog is localized, new services which need mobile networks supporting high data rates and low latency become possible, such as Virtual Reality. The vehicle can be employed as the fog node to make the best utilization of these vehicular communications and computational resources [60]. The mobile fog node can communicate with other fog nodes or provide services including infotainment, advanced driver assistance systems

(ADAS), autonomous driving, collision avoidance, navigation, etc. Emergency, health care, and other latency-sensitive and security-/privacy-sensitive services require fog nodes to be executed between the underlying nodes and the distant cloud [61]. Extensive experiment results validate fog computing supporting medical cyber-physical system can improve significantly cost efficiency by jointly considering base station association, task distribution, and VM deployment [62]. Fog computing provides business value for some applications that require real-time decision making, low latency, improved security, and are network-constrained.

C. System Architecture of Fog Computing

A typical hierarchical architecture based on fog computing is shown in Fig. 5. From the functional point of view, a fog node has several functions, including networking, computing, accelerating, storing and control. Fog nodes can communicate with each other through wired or wireless transmit. Moreover, fog nodes have some general compute capabilities, especially those fog nodes, engaged in enhanced analytic, need to configure accelerator modules such as graphics processing units (GPUs), field programmable gate arrays (FPGAs), and digital signal processors (DSPs), to provide supplementary computational throughput. Many types of storage are required in fog nodes to meet the required reliability, and data integrity of the system and scenario. Generally, there is a rich set of sensors and actuators at the edge of the network in an application scenario. These sensors and actuators are connected to the fog node via a multitude of interfaces, such as PCIe, USB, Ethernet, etc. Fog nodes can be worked in a mesh manner to provide load balancing, resilience, fault tolerance, data sharing, and minimization of cloud communication. There are often three tiers in a fog computing system, but more tiers can be allowed for the special application scenario. At the edge of the network, fog nodes are typically focused on sensor data acquisition/collection, data normalization, and command/control of sensors and actuators. At the next higher tier, fog nodes are focused on data filtering, compression, and transformation. At the higher tiers or nearest the backend cloud, fog nodes are focused on aggregating data and turning the data into knowledge. Architecturally, fog nodes at the edge may be less processing, communications, and storage than nodes at high levels. However, input and output (I/O) accelerators required to facilitate sensor data intake at the edge are much larger in aggregate than I/O accelerators designed for higher level nodes. With the increase in the number of tiers, each tier would be sifting and extracting meaningful data to create more intelligence.

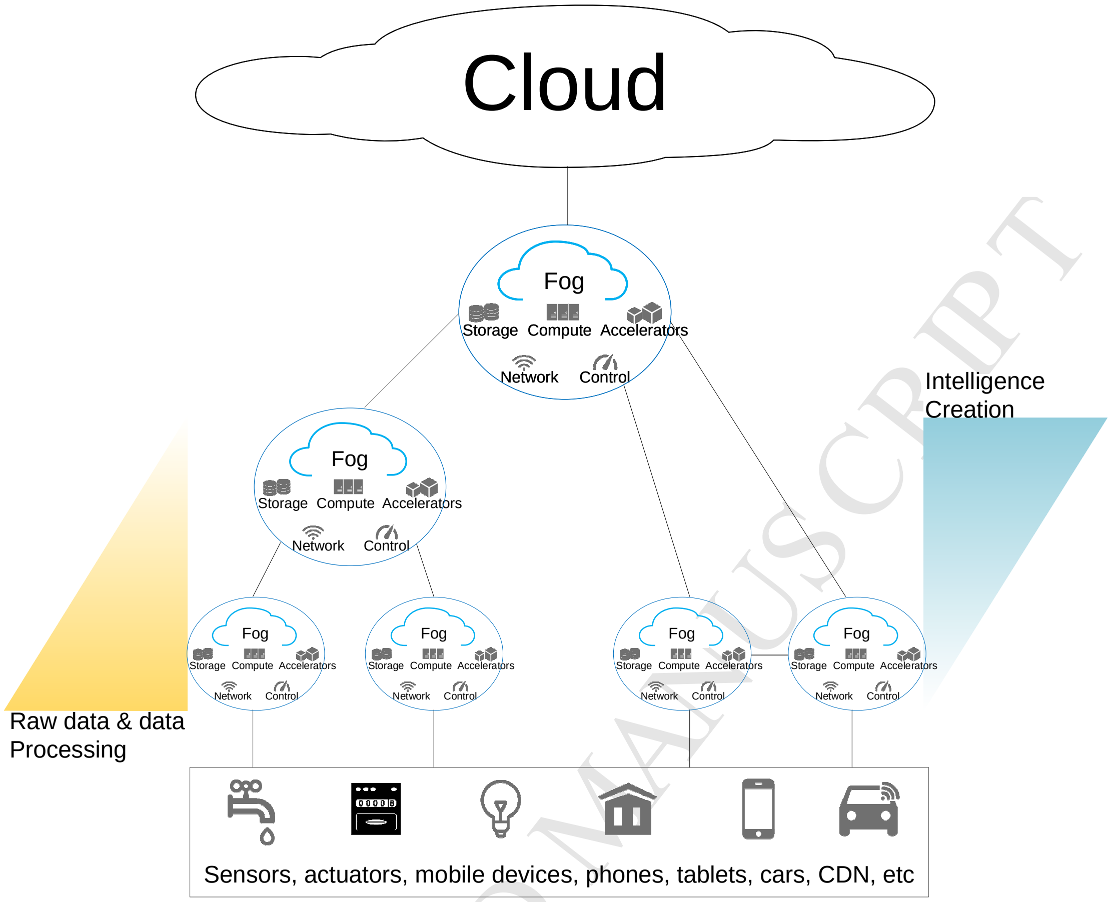


Fig. 5. A typical hierarchical architecture based on fog computing

The traditional centralized cloud computing continues to remain an important part of computing systems as fog computing emerges. Cloud and edge computing complement each other to form a mutually beneficial and inter-dependent service continuum. Some functions are naturally more advantageous to carry out in centralized cloud, while others are better suited to the edge. In [63], a quantitative analysis of energy consumption in a scenario where 25% of the IoT applications demand real-time and low-latency services is presented, and it is shown that the mean energy expenditure in fog computing is 40.48% less than the conventional cloud computing model. Evaluation results show fog computing as an improved, eco-friendly computing platform that can support IoT better compared to the existing cloud computing paradigm [64]. To take advantage of edge computing and to complement centralized cloud computing, a portion of IoT applications that are energy-efficient in fog computing architecture should be identified. In order to compare the energy consumption of applications using centralized data centers (DCs) in cloud computing with applications using nano data centers (nDCs) based on the fog computing, flow-based and time-based energy consumption models for shared and unshared network equipment are proposed in [65]. Correspondingly, a set of measurements and experiments are used to provide data for the models, in which the nano-servers in fog computing are implemented.

The results indicate that the best energy savings using nDCs is for some applications that generate and distribute a large amount of data in end-user premises with low access data rate, such as video surveillance in end-users' homes. The tradeoff between power consumption and transmission delay in the fog-cloud computing system is investigated in [66]. The segmentation of what tasks go to edge and what goes to the backend cloud are application specific. Simulation results about the user case of medical emergency service demonstrate the benefits of coordinated control and management of a combined fog and cloud system. Thus, the design of a coordinated management strategy becomes critical and needs to address the different cloud/edge resources in a joint framework capable of managing the emerging edge-to-cloud computing and network architecture [67].

D. Fog-RAN: Fog Computing in Radio Access Networks

The 5G radio access network (RAN) seamlessly and ubiquitously connects everything, brings a 1000-fold increase in terms of area capacity, supports 100 billion connected wireless devices, and provides diversified use cases as well as high QoS requirements of multimedia applications, compared with current 4G LTE networks [68]. To achieve the above goals, C-RAN has been proposed as a combination of emerging technologies from both the wireless network and cloud computing [69]. The traditional BS is decoupled into two parts: the baseband resources are pooled at BBUs in a centralized location, RRHs with radio frequency functions connect with BBU pool through the wire/wireless fronthaul links. Based on centralized cloud principle of sharing storage and computing resources via virtualization, C-RANs bring the advantage of high spectral efficiency and energy efficiency while at the same time reducing the cost of network deployment and operation. However, the practical fronthaul is often capacity constrained or time-delay constrained, which presents a bottleneck to the capacity of C-RANs [70]. To overcome the disadvantages of C-RANs with the fronthaul constraints, H-CRANs are proposed as a potential solution [71]. Compared to the C-RAN architecture, the proposed H-CRAN alleviates the fronthaul requirements with the participation of high power node (e.g., macro or micro base station). The user and control planes are decoupled in H-CRANs. HPNs execute the functions of the control plane, which deliver all control signaling and system broadcasting data to UEs. HPNs are connected to the BBU pool via the backhaul links, which alleviates constraints on fronthaul. However, H-CRANs still have same challenges in practice. More and more IoT applications generate unprecedented amounts of data at the edge of network, which worsens the fronthaul constraints. In additions, processing and storage capabilities in edge devices has not been fully exploited, which is a promising approach to successfully alleviate the burden of the fronthaul and BBU pool. Taking full advantage of fog computing and C-RANs, fog radio access networks (F-RANs) have been proposed to tackle these aforementioned disadvantages of C-RANs and H-CRANs [72]. In F-RANs, local radio signal processing, cooperative radio resource management, and distributed storing capabilities in edge devices, which can decrease the heavy burden on fronthaul and avoid large-scale radio signal processing in the centralized baseband unit pool. Therefore, F-RANs can achieve high SE/EE, low latency, and fantastic reliability to meet 5G requirements.

Table II shows the differences between MEC and F-RANs. First, F-RANs have been proposed as an enhancement and evolution of C-RANs to overcome the disadvantages of C-RANs with the fronthaul constraints. MEC is based on a virtualized platform to enable an open radio access network which can host third party innovative applications and content at the edge of the network. Operators can open their networks to authorized third parties, exposing capabilities to the over the top (OTT) players and application developers to flexibly and rapidly deploy innovative applications and services towards mobile subscribers, enterprises and vertical segments. MEC does not contradict with C-RANs but rather complement them. MEC is mainly compute oriented, which enables running computation-intensive tasks for edge users by offloading energy consuming computations of the applications to the MEC server to enhance user experience. In F-RANs, the cooperative radio signal processing (CRSP) and cooperative radio resource management (CRRM) can be executed in fog-computing-based access point (F-AP), and the edge caching in F-APs is a key technology to improve SE and EE under maintaining a low latency level [73]. In addition, MEC sever is compatible with the traditional wireless network architecture, which can be deployed at multiple locations, such as at the eNodeB, RNC, and RAT cell aggregation sites. The system architecture of F-RANs is evolved from HetNets and C-RANs, and F-AP is introduced to integrate not only the front radio frequency (RF), but also the limited caching, the local distributed CRSP and simple CRRM functions capabilities.

TABLE II

THE DIFFERENCE BETWEEN MEC AND F-RANS

	MEC	F-RANs
Motivation	enable an open radio access network which can host third party innovative applications and content at the edge of the network	to overcome the disadvantages of the fronthaul constraints with limited capacity and long delay
The relationship with C-RANs	complement with C-RANs	an enhancement and evolution of C-RANs
Key technology	computation offloading	edge caching
Deployment Scheme	be compatible with traditional wireless network architecture	an new system architecture is evolved from HetNets and C-RANs by introducing F-AP

V. OPEN ISSUES AND CHALLENGES

Table III summarizes the main features of these three edge computing technologies. Fog computing is initiated to address some challenges in meeting new requirements about IoTs. MEC is recognized as one of the key technologies to meet 5G requirements. Cloudlets is proposed to address some challenges in mobile computing. From the application point of view, MEC enables an open radio access network which can host third party innovative applications and content at the edge of the network. Cloudlet enables new classes of mobile applications that are both compute-intensive and latency-sensitive in an open ecosystem based on cloudlets. Fog computing enables high-performance, interoperability and security in a multi-vendor fog computing based ecosystem. The similarities of the three technologies are openness. Operators open their networks to third parties to deploy innovative applications and services.

TABLE III

COMPARISON OF CLOUDLETS, MEC, AND FOG COMPUTING

Item	Mobile Edge Computing	Cloudlets	Fog computing
Organization	ETSI MEC supported by Huawei, IBM, Intel, Nokia Networks, NTT DoCoMo and Vodafone	OEC launched by Vodafone, Intel, Huawei and Carnegie Mellon University	OpenFog Consortium founded by ARM, Cisco, Dell, Intel, Microsoft and Princeton University
Is the corresponding consortium a standards body?	Yes	No. (influences other standards organizations such as ETSI MEC)	No. (OpenFog has an affiliation agreement with IEEE and will be establishing liaisons with other standards organizations, with the objective of collaborating in the creation of standards)
Which business interests are driving?	5G requirements in the telecommunications industry	some applications based on mobile computing	Internet of Things
Motivation from the application point of view	enable an open radio access network which can host third party innovative applications and content at the edge of the network	enable new classes of mobile applications that are both compute-intensive and latency-sensitive in an open ecosystem based on cloudlets	enable high-performance, interoperability and security in a multi-vendor fog computing based ecosystem
Features about openness	Operators open their networks to third parties to deploy innovative applications and services	OPENSTACK++ is an open source platform that extends the functionality of Openstack to support cloudlets	The OpenFog Reference Architecture is used as a common baseline for achieving a multivendor interoperable fog computing ecosystem

In this section, some of the pertinent open issues which require additional investigations for edge computing are summarized.

A. Big Data Mining in Edge Computing

To adapt the massive kinds of packet traffic and the time-varied radio channel, the edge computing should be information-aware and reconfigured. The big data, described by volume, variety, velocity, and value, includes subscriber-level, cell-level, core-network-level, and other level data, can facilitate the network towards a more proactive one [74]. Owing to the fast development of big data mining, it is feasible to utilize big data technology to extract interesting patterns or knowledge to enhance the self-organizing capabilities in edge computing. The hierarchical data mining techniques should be used. The transmission of the large volume of data collected by edge devices puts a heavy burden on the fronthaul/backhaul, as a result, the data mining can be pre-executed in the edge devices. In the centralized cloud computing, these pre-executed information is reprocessed. Meanwhile, the computing of sparse, uncertain and incomplete data is a big problem, which requires advanced data mining algorithms [75].

B. Network Slicing in Edge Computing

To meet the diverse use cases and business models for the emerging applications of mobile Internet and IoT, both revolutionary wireless network architectures and advanced technologies are anticipated. As a result, network slicing is proposed recently to flexibly provide a soft-defined networking in a cost-efficient way. In the concept of network slicing [76], the network entity are sliced into multiple isolated network slice instances, and each slice instance has appropriate network functions and uses advanced radio access technologies for a specific use case or business model. By exploring software defined networking and network function virtualization, network slice instances and the isolation between them can be conveniently realized [77]. Despite the evident attractive advantages in the centralized cloud computing, network slicing comes with several severe challenges when applying in edge computing. First, the conventional creation of network slice instance is mainly business driven. The network slicing solution mainly addresses the needs of different services, which does not highlight the characteristics of edge computing on network slicing creation. For example, when the radio resource in RANs is in shortage, the requested network slicing may not be effective. As a result, the network slicing should consider the radio transmission impacts, and the corresponding network slicing jointly considering the status of RANs should be defined. Second, most of the existing work on network slicing is purely based on CNs, while network slicing as an end-to-end solution should cover the specific characteristics of RANs. To overcome these challenges, a framework solution of a new network slicing technique for edge computing is anticipated.

VI. Conclusion

This paper outlines and surveys the state-of-the-art edge computing technologies. With the goal of understanding further intricacies of the key technologies, we have broadly divided the body of knowledge into cloudlet, mobile edge computing, and fog computing. Within each of these aspects, we have given a detailed tutorial on the principle, system architecture, standards, and applications. Nevertheless, given the relative infancy of the field, there are still quite a number of outstanding problems that need further investigation from the perspective of key techniques and advanced solutions. Given the extensiveness of the research areas, it is also concluded that more rigorous investigations are required with greater attention to be focused on transforming well established fog computing into the fog computing bases radio access networks. Furthermore, with the introduction of the advanced big data mining and network slicing, the availability of varied degrees of freedom along with the associated constraints further beckon the design and validation of the original models in the context of edge computing.

References

[1] M. Peng, K. Zhang, Recent advances in fog radio access networks: Performance analysis and radio resource allocation, IEEE Access J. 4 (2016) 5003-5009.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications, IEEE Commun. Surv. Tutorials. 17 (2015) 2347–2376.

[3] CISCO, The Internet of Things how the next evolution of the Internet is changing everything, White Pap. (2011). http://www.cisco.com/c/dam/en us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

[4] M. Chiang, T. Zhang, Fog and IoT: An Overview of Research Opportunities, IEEE Internet Things J. 3 (2016) 854–864.

[5] F. Ganz, D. Puschmann, P. Barnaghi, F. Carrez, A Practical Evaluation of Information Processing and Abstraction Techniques for the Internet of Things, IEEE Internet Things J. 2 (2015) 340–354.

[6] M.A. Razzaque, M. Milojevic-Jevric, A. Palade, S. Cla, Middleware for internet of things: A survey, IEEE Internet Things J. 3 (2016) 70–95.

[7] A. Brogi, S. Forti, QoS-aware deployment of IoT applications through the fog, IEEE Internet of Things Journal, (2017).

[8] J. Lin; W. Yu; N. Zhang; X. Yang; H. Zhang; W. Zhao, A survey on internet of things: architecture, enabling Technologies, security and privacy, and applications, IEEE Internet of Things Journal, (2017).

[9] M. Satyanarayanan, The emergence of edge computing, Computer (Long. Beach. Calif). 50 (2017) 30–39.

[10] N. Bizanis, F. Kuipers, SDN and virtualization solutions for the Internet of Things: A survey, IEEE Access. J.4 (2016) 5591-5606.

[11] J. Li, M. Peng, A. Cheng, Y. Yu, C. Wang, Resource allocation optimization for delay-sensitive traffic in fronthaul constrained cloud radio access networks, IEEE Syst. J. PP (2014) 1-12.

[12] ETSI, Mobile-edge computing introductory technical white paper, White Paper, Mobile-edge Computing Industry Initiative, (2014). https://portal.etsi.org/portals/0/tbpages/mec/docs/mobile-edge_computing_–

_introductory_technical_white_paper_v1

[13] Y. Mao, C. You, J. Zhang, K. Huang, K. B. Letaief, A survey on mobile edge computing: the communication perspective, (2017). https://arxiv.org/abs/1701.01090

[14] Guenter I. Klas, Fog computing and mobile edge cloud gain momentum open fog consortium, ETSI MEC and Cloudlets, (2015). http://yucianga.info/?p=938

[15] M. Satyanarayanan, P. Bahl, R. Caceres, N. Davies, The Case for VM-Base Cloudlets in Mobile Computing, Pervasive Comput. 8 (2009) 14–23.

[16] Nokia Solutions and Networks, Increasing Mobile Operators' Value Proposition With Edge Computing, White Pap. (2013) 1–6. http://nsn.com/portfolio/liquid-net/intelligent-broadband-management/liquid-applications.

[17] Y. Ch. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, Mobile Edge Computing a key technology towards 5G, (2015). http://10.3.200.202/cache/8/03/etsi.org/6e14a9668574b8b935111768d9f6e501/etsi_wp11_mec_a_key_technology_towards_5g.pdf

[18] N. Sprecher, J. Friis, R. Dolby, and J. Reister, Edge computing prepares for a multi-access future, (2016). http://www.telecomtv.com/articles/mec/edge-computing-prepares-for-a-multi-access-future-13986/

[19] F. Bonomi, R. Milito, J. Zhu, S. Addepalli, Fog Computing and Its Role in the Internet of Things, Proc. First Ed. MCC Work. Mob. Cloud Comput. (2012) 13–16.

[20] Z. Sanaei, S. Abolfazli, A. Gani, R. Buyya, Heterogeneity in mobile cloud computing: Taxonomy and open challenges, IEEE Commun. Surv. Tutorials. 16 (2014) 369–392.

[21] K. Ha, P. Pillai, W. Richter, Y. Abe, M. Satyanarayanan, Just-in-time provisioning for cyber foraging, MobiSys 2013 - Proc. 11th Annu. Int. Conf. Mob. Syst. Appl. Serv. (2013) 153–166.

[22] Y. Gao, W. Hu, K. Ha, B. Amos et al., Are cloudlets necessary? School of Computer Science Carnegie Mellon University Pittsburgh, (2015).

[23] K. Ha, Y. Abe, Z. Chen, W. Hu et al., Adaptive VM handoff across cloudlets, Technical Report CMU-CS-15-113, CMU School of Computer Science, (2015).

[24] M. Satyanarayanan, R. Schuster, M. Ebling, G. Fettweis, H. Flinck, K. Joshi, K. Sabnani, An open ecosystem for mobile-cloud convergence, IEEE Commun. Mag. 53 (2015) 63–70.

[25] K. Ha, Z. Chen, W. Hu, W. Richter, P. Pillai, M. Satyanarayanan, Towards wearable cognitive assistance, in: Proc. 12th Annu. Int. Conf. Mob. Syst. Appl. Serv. - MobiSys '14, 2014: pp. 68–81.

[26] M. Satyanarayanan, P. Simoens, Y. Xiao, P. Pillai, Z. Chen, K. Ha, W. Hu, B. Amos, Edge analytics in the internet of things, IEEE Pervasive Comput. 14 (2015) 24–31.

[27] M. Satyanarayanan, G. Lewis, E. Morris, S. Simanta, J. Boleng, K. Ha, The role of cloudlets in hostile environments, IEEE Pervasive Comput. 12 (2013) 40–49.

[28] K. Ha, M. Satyanarayanan, OpenStack++ for cloudlet deployment, School of Computer Science Carnegie Mellon University Pittsburgh, (2015).

[29] ETSI GS MEC 001, Mobile Edge Computing (MEC) Terminology V1.1.1, (2016). http://www.etsi.org/deliver/etsi_gs/MEC/001_099/001/01.01.01_60/gs_MEC001v010101p.pdf

[30] ETSI GS MEC 002, Mobile Edge Computing (MEC) Technical Requirements V1.1.1, (2016). http://www.etsi.org/deliver/etsi_gs/MEC/001_099/002/01.01.01_60/gs_MEC002v010101p.pdf

[31] ETSI GS MEC 003, Mobile Edge Computing (MEC) Framework and Reference Architecture V1.1.1, (2016). http://www.etsi.org/deliver/etsi_gs/MEC/001_099/003/01.01.01_60/gs_MEC003v010101p.pdf

[32] ETSI GS MEC-IEG 004, Mobile Edge Computing (MEC) Service Scenarios V1.1.1, (2015). http://www.etsi.org/deliver/etsi_gs/MEC/001_099/004/01.01.01_60/gs_MEC003v010101p.pdf

[33] ETSI GS MEC-IEG 005, Mobile Edge Computing (MEC) Proof of Concept Framework V1.1.1, (2015).

http://www.etsi.org/deliver/etsi_gs/MEC-IEG/001_099/005/01.01.01_60/gs_MEC-IEG005v010101p.pdf

[34] ETSI GS MEC-IEG 006: Mobile Edge Computing Market Acceleration MEC Metrics Best Practice and Guidelines V1.1.1, (2017) http://www.etsi.org/deliver/etsi_gs/MEC-IEG/001_099/006/01.01.01_60/gs_MEC-IEG006v010101p.pdf

[35] T. Taleb, S. Dutta, A. Ksentini, M. Iqbal and H. Flinck, Mobile edge computing potential in making cities smarter, IEEE Commun. Mag. 55 (2017) 38-43.

[36] S. Lin, H.F. Cheng, W. Li, Z. Huang, P. Hui, C. Peylo, Ubii: Physical World Interaction Through Augmented Reality, IEEE Trans. Mob. Comput. 16 (2017) 872–885.

[37] N. Kumar, S. Zeadally, J.J.P.C. Rodrigues, Vehicular delay-tolerant networks for smart grid data management using mobile edge computing, IEEE Commun. Mag. 54 (2016) 60-66.

[38] X. Sun and N. Ansari, EdgeIoT: mobile edge computing for the Internet of Things, IEEE Commun. Mag. 54 (2016) 22-29.

[39] A. Ceselli, M. Premoli, S. Secci, Mobile edge cloud network design optimization, IEEE/ACM Trans. Netw. 25 (2017) 1818-1831.

[40] S. Barbarossa, S. Sardellitti, P. Di Lorenzo, Communicating while computing: Distributed mobile cloud computing over 5G heterogeneous networks, IEEE Signal Process. Mag. 31 (2014) 45–55.

[41] P. Mach, Z. Becvar, Mobile Edge Computing: a survey on architecture and computation offloading, IEEE Commun. Surv. Tutorials. (2017).

[42] H. Flores, P. Hui, S. Tarkoma, Y. Li, S. Srirama, R. Buyya, Mobile code offloading: From concept to practice and beyond, IEEE Commun. Mag. 53 (2015) 80–88.

[43] K. Zhang, Y. Mao, S. Leng, Q. Zhao, L. Li, X. Peng, L. Pan, S. Maharjan, Y. Zhang, Energy-efficient offloading for mobile edge computing in 5G heterogeneous networks, IEEE Access. J. 4 (2016) 5896-5907.

[44] X. Chen, L. Jiao, W. Li, X. Fu, Efficient multi-user computation offloading for mobile-edge cloud computing, IEEE/ACM Trans. Netw. 24 (2016) 2795-2808.

[45] C. You, K. Huang, H. Chae, B.H. Kim, Energy-Efficient Resource Allocation for Mobile-Edge Computation Offloading, IEEE Trans. Wirel. Commun. 16 (2017) 1397–1411.

[46] S. Sardellitti, G. Scutari, S. Barbarossa, Joint optimization of radio and computational resources for multicell mobile-edge computing, IEEE Trans. Signal Inf. Process. over Networks. 1 (2015) 89–103.

[47] S. Secci, P. Raad and P. Gallard, Linking virtual machine mobility to user mobility, IEEE Trans. Netw. Serv. Manag. 13 (2016) 927–940.

[48] P. Raad, S. Secci, D.C. Phung, A. Cianfrani, P. Gallard, G. Pujolle, Achieving sub-second downtimes in large-scale virtual machine migrations with LISP, IEEE Trans. Netw. Serv. Manag. 11 (2014) 133–143.

[49] T. Taleb, A. Ksentini, P. Frangoudis, Follow-Me Cloud: When Cloud Services Follow Mobile Users, IEEE Trans. Cloud Comput. (2016) 1–1.

[50] X. Sun, N. Ansari, PRIMAL: PRofIt Maximization Avatar pLacement for mobile edge computing, in: 2016 IEEE Int. Conf. Commun. ICC 2016, (2016) 1-6.

[51] J. Plachy, Z. Becvar, E.C. Strinati, Dynamic resource allocation exploiting mobility prediction in mobile edge computing, in: IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC, 2016.

[52] Open Fog Consortium, [Online]. https://www.openfogconsortium.org/

[53] OpenFog Consortium Architecture Working Group, OpenFog architecture overview white paper, https://www.openfogconsortium.org/wp-content/uploads/OpenFog-Architecture-Overview-WP-2-2016.pdf

[54] Open Fog Consortium, OpenFog reference architecture for fog computing, https://www.openfogconsortium.org/wp-content/uploads/OpenFog_Reference_Architecture_2_09_17-FINAL-1.pdf

[55] A. V. Dastjerdi and R. Buyya, Fog computing: helping the Internet of Things realize its potential, Computer. 49 (2016) 112–116.

[56] S. K. Sharma, X. Wang, Live data analytics with collaborative edge and cloud processing in wireless IoT networks, IEEE Access, 5(2017) 4621-4635.

[57] B. Tang, Z. Chen, G. Hefferman, S. Pei, W. Tao, H. He, Q. Yang, Incorporating intelligence in fog computing for big data analysis in smart cities, IEEE Trans. Ind. Informatics. (2017).

[58] S. Basudan, X. Lin, K. Sankaranarayanan, A privacy-preserving vehicular crowdsensing based road surface condition monitoring system using fog computing, IEEE Internet Things J. 4 (2017) 772–782.

[59] P. Hu, H. Ning, T. Qiu, H. Song, Y. Wang, X. Yao, Security and Privacy Preservation Scheme of Face Identification and Resolution Framework Using Fog Computing in Internet of Things, IEEE Internet Things J. (2017) 1–1.

[60] X. Hou, Y. Li, M. Chen, D. Wu, D. Jin, S. Chen, Vehicular Fog Computing: A Viewpoint of Vehicles as the Infrastructures, IEEE Trans. Veh. Technol. 65 (2016) 3860–3873.

[61] M. Aazam, E.N. Huh, Fog Computing: The Cloud-IoT/IoE Middleware Paradigm, IEEE Potentials. 35 (2016) 40–44.

[62] L. Gu, D. Zeng, S. Guo, A. Barnawi, Y. Xiang, Cost efficient resource management in fog computing supported medical cyber-physical system, IEEE Trans. Emerg. Top. Comput. 5 (2017) 108-119.

[63] S. Misra, S. Sarkar, Theoretical modelling of fog computing: a green computing paradigm to support IoT applications, IET Networks. 5 (2016) 23–29.

[64] S. Sarkar, S. Chatterjee, S. Misra, Assessment of the suitability of fog computing in the context of internet of things, IEEE Trans. Cloud Comput. PP (2015).

[65] F. Jalali, K. Hinton, R. Ayre, T. Alpcan, R.S. Tucker, Fog computing may help to save energy in cloud computing, IEEE J. Sel. Areas Commun. 34 (2016) 1728–1739.

[66] R. Deng, R. Lu, C. Lai, T.H. Luan, H. Liang, Optimal Workload Allocation in Fog-Cloud Computing Toward Balanced Delay and Power Consumption, in: IEEE Internet Things J., 2016: pp. 1171–1181.

[67] X. Masip-Bruin, E. Marín-Tordera, G. Tashakor, A. Jukan, G.J. Ren, Foggy clouds and cloudy fogs: A real need for coordinated management of fog-to-cloud computing systems, IEEE Wirel. Commun. 23 (2016) 120–128.

[68] M. Peng, Y. Li, Z. Zhao, C. Wang, System architecture and key technologies for 5G heterogeneous cloud radio access networks, IEEE Netw. 29 (2015) 6–14.

[69] M. Peng, Y. Sun, X. Li, Z. Mao, C. Wang, Recent Advances in Cloud Radio Access Networks: System Architectures, Key Techniques, and Open Issues, IEEE Commun. Surv. Tutorials. 18 (2016) 2282–2308.

[70] M. Peng, C. Wang, V. Lau, H.V. Poor, Fronthaul-constrained cloud radio access networks: Insights and challenges, IEEE Wirel. Commun. 22 (2015) 152–160.

[71] M. Peng, Y. Li, J. Jiang, J. Li, C. Wang, Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies, IEEE Wirel. Commun. 21 (2014) 126–135.

[72] M. Peng, S. Yan, K. Zhang, C. Wang, Fog-computing-based radio access networks: Issues and challenges, IEEE Netw. 30 (2016) 46–53.

[73] S. Jia, Y. Ai, Z. Zhao, M. Peng and C. Hu, Hierarchical content caching in fog radio access networks: ergodic rate and transmit latency, China Commun. 13 (2016) 1–14.

[74] A. Imran, A. Zoha, Challenges in 5G: How to empower SON with big data for enabling 5G, IEEE Netw. 28 (2014) 27–33.

[75] X. Wu, X. Zhu, G.Q. Wu, W. Ding, Data mining with big data, IEEE Trans. Knowl. Data Eng. 26 (2014) 97–107.

[76] R. Hattachi and J. Erfanian, 5G white paper, NGMN Alliance, (2015).

[77] Q. Li, G. Wu, A. Papathanassiou, and U. Mukherjee, An end-to-end network slicing framework for 5G wireless communication systems, (2016) https://arxiv.org/abs/1608.00572