Hierarchical Edge Cloud Enabling Network Slicing for 5G Optical Fronthaul

Chuang Song, Min Zhang, Yueying Zhan, Danshi Wang, Luyao Guan, Wei Liu, Lin Zhang, and Siya Xu

***Abstract*—Network slicing has been proposed as an innovative concept to partition a physical network into different configurable slices, thereby allowing the network to meet the diverse requirements of expected services in the 5G era. Existing network-slicing schemes focus almost exclusively on the partitioning of the core network or the radio access network. However, to achieve greater flexibility and better resource utilization in service provisioning, the potential of network slicing in the 5G optical mobile fronthaul (MFH) network should be jointly exploited. To address this challenge, we propose a flexible hierarchical edge cloud architecture to enable 5G optical MFH network slicing. Additionally, a corresponding integrated network resource management scheme is proposed to support various quality-of-service (QoS) requirements and diverse resource requests of the network slices. Simulation results indicate that the proposed scheme is able to jointly allocate the bandwidth resources to network slices and realize cloud-computing offloading to meet various QoS requirements, and therefore reduce the fronthaul bandwidth burden.**

***Index Terms*—Cloud computing offloading; Edge cloud; Network resource management functional split; Network slicing.**

I. Introduction

With the rapid development of cloud computing, the cloud/centralized radio access network (C-RAN) architecture is viewed as a promising solution for mobile networks in the 5G era [1], where centralized baseband units (BBUs) and the remote radio heads (RRHs) are connected by optical mobile fronthaul (MFH) links [2]. The centralized BBUs can be used as shared resources, reducing operating and capital expenditures and energy consumption [3,4]. However, legacy network architectures were designed for mobile broadband consumers without considering the characteristics of emerging 5G scenarios, such as massive machine-type communications (mMTC) and ultra-reliable low-latency communications (uRLLC) [5]. With the deployment of diverse use cases in 5G networks, a tremendous amount of traffic will be generated at the network edge with a wide range of quality-of-service (QoS) requirements in terms of latency, bandwidth, and reliability [6]. The conventional fully centralized C-RAN design follows a “one size fits all” architectural approach to support all types of traffic, which makes it difficult to address the wide range of QoS requirements [7]. In addition, a fully centralized architecture imposes very high capacity requirements on the MFH links [8].

In order to efficiently accommodate different use cases over the same network infrastructure while reducing the bandwidth requirements of optical fronthaul, 5G networks require architectural optimization and restructuring with respect to current deployments. Network slicing has been proposed recently to provide flexible networks in a cost-efficient manner [9]. Network slicing constructs different logical networks (i.e., slices) on a unified physical infrastructure via software-defined networking (SDN), network function virtualization (NFV), and cloud computing [10,11]. Network slicing allows the support of highly diverse services, with possibly contradictory requirements, on a single infrastructure through independent configuration of each network slice [12]. Each slice in such an architecture is an end-to-end virtualized network instance [3], and is tailored in terms of resources to meet the requirements of the services. Here, we define a virtualized network instance as a set of run-time network functions, and resources to run these functions, forming a complete instantiated logical network to meet the required network characteristics.

It is important to note that NFV plays a vital role in network slicing. Network slicing needs to leverage NFV technology to realize both infrastructure virtualization and NFV. The infrastructure and network functions will be appropriately instantiated, connected, and combined over the underlying 5G substrate networks by network-slicing technology [12]. Moreover, note that network slicing is different from another virtual network technology, virtual private networks (VPNs). Network slicing not only needs to isolate different networks as in VPN, but also needs to provide computing resources and storage resources for the different network slices by virtualizing the underlying physical infrastructure.

Network slicing has attracted extensive attention from academia and industry. We summarize some of the existing

Manuscript received September 20, 2018; revised February 16, 2019; accepted February 18, 2019; published March 26, 2019 (Doc. ID 346290).

C. Song, M. Zhang (e-mail: mzhang@bupt.edu.cn), D. Wang, L. Guan, W. Liu, and L. Zhang are with the State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications (BUPT), Beijing 100876, China.

Y. Zhan is with the Technology and Engineering Center for Space Utilization, Chinese Academy of Science (CAS), Beijing 100094, China.

S. Xu is with the State Key Laboratory of Networking and Switching Technology, BUPT, Beijing 100876, China.

https://doi.org/10.1364/JOCN.11.000B60

research here; more detailed coverage of prior work is provided in Section II.

Various works have proven the potential gains of deploying network slicing [13,14]. Many schemes address RAN-based network slicing [15–17]. In [15], a novel network-slicing architecture for integrated 5G communications is presented, featuring a heterogeneous wireless domain. In [16], a flexible RAN slicing architecture named Orion is presented, which can be deployed dynamically over the network infrastructure, coexisting in a fully isolated manner in terms of radio resources and control functions. In [17], a framework is reported to describe the configuration descriptors and the extended protocol features within a next-generation RAN (NG-RAN) node for RAN slicing.

There is also existing research focusing on the management and orchestration of network functions [18,19]. For example, an end-to-end (E2E) slicing orchestration system is presented in [20,21]. Furthermore, there is a lot of existing work on core network (CN)-based network slicing [22–24]. In [22], two examples of network slicing are illustrated to explain the impact of use case requirements on the network-slice design. In [23], the authors present an in-operation network-slice resource allocator for data center networks and core networks, which is able to consider networking and cloud infrastructure. In [24], the authors present a solution for the 5G optical transport slicing problem in terms of both mixed integer linear programming (MILP) formulations and heuristic algorithms.

However, the reported work above focuses almost exclusively on the access or core segment of the network, while the MFH segments used to transport data between access and core have received little attention. Note that the edge cloud is emerging as a promising solution for satisfying diverse QoS requirements, such as low latency and high reliability, for different service types in 5G. Conventional CN-based network slicing cannot take full advantage of the characteristics of the edge cloud to support mMTC and uRLLC service types efficiently. Edge clouds are deployed at the edge of the metro network, closer to end users. From the perspective of network architecture evolution, the edge cloud needs to be highly integrated with the optical MFH network and C-RAN to support 5G low-latency services. The investigation of network slicing in the fronthaul network is still in its infancy. Specifically, the impact of edge-cloud-based optical MFH on different network-slicing deployment scenarios is not well studied.

The integration of MFH and edge cloud is a critical challenge that requires efficient coordination and dynamic provisioning of heterogeneous network resources. In this paper, to cope with the aforementioned issues of network slicing, we propose a hierarchical edge cloud (HEC)-based MFH network-slicing framework. The HEC can effectively shift the remote processing resources closer to the end users to implement cooperative resource management and enhance the responsiveness (and perhaps resilience) of dynamic end-to-end 5G service demands. Furthermore, we address functional split schemes based on HEC, in which different baseband functions can be deployed on different tiers of cloud servers to satisfy diverse QoS requirements of different slices optimally.

The contributions of this paper are summarized as follows:

First, we propose an HEC architecture for optical MFH network slicing, in which a two-tier HEC is deployed at the BBU and RRH portions of the network. Different services can be forwarded and processed in the corresponding edge cloud.

Second, we propose an integrated network resource management (INRM) scheme that jointly considers cloud-computing offloading and MFH bandwidth allocation. The focus is on network slicing and bandwidth reduction for the HEC architecture applied to 5G optical fronthaul. The scheme is able to satisfy the QoS requirements of different network-slicing configurations and reduce the MFH bandwidth burden.

The remainder of this paper is organized as follows. Section II presents a more comprehensive list of existing related work. Section III illustrates the proposed network architecture and signaling procedure of the slicing management. We then elaborate on the analysis formulations and our proposed INRM scheme in Section IV. Section V provides simulation results to evaluate our design objectives, and analyzes the results. Section VI presents open issues and future research. Finally, conclusions are presented in Section VII.

II. Related Work

As noted above, existing network-slicing research mainly focuses on the following aspects: RAN slicing, CN slicing, E2E slicing, and management and network orchestration (MANO).

RAN-based network slicing can be traced back to earlier research on active RAN sharing [25]. Many papers focus on radio resource virtualization and sharing [26–28]. In [26], the authors have analyzed a "share-constrained proportional allocation" framework for network slicing. The authors of [27] analyzed the RAN slicing problem in a multi-cell network in relation to the radio resource management (RRM) functionalities that can be used to support splitting the radio resources among the RAN slices. In [28], the network-slicing process is formulated as a weighted throughput maximization problem that involves sharing of computational resources, fronthaul capacity, physical RRHs, and radio resources. In addition, some research has focused on the convergence of network slicing and existing wireless technologies. The authors of [29] demonstrate its realization for the case of evolved LTE using state-of-the-art technologies. Software-defined wireless networking enabled WLAN slices are introduced in [30] to provide fine-grained spectrum.

Existing literature on RAN slicing focuses mainly on efficient sharing of resources; however, functional split is essential in the RAN slicing context, especially in the 5G era. Functional split is a promising technique to reduce the fronthaul bandwidth requirement. Different functional blocks can be designed as "distributed" and deployed in different nodes. Different functional split schemes have different processing costs and fronthaul transmission rates [31]. Hence, functional split should be considered when studying network slicing.

There has been significant progress on CN-based slicing to the point that it is fairly mature. The 3GPP SA group is discussing a solution to implement network slicing using the eDECOR concept [32]. The use of mature virtualization and cloud-computing technologies has led to systems that realize CN slicing [22,27,33]. In [22], the authors discussed the feasibility of designing a flexible and adaptive mobile core network based on functional decomposition and network-slicing concepts. The authors of [27] present a slice-based 5G architecture and "Network Store in a programmable cloud" that efficiently manages network slices; this work combines many existing paradigms, such as NFV, SDN, and cloud computing. In [33], the authors introduce a new architecture named V-core for mobile core networks, which takes advantage of both the SDN and NFV concepts. Moreover, the possibility of an evolved packet core (EPC) network as a service over the cloud has been explored [34]. The authors of [7] propose a model introducing DevOps for both optical network virtualization and network slicing.

Additionally, end-to-end slicing automated MANO testbeds are presented in [35–38]. The authors of [35] develop a proof of concept based on OpenAirInterface (OAI) to derive key performance results of network slicing. In [36], a slicing solution is proposed enabling the automation of end-to-end network-slice management and orchestration in multiple resource domains. This has several open-source MANO framework implementations. In [37], the authors present an example scenario that combines SDN and NFV technologies to address the realization of network slices. In the 5G-NORMA project, SDN and NFV technologies are exploited to enable dynamic sharing of network resources among operators [38].

Little existing research focuses on the MFH or on backhaul network slicing. In [39], the authors propose a unified control and network-slicing architecture for a multi-vendor multi-standard passive optical network (PON)-based 5G fronthaul network and experimentally demonstrate its flexible resource management and slicing capability. Moreover, they propose a multi-vendor network-slicing scheme for converged vehicular and fixed access networks in [40]. In [41], some implementations and experimental validations were conducted addressing specific 5G and fixed mobile converged (FMC) objectives within an aggregation (backhaul) network segment.

As stated earlier, more research is needed on the application of network slicing in optical MFH. Specifically, the edge-cloud-based optical MFH should be studied. In addition, functional split needs to be considered in conjunction with network slicing. The approach in this paper is to focus on the HEC-based MFH segments, while considering functional split.

III. Hierarchical Network Architecture and Network Resource Management

A. HEC Architecture Considering Function Split

In order to provide various services with desired QoS and reduce the capacity burden of MFH, a two-tier HEC architecture is proposed for optical MFH network slicing, as shown in Fig. 1. We define two types of clouds: BBU-clouds and RRH-clouds. It is emphasized that both types

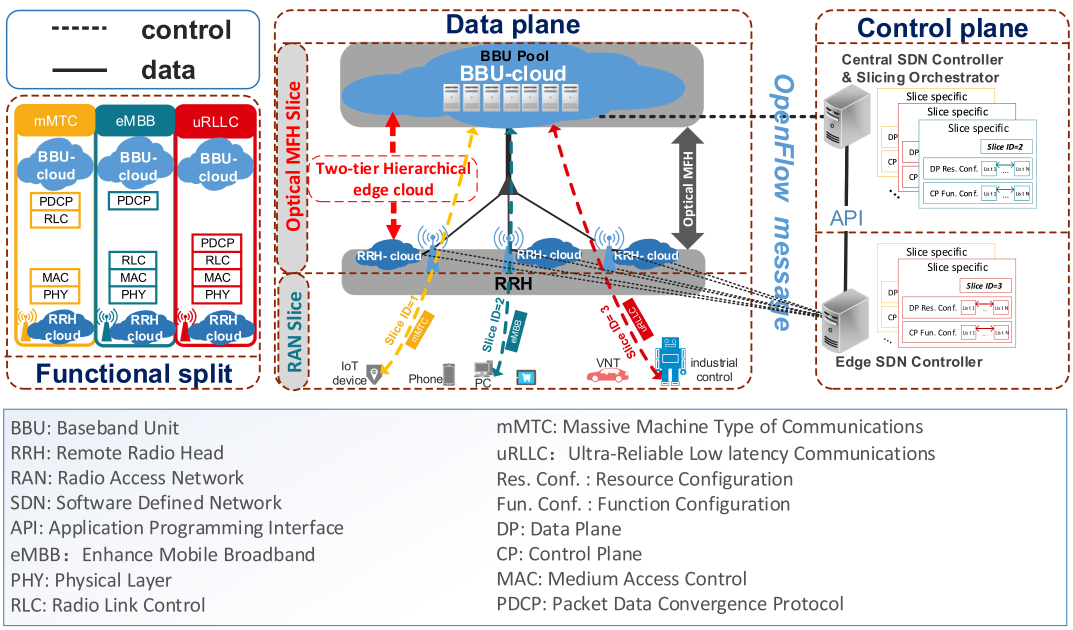

Fig. 1. Hierarchical edge-cloud-based optical MFH network architecture.

of clouds are located at the network edge. Specifically, the BBU-clouds are centralized and located in a BBU pool (or central office), whereas the RRH-clouds (i.e., the distributed RRHs) are closer to users. The capacity of RRH-clouds is far less than that of BBU-clouds. To avoid all traffic being offloaded directly from the RRHs to the BBU-cloud servers and to reduce the long transmission delay and heavy capacity burden on the fronthaul, a subset of centralized control functions and baseband processing is shifted from the BBU-clouds to the RRH-clouds (the particular partitioning of functions depends on the chosen functional split). Therefore, some delay-sensitive services are able to be delivered to the adjacent RRHs and processed at the local RRHs to reduce fronthaul bandwidth usage.

The proposed HEC architecture is an effective solution to realize network slicing and functional split. A fundamental feature of network slicing is the ability to provide independent network instances that support services with distinct requirements. In this regard, the HEC can facilitate the realization of the flexible network-slicing concept in the MFH network. In particular, as mentioned earlier, the functional split chosen between the RRH-cloud and BBU-cloud greatly affects the characteristics of network slices and the MFH bandwidth requirements. In the following, we elaborate on the conceptual connection between the flexible HEC architecture and MFH network slicing. We illustrate two levels of flexible function split by projecting them onto three major slice types, namely, eMBB, mMTC, and uRLLC. A detailed view is shown on the left side of Fig. 1. The definitions of the abbreviations used are shown below the figure.

The functional blocks mentioned below are used in 4G RAN, and the definition of functional split is provided by 3GPP for 5G. Thus, we describe the corresponding functional split options only briefly for the three network slices.

The eMBB slice requires large fronthaul bandwidth to support service such as 4K video and virtual reality (VR). As a result, eMBB mainly involves network functions that facilitate increasing the throughput. The RLC-PDCP split is adopted for eMBB slicing, because it can significantly reduce the optical fronthaul bandwidth requirements by forwarding MAC frame data instead of IQ data to the BBU-cloud [8]. In addition, the PDCP and PDCP split-bearer blocks are located at the BBU-cloud, facilitating the implementation of multi-connectivity and minimizing mobility signaling to the core network [42].

mMTC requires a large amount of data exchange in the network. Hence, a distributed implementation of the RLC functionality would result in a large amount of signaling overhead between the involved nodes. Moreover, the traffic is relatively small, which will not significantly increase the optical fronthaul bandwidth burden. Considering the latency requirements and deployment cost, the execution of the PDCP and RLC functionalities is envisioned as taking place at the BBU-cloud. Hence, MAC-RLC function split is chosen for the mMTC slice [43].

Regarding an uRLLC slice, applications with low-latency requirements are often not demanding in data rates, so we mainly focus on the ultra-low-latency requirements. The tight latency requirements for this type of slice favor its implementation as close to the radio access as possible. To this end, all functions are executed in the RRH-cloud for this type of service. In other words, the BBU-cloud is “unused” for the uRLLC slice.

B. Control Plane

To achieve orchestration and control, corresponding two-layer controllers are designed in the proposed MFH network-slicing architecture, i.e., the edge SDN controller (ESC) and the central SDN controller and slicing orchestrator (CSC&SO). Note that the CSC is integrated with the SO, as shown in Fig. 1.

The ESC abstracts the physical infrastructure and performs resource management at the network edge. The CSC&SO has similar functions as the ESC. In addition, the CSC&SO is responsible for dynamic provisioning of the slices and the resource management between the implemented network-slice instances with a global view. An application programming interface (API) is used to connect the ESC and CSC&SO to achieve signaling interaction and synchronization. The resource management scheme is configured by the CSC&SO and ESC in a coordinated manner.

In order to achieve security isolation among network slices, each access slice instance has specified control/data planes and operates as a logically separate network through NFV and SDN. As for the different network slices, significant performance gains can be achieved by configuring the specific network functions and transport resources. Specifically, the CSC has a global view of the network, which allows for optimizing the resource allocation and scheduling across multiple RRHs. When user equipment (UE) requests service resources, the ESC and CSC&SO will identify the traffic type (i.e., by slicing ID recognition [44], artificial intelligence [45], big data analysis [46], or deep packet inspection (DPI) [47]) and then determine how to place traffic load among different tiers of clouds and allocate the network resources so as to realize different adaptive traffic offloading strategies.

C. Signaling Procedure for Slicing Management

The network-slicing management strategies between the UE and HECs are able to realize the QoS requirements of different slices. The signaling procedures of the slicing management are presented in Fig. 2. Network-slicing management can be performed in the ESC and/or the CSC&SO according to the QoS requirements, real-time network status, cloud resources, and bandwidth utilization. This information is perceived by the ESC from the cloud nodes and other transport/process nodes in the MFH network and forwarded to the CSC&SO for synchronization.

Specifically, for the mMTC slice and eMBB slice, UE initiates a *Resource Request* to the controller (*step 1 in blue*). The ESC will collect and identify the slicing type and then

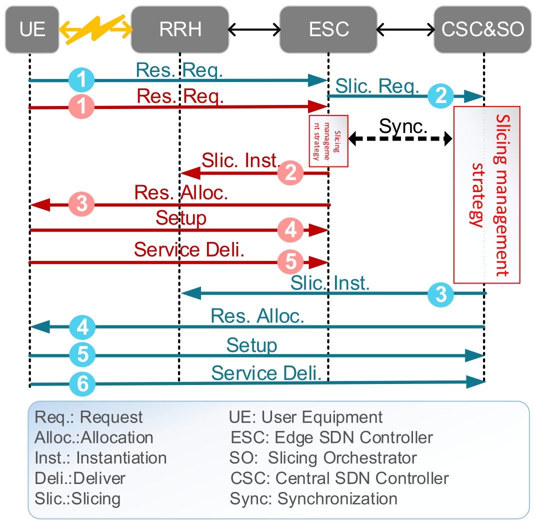

Fig. 2. Signaling procedure for the slicing management.

forward a *Slicing Request* to the CSC&SO via the API (*step 2 in blue*). The CSC&SO will perform the slicing management strategy by configuring the baseband functions and allocating network resources (i.e., cloud-computing resources, storage resources, and bandwidth resources), and then perform *slicing instantiation* (*step 3 in blue*). When the *resource allocation* (*step 4 in blue*) is fulfilled and the resource request is accommodated successfully, the UE and RRHs can send successful replies, i.e., *setup*, to the CSC&SO (*step 5 in blue*). Finally, *service delivery* begins (*step 6 in blue*).

As for the uRRLC slice, the ESC receives the resource *Requests* and identifies the delay-sensitive requirements (*step 1 in red*). Then the ESC instantiates *slicing* (*step 2 in red*) instead of forwarding the requests to the CSC&SO. When the *resource allocation* (*step 3 in red*) is fulfilled, the UE sends *setup* to the ESC (*step 4 in red*). Finally, the delay-sensitive traffic is handled in local or adjacent sites (*step 5 in red*). This process can significantly reduce the latency and thus meet the very strict latency requirements of an uRLLC slice. If the RRH-clouds do not have sufficient compute or storage resources to cope with the slicing resource requests, the ESC will forward the requests to the CSC&SO, which will allocate the required resource to the slice. Therefore, the CSC&SO has a high-level view of how to manage resources and instantiate slicing dynamically.

IV. Analysis

In the HEC-based optical MFH network, multiple types of resources, namely, the wireless network, optical network, and cloud processing resources, are deployed. In this section, based on the proposed network architecture, we propose an INRM to realize network resource management and optimization. The INRM jointly considers the optical bandwidth resource allocation and hierarchical cloud resource offloading to guarantee the QoS demands of different network slices while reducing the MFH bandwidth burden.

The network-slicing problem is a combined optimization problem of placing network functions over a set of candidate locations and deciding on their interconnections.

A. Network Modeling

We are given a physical network $G = (V, E, B, D)$, where $V = \{v_1, v_2, \ldots, v_n\}$ is a set of nodes and $E = \{e_1, e_2, \ldots, e_m\}$ is a set of links. For formulation purposes, V includes all RRH nodes, optical nodes, RRH-cloud nodes, and the BBU-cloud connected to the fronthaul network, namely, all the nodes shown in Fig. 1. E consists of not only fronthaul links but also links between the RRH-cloud and BBU-cloud. Nodes $v \in V$ and links $e \in E$ are characterized by capacities $B_v \geq 0$ and $B_e \geq 0$, respectively.

In the online operation of the network, B plays the role of residual capacity. That is, B represents the remaining resources of the nodes or links after subtracting out the current utilization. Each node v and link e is also associated with a cost in terms of latency, D_v and D_e, respectively. In the RRH-cloud and BBU-cloud servers, two time-varying cloud processing parameters indicate the status of computing and storage resources, which are composed of the CPU usage U_c and storage utilization U_m. From the perspective of the network, the parameter of concern is the occupied network bandwidth weight W_e of each link, which is related to the traffic cost of the corresponding link. From the users' perspective, the QoS experience is more important than which server provides the services. Therefore, each request from the source node s is translated into the needed network bandwidth and cloud resources. Note that these resources include the required optical bandwidth o and cloud resources c. We denote the ith traffic request as $\mathrm{TR}_i(s, o, c)$. Additionally, depending on the connection request and status of resources, the appropriate cloud server can be selected as the destination node based on the strategy.

B. Integrated Network Resource Management

The INRM can select the appropriate cloud according to the service QoS requirements of each slice and the radio and optical resource-usage condition provided by the ESC and CSC. To measure the choice rationality of service provisioning, we define F as the integrated evaluation factor that considers all multiple stratum parameters. As described above, U_c and U_m describe the current usage of cloud resources, while optical network parameters are composed of the traffic engineering weight W_e of the link. The overall RRH-cloud function is expressed in Eq. (1) and the BBU-cloud function in Eq. (2), where ψ and φ are adjustable parameters that weight the CPU usage and the storage utilization. CPU and storage utilizations are not directly comparable. Thus, ψ and φ provide a means of combining the two into one function:

$$f_{ac}(U_c^{\mathrm{BBU}}, U_m^{\mathrm{BBU}}, \psi) = U_c^{\mathrm{BBU}} \times \psi + U_m^{\mathrm{BBU}} \times (1 - \psi), \tag{1}$$

$$f_{bc}(U_c^{\mathrm{RRH}}, U_m^{\mathrm{RRH}}, \varphi) = U_c^{\mathrm{RRH}} \times \varphi + U_m^{\mathrm{RRH}} \times (1 - \varphi). \tag{2}$$

In addition, the optical network function is expressed by Eq. (3):

$$f_{cc}(W_e) = W_e. \tag{3}$$

$f_{a1}, f_{a2}, \ldots, f_{ak}$ and $f_{b1}, f_{b2}, \ldots, f_{bk}$ are the cloud parameters among the K candidate cloud nodes, while $f_{c1}, f_{c2}, \ldots, f_{ck}$ are the optical parameters associated with the K candidate paths. The integrated evaluation factor F is represented by

$$F = \frac{f_{ac}(U_c^{\mathrm{BBU}}, U_m^{\mathrm{BBU}}, \psi)}{\max\{f_{a1}, f_{a2}, \ldots, f_{ak}\}}\alpha + \frac{f_{bc}(U_c^{\mathrm{RRH}}, U_m^{\mathrm{RRH}}, \varphi)}{\max\{f_{b1}, f_{b2}, \ldots, f_{bk}\}}\beta + \frac{f_{cc}(w_e)}{\max\{f_{c1}, f_{c2}, \ldots, f_{ck}\}}\gamma \quad (\alpha + \beta + \gamma = 1), \tag{4}$$

where α and β are the adjustable weights of the cloud terms and γ is the weight of the optical term.

Note that the various weights can be selected based on the properties of the network. If the number of cloud nodes is relatively large compared to other resources (i.e., optical nodes and bandwidth capacity), the value of α or β should be low, because the cloud resources are a subordinate consideration. Conversely, if the number of optical nodes is greater than the number of cloud nodes, the value of γ should be low. Since the number of optical and cloud nodes is fixed, the weights can be set up in advance to control the importance of these parameters.

The flow chart of the proposed INRM scheme is illustrated in Fig. 3. When a UE sends a resource request, the INRM scheme chooses the cloud offloading strategy (cloud resource allocation) for the request according to the slicing identification and QoS requirements. To ensure that the critical latency requirements of low-latency services are met, the INRM scheme chooses RRH-cloud offload preferentially for uRLLC services. If the RRH-cloud resources are insufficient, the resource requests of uRLLC will be forwarded to a BBU-cloud with enough capacity. For eMBB and mMTC services, the INRM scheme chooses the best of K candidate BBU-cloud nodes according to their resource utilization. With regard to MFH bandwidth allocation, the node with minimum F based on the integrated evaluation factor will be selected from the K candidates. After the BBU-cloud is selected, the path is established with spectrum and modulated frequency allocation using the OpenFlow protocol between the source and destination nodes. As described, the resource allocation for the three types of network slices is based on hybrid bandwidth allocation to guarantee different requirements. A preemptive bandwidth allocation strategy was adopted for uRRLC because of its strict delay requirement and relatively small amount of data.

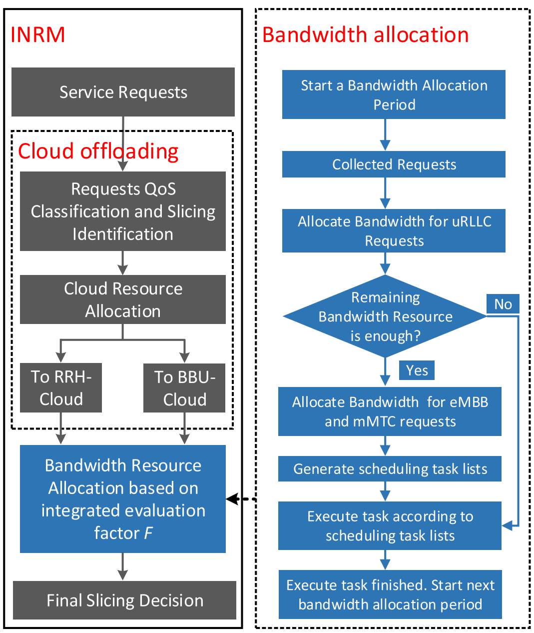

Fig. 3. Flow chart of the proposed INRM scheme.

C. Delay Formulation

The end-to-end delay of each packet is calculated in the proposed scheme to compute the optimal network resource management strategy. The E2E delay in a fronthaul network is denoted as D_{e2e}^F and formulated as

$$D_{e2e}^F = D_{\mathrm{proc}}^F + D_{\mathrm{prop}}^F + D_{\mathrm{link}}^F + D_{\mathrm{queue}}^F. \tag{5}$$

D_{proc}^F is the total processing delay, which is treated as a fixed value required to forward a packet as determined by the processing speed of the cloud. D_{prop}^F is the propagation delay, which is a symmetric and static value determined by the fiber length. Serialization delay D_{link}^F is proportional to the frame size and is inversely proportional to the link bandwidth capacity. Queuing delay D_{queue}^F is caused by the competition among fronthaul packets; i.e., a cloud receives multiple packets at the same time and they are forwarded in sequence. Queuing delay increases in proportion to the number of competing packets at each link. The maximum queuing delay of a link is formulated as

$$D_{\mathrm{queue}}^F = \frac{p\lambda}{b}, \tag{6}$$

where p denotes the packet size, b is the link speed, and λ denotes the maximum number of competing packets at that link.

V. Numerical Evaluation and Results Analysis

To evaluate the feasibility and efficiency of the proposed INRM scheme on the HEC-based optical MFH architecture, we developed a simulation platform using *Opnet*. The evaluations were executed on a laptop with an Intel Core i5-4210 processor with 8 GB of memory.

Note that the simulations are performed at a high level to demonstrate proof of concept. All implementation details that would be necessary in a real system were not included.

The important parameters employed in the simulations are summarized in Table I. The RRHs connect with RRH-clouds via a dedicated link, which means that the RRH-clouds are located at RRH sites. In addition, the RRHs are connected with the BBU-cloud via a fiber link, where the uplink bandwidth of each link is 10 Gbps. The total number of fronthaul network wavelengths is 10; thus, the total fronthaul uplink bandwidth of the MFH network is 100 Gbps. The number of UEs is 100. eMBB, mMTC, and uRLLC occupied, respectively, 20%, 60%, and 20% of the total number of UEs. We set the number of services as 30 per UE, which is usually much less than the number of users.

The service generation interval follows a Pareto distribution with mean arrival rate δ; the total peak traffic load is 100 Gbps. We simulate the total traffic load from 10 to 100 Gbps at 10 Gbps intervals by adjusting the service generation interval; that is to say, the normalized traffic load is 0.1 to 1. As for various services in different network slices, the bandwidth demand of eMBB services is uniformly distributed between 30M and 70M bytes, capturing bursty and high-bandwidth and cloud-computing demands. For this type of slice, the resource requests may not be completely satisfied due to network resource constraints. For the mMTC, the bandwidth demand is uniformly distributed between 0.5M and 1.5M bytes to model periodic, low-rate traffic. The bandwidth demand of uRRLC services is uniformly distributed between 5M and 15M bytes; these services have the most stringent latency requirements. The slicing resource manager of both eMBB and uRLLC slices uses a proportional fair algorithm, while mMTC slices use a pre-fixed scheduling scheme (semi-persistent scheduling).

We also ran a benchmark test case, where no INRM scheme is utilized. That is to say, there is no network slicing and all of the services have the same processing; i.e., we do not consider the different QoS requirements of different slices, all of the traffic has the same processing, and all of the traffic is processed in the BBU-cloud; i.e., the RRH-clouds are not used. It is assumed that the computing resources of the RRH-clouds are transferred to the BBU-cloud.

One-hundred runs of each simulation were performed. Unless otherwise stated, the numbers presented in the graphs represent the averages over the 100 runs.

We focus on the analysis of the flexibility and dynamicity of the proposed architecture and INRM scheme. The average delay experienced by each slice type is depicted in Fig. 4.

TABLE I
SIMULATION PARAMETERS

Parameters	Values
Number of BBU-clouds	1
Number of RRH-clouds	10
Number of RRHs	10
Number of UEs	100
Number of services per UE	30
Bandwidth of optical fronthaul	$10 * 10$ Gbps

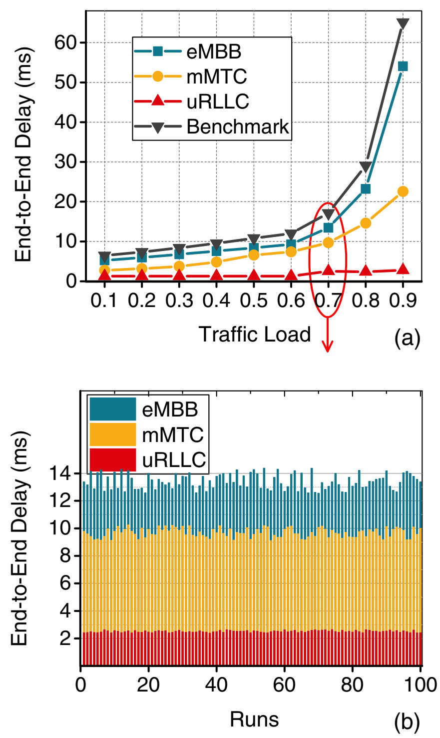

Fig. 4. Delay performance: (a) delay per slice versus traffic load; (b) delay under conditions of 70% traffic load.

Figure 4(a) illustrates end-to-end delay versus traffic load. Compared with the benchmark, the delays of each slice type are decreased significantly. The uRRLC slice has the lowest end-to-end delay of the three slice types. Even when the traffic load is 70% and network congestion occurs, the end-to-end delay of uRRLC is still lower than that of eMBB and mMTC slices. The main reason is that the traffic of uRRLC is usually processed in RRH-clouds, which avoids transporting the traffic on the heavily loaded optical fronthaul link, resulting in lower queuing delay and propagation delay.

Figure 4(b) represents the end-to-end delay statistics for each of the 100 simulation runs when the traffic load is 0.7. Network performance starts to deteriorate at this load, and the end-to-end delays are distinct for the different network slices. For uRLLC slices, the end-to-end delay is about 2 ms, whereas it is 9.5 ms and 13 ms for mMTC and eMBB slices, respectively. In brief, the INRM scheme is able to meet the diverse delay requirements for different slices.

Figure 5 shows the jitter versus traffic load. It has similar trends to the latency performance. The jitter is lower than 0.1 ms for each slice when the traffic load is below 0.7. However, the jitter increases sharply when the traffic load exceeds 0.7. The mean jitter is about 1.25 ms for eMBB and 0.77 ms for mMTC when the traffic load exceeds 0.7; the jitter for uRLLC remains low and steady at less than 0.1 ms. Hence the uRLLC traffic can achieve low jitter using the network-slicing-based INRM scheme. In the benchmark case, the jitter is larger than that of all the slice types.

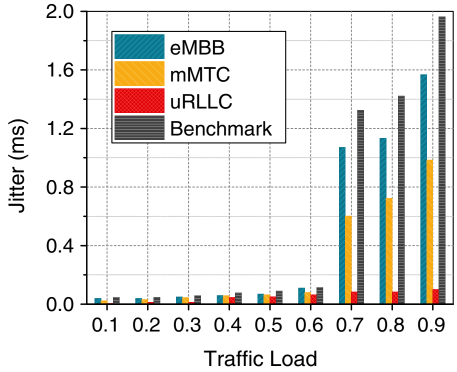

Fig. 5. Jitter performance for each slice type.

The throughput obtained by each slice type and the benchmark case is shown in Fig. 6. Figure 6(a) compares the performance of the proposed scheme and the benchmark in terms of the total throughput. As shown in Fig. 6(a), the INRM scheme can enhance the throughput more effectively than the benchmark. The reason is that the INRM scheme is able to globally optimize the network resources. In addition, we also calculate the throughput gain of the proposed INRM scheme as compared to the benchmark case. As shown in Fig. 6(a), the throughput gain increases at an early stage and then decreases with the addition of traffic load; when the network is heavily loaded, the throughput performance gain is relatively low. As expected, when the traffic load is low, the network resources are sufficient and the throughput performance gain is higher; as the traffic load increases, the network resources are constrained and the throughput performance gain declines accordingly.

In addition, the throughput of each slice type versus traffic load is illustrated in Fig. 6(b). Figure 6(c) represents the throughput statistics of 100 simulation runs where the traffic load is 0.7. These figures clearly illustrate the adaptability of the proposed INRM scheme to share the network resources. We observe that the eMBB slices always obtain the highest throughput to satisfy its huge bandwidth requirement, whereas the throughput of mMTC slices is the lowest.

Figure 7 compares the performance of the proposed scheme and benchmark in terms of resource utilization and occupied MFH bandwidth. Resource utilization reflects the percentage of the entire bandwidth and cloud-computing resources that are utilized. As shown in Fig. 7(a), the INRM scheme can improve resource utilization more effectively than the benchmark, especially when the network is heavily loaded. The reason is that the INRM

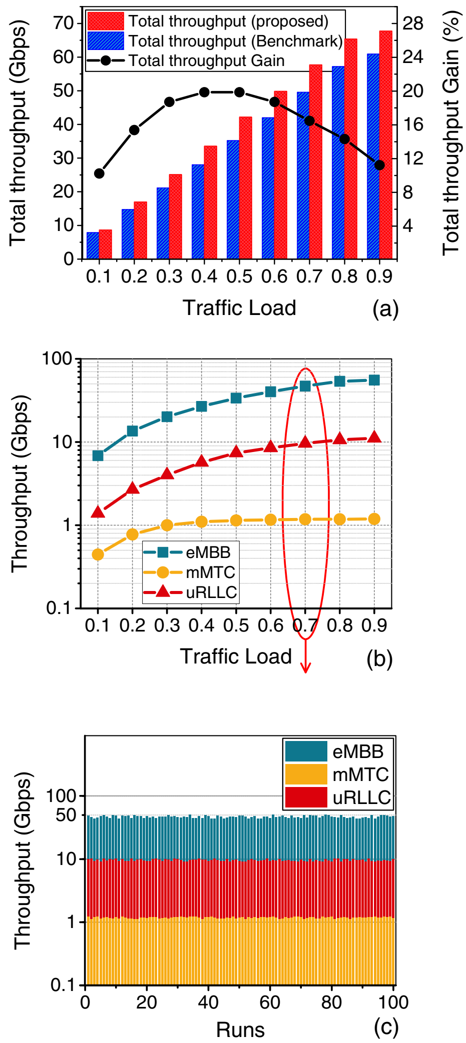

Fig. 6. Throughput performance: (a) comparison of throughput and throughput gain, (b) throughput per slice versus traffic load, and (c) throughput under conditions of 70% traffic load.

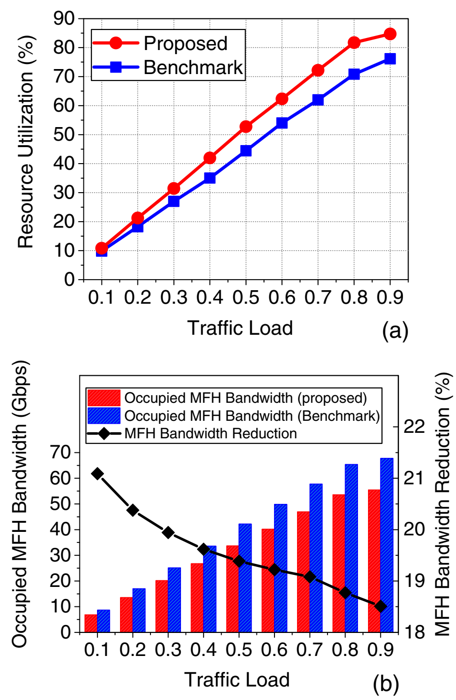

Fig. 7. Resource utilization performance: (a) resource utilization versus traffic load; (b) occupied MFH bandwidth and MFH bandwidth reduction ratio.

scheme can globally optimize the bandwidth and cloud-computing resources to maximize the service processing and transmission. Figure 7(b) shows that the INRM scheme can reduce the occupied MFH bandwidth and decrease the MFH bandwidth burden compared to the benchmark. What is noteworthy is that the MFH bandwidth reduction ratio decreases with increasing traffic load. The reason is that the proportion of offloaded services decreases as the capacity-limited RRH-clouds tend to saturate.

VI. Open Issues and Future Research Work

Although network slicing in HEC-based MFH has been analyzed here, network slicing remains at an early stage in terms of its development. Therefore, further enhancements and studies are needed for it to become a mature technology and thus be adopted over various domains of emerging 5G networks. Aside from the key techniques addressed here, business profit and field trials are two other important issues.

A. Business and Profit

The integration of various slices, each providing services at different stages and targeting specific end users, and partnership between several operators through infrastructure sharing lead to new challenges for total network investment, the service level agreement between owner and tenant, and expected generated revenue. Considering such an integrated business-oriented approach, different economic strategies and profit modeling should be extensively analyzed and developed in order to meet 5G network requirements. To reach this goal, a comprehensive study of the existing telecom regulatory framework should be conducted. New innovative ways of pricing, cost of infrastructure sharing, the service level agreement between the owner and the tenant of the slice, and expected generated revenue should be addressed and furthermore standardized.

B. Testbed and Field Trials

The industry has made a lot of progress in the realization and testing of CN-based network slicing. During the Mobile World Congress 2016, Deutsche Telekom and Huawei demonstrated the world's first end-to-end 5G network to validate CN-based network slicing for diverse 5G use cases. Autonomous end-to-end network slicing, including the dynamic and real-time slicing of the 5G RAN, CN, and the interconnecting transmission network, was demonstrated. These demonstrations showed that CN-based network slices can be automatically created in an optimized approach with a cost of sub-minute time. Motivated by the CN-based network-slicing testbed, it can be anticipated that MFH network slicing towards future 5G scenarios can be tested in future 5G field trials, and the gains will be significant.

To show the performance gains achieved by the proposed MFH network slicing, the first step is to build a powerful 5G HEC-based MFH testbed. Then the proposed architecture and key techniques for MFH network slicing should be fulfilled in the implemented testbed. Notable achievements are foreseen to be gained.

VII. Conclusion

In this paper, we have proposed an INRM scheme featuring HEC-based MFH abstraction. The proposed architecture relies on HEC and functional split to direct the traffic toward the appropriate network location, and uses an INRM scheme to allocate and share the network resources among slices. Moreover, the proposed architecture enables flexibility and dynamicity, using the HEC concept, to enforce network slicing in the fronthaul network, and adapt the resource allocation strategy according to the slice requirements. Simulation results indicate the feasibility and efficiency of our proposed INRM scheme. The obtained simulation results show that offloading traffic will help support the heterogeneous needs of various slices, while ensuring efficient and fair resource sharing among slices. These analyses and results provide a foundation for further studies of network slicing. Future work will focus on evaluating the model in a large-scale network scenario, and developing resource scheduling in more detail that will be configured to the network resource management platform.

Acknowledgment

This work is supported by the National Key R&D Program of China (2016YFB0901200) and the Open Research Fund of the Key Laboratory of Space Utilization, Chinese Academy of Sciences (No. LSU-DZXX-2017-01).

References

[1] 3GPP, "Study on architecture for next generation," Release 14, TR 23.799, Nov. 2016 [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/23_series/23.799/.

[2] 3GPP, "Feasibility study on new services and markets technology enablers stage 1," Release 14, TR 22.891, Sept. 2016 [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/22_series/22.891/.

[3] "Description of network slicing concept," NGMN Alliance 5G P1 [Online]. Available: https://www.ngmn.org/fileadmin/user_upload/161010_NGMN_Network_Slicing_framework_v1.0.8.pdf.

[4] S. Vassilaras, L. Gkatzikis, N. Liakopoulos, I. N. Stiakogiannakis, M. Qi, L. Shi, L. Liu, M. Debbah, and G. S. Paschos, "The algorithmic aspects of network slicing," *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 112–119, Aug. 2017.

[5] C.-L. I, Y. Yuan, J. Huang, S. Ma, C. Cui, and R. Duan, "Rethink fronthaul for soft RAN," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 82–88, Sept. 2015.

[6] B. Bangerter, S. Talwar, R. Arefi, and K. Stewart, "Networks and devices for the 5G era," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 90–96, Feb. 2014.

[7] M. R. Raza, M. Fiorani, A. Rostami, P. Öhlen, L. Wosinska, and P. Monti, "Dynamic slicing approach for multi-tenant

5G transport networks," *J. Opt. Commun. Netw.*, vol. 10, no. 1, pp. A77–A90, Oct. 2018.

[8] K. Miyamoto, S. Kuwano, T. Shimizu, J. Terada, and A. Otaka, "Performance evaluation of Ethernet-based mobile fronthaul and wireless CoMP in split-PHY processing," *J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. A46–A54, Jan. 2017.

[9] O. Chabbouh, S. B. Rejeb, N. Agoulmine, and Z. Choukair, "Cloud RAN architecture model based upon flexible RAN functionalities split for 5G networks," in *31st Advanced Information Networking and Applications Workshops (WAINA)*, Taipei, Taiwan, 2017, pp. 184–188.

[10] Open Networking Foundation, "Applying SDN architecture to 5G slicing," issue 1, TR-526, July 2016 [Online]. Available: https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/Applying_SDN_Architecture_to_5G_Slicing_TR-526.pdf.

[11] X. Zhou, R. Li, T. Chen, and H. Zhang, "Network slicing as a service: Enabling enterprises' own software-defined cellular networks," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 146–153, July 2016.

[12] U. Dötsch, M. Doll, H. P. Mayer, F. Schaich, J. Segel, and P. Sehier, "Quantitative analysis of split base station processing and determination of advantageous architectures for LTE," *Bell Labs Tech. J.*, vol. 18, no. 1, pp. 105–128, Jan. 2013.

[13] P. Rost, A. Banchs, I. Berberana, M. Breitbach, M. Doll, H. Droste, C. Mannweiler, M. A. Puente, K. Samdanis, and B. Sayadi, "Mobile network architecture evolution toward 5G," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 84–91, May 2016.

[14] P. Rost, C. Mannweiler, D. S. Michalopoulos, C. Sartori, V. Sciancalepore, N. O. Sastry, O. Holland, S. Tayade, B. Han, D. Bega, and D. Aziz, "Network slicing to enable scalability and flexibility in 5G mobile networks," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 72–79, May 2017.

[15] K. Katsalis, N. Nikaein, E. Schiller, A. Ksentini, and T. Braun, "Network slices toward 5G communications: Slicing the LTE network," *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 146–154, Aug. 2017.

[16] X. Foukas, M. K. Marina, and K. Kontovasilis, "Orion: RAN slicing for a flexible and cost-effective multi-service mobile network architecture," in *23rd Annu. Int. Conf. on Mobile Computing and Networking (MobiCom)*, Salt Lake City, Utah, 2017, pp. 127–140.

[17] R. Ferrus, O. Sallent, J. Perez-Romero, and R. Agusti, "On 5G radio access network slicing: Radio interface protocol features and configuration," *IEEE Commun. Mag.*, vol. 56, no. 5, pp. 184–192, May 2018.

[18] A. Mayoral, R. Muñoz, R. Vilalta, R. Casellas, R. Martínez, and V. López, "Need for a transport API in 5G for global orchestration of cloud and networks through a virtualized infrastructure manager and planner [Invited]," *J. Opt. Commun. Netw.*, vol. 9, pp. A55–A62, Sept. 2017.

[19] K. Samdanis, X. C. Perez, and V. Sciancalepore, "From network sharing to multi-tenancy: The 5G network slice broker," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 32–39, July 2016.

[20] T. Taleb, B. Mada, M. Corici, A. Nakao, and H. Flinck, "PERMIT: Network slicing for personalized 5G mobile telecommunications," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 88–93, May 2017.

[21] W. Guan, X. Wen, L. Wang, Z. Lu, and Y. Shen, "A service-oriented deployment policy of end-to-end network slicing based on complex network theory," *IEEE Access*, vol. 6, pp. 19691–19701, June 2018.

[22] M. R. Sama, X. An, Q. Wei, and S. Beker, "Reshaping the mobile core network via function decomposition and network slicing for the 5G era," in *Wireless Communications and Networking Conf. Workshops (WCNC)*, Doha, Qatar, 2016, pp. 1–7.

[23] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network slicing in 5G: Survey and challenges," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 94–100, May 2017.

[24] R. Martínez, R. Vilalta, R. Casellas, R. Muñoz, L. Fei, P. Tang, and V. López, "Network slicing resource allocation and monitoring over multiple clouds and networks," in *Optical Fiber Communications Conf. and Expo. (OFC)*, San Diego, California, Mar. 2018. pp. 1–3.

[25] X. Costa-Pérez, J. Swetina, T. Guo, R. Mahindra, and S. Rangarajan, "Radio access network virtualization for future mobile carrier networks," *IEEE Commun. Mag.*, vol. 51, no. 7, pp. 27–35, July 2013.

[26] P. Caballero, A. Banchs, G. de Veciana, and X. Costa-Pérez, "Network slicing games: Enabling customization in multi-tenant networks," in *Conf. on Computer Communications (INFOCOM)*, Atlanta, Georgia, 2017, pp. 1–9.

[27] O. Sallent, J. Perez-Romero, R. Ferrus, and R. Agusti, "On radio access network slicing from a radio resource management perspective," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 166–174, May 2017.

[28] Y. L. Lee, J. Loo, and T. C. Chuah, "A new network slicing framework for multi-tenant heterogeneous cloud radio access networks," in *Int. Conf. on Advances in Electrical, Electronic and Systems Engineering (ICAEES)*, Putrajaya, Malaysia, 2016, pp. 414–420.

[29] V. G. Nguyen, T. X. Do, and Y. Kim, "SDN and virtualization based LTE mobile network architectures: A comprehensive survey," *Wireless Pers. Commun.*, vol. 86, no. 3, pp. 1401–1438, Mar. 2016.

[30] W. Wang, Y. Chen, Q. Zhang, and T. Jiang, "A software-defined wireless network enabled spectrum management architecture," *IEEE Commun. Mag.*, vol. 54, no. 1, pp. 33–39, Jan. 2016.

[31] C.-L. I, H. Li, J. Korhonen, J. Huang, and L. Han, "RAN revolution with NGFI (xHaul) for 5G," *J. Lightwave Technol.*, vol. 36, no. 2, pp. 541–550, Jan. 2018.

[32] 3GPP, "Technical specification group services and system aspects: Enhancements of dedicated core networks selection mechanism," Release 14, TR 23.711, Sept. 2016 [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/23_series/23.711/.

[33] Y. H. Kim, "Slicing the next mobile packet core network," in *11th Int. Symp. On Wireless Communications Systems (ISWCS)*, Barcelona, Spain, 2014, pp. 901–904.

[34] T. Taleb, M. Corici, C. Parada, A. Jamakovic, S. Ruffino, G. Karagiannis, and T. Magedanz, "EASE: EPC as a service to ease mobile core network deployment over cloud," *IEEE Netw.*, vol. 29, no. 2, pp. 78–88, 2015.

[35] A. Ksentini and N. Nikaein, "Toward enforcing network slicing on RAN: Flexibility and resources abstraction," *IEEE Commun. Mag.*, vol. 55, no. 6, pp, 102–108, June 2017.

[36] A. Devlic, A. Hamidian, D. Liang, M. Eriksson, A. Consoli, and J. Lundstedt, "NESMO: Network slicing management and orchestration framework," in *Int. Conf. on Communications Workshops (ICC Workshops)*, Paris, France, 2017, pp. 1202–1208.

[37] J. Ordonez-Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca, and J. Folgueira, "Network slicing for 5G with SDN/NFV: Concepts, architectures, and challenges," *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 80–87, May 2017.

[38] P. Rost, C. J. Bernardos, A. De Domenico, M. Di Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wübben, "Cloud technologies for flexible 5G radio access networks," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 68–76, May 2014.

[39] D. Liang, R. Gu, Q. Guo, and Y. Ji, "Demonstration of multi-vendor multi-standard PON networks for network slicing in 5G-oriented mobile network," in *Asia Communications and Photonics Conf. (ACP)*, Guangzhou, China, Nov. 2017, paper M3B55.

[40] R. Gu, S. Zhang, Y. Ji, and Z. Yan, "Network slicing and efficient ONU migration for reliable communications in converged vehicular and fixed access network," *Veh. Commun.*, vol. 11, pp. 57–67, Sept. 2018.

[41] R. Martínez, R. Vilalta, M. Requena, R. Casellas, R. Muñoz, and J. Mangues, "Experimental SDN control solutions for automatic operations and management of 5G services in a fixed mobile converged packet-optical network," in *Int. Conf. on Optical Network Design and Modeling (ONDM)*, Dublin, Ireland, May 2018, pp. 214–219.

[42] 3GPP, "Study on new radio access technology: Radio access architecture and interfaces," Release 14, TR 38.801, Apr. 2017 [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/38_series/38.801/.

[43] P. Chanclou, L. A. Neto, K. Grzybowski, Z. Tayq, F. Saliou, and N. Genay, "Mobile fronthaul architecture and technologies: A RAN equipment assessment," *J. Opt. Commun. Netw.*, vol. 10, no. 1, pp. A1–A7, Oct. 2018.

[44] S. Kumar, S. Dharmapurikar, F. Yu, P. Crowley, and J. Turner, "Algorithms to accelerate multiple regular expressions matching for deep packet inspection," in *Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM)*, New York, Sept. 2006, pp. 339–350.

[45] D. Wang, M. Zhang, J. Li, and X. Chen, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Opt. Express*, vol. 25, no. 15, pp. 17150–17166, July 2017.

[46] S. Han, C.-L. I, G. Li, S. Wang, and Q. Sun, "Big data enabled mobile network design for 5G and beyond," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 150–157, Sept. 2017.

[47] R. Challa, S. Jeon, S. M. Raza, P. Thorat, and H. Choo, "SuperFlex: Network slicing based super flexible 5G architecture," in *12th Int. Conf. on Ubiquitous Information Management and Communication (IMCOM)*, Langkawi, Malaysia, Jan. 2018, p. 98.

Chuang Song is currently working toward the Ph.D. degree in electronic science and technology at the Key Lab of Information Photonics & Optical Communications (IPOC), BUPT. His main research interests include 5G emerging technologies, software-defined optical networking, cloud-computing network and network resource management, and orchestration with machine learning.

Min Zhang received his Ph.D. degree in optical communications from Beijing University of Posts and Telecommunications (BUPT). He is now a professor at BUPT. His main research interests include advanced optical communication systems and networks, especially high-speed switching, optical signal processing, and optical wireless communications. He holds 32 China patents. He has authored or co-authored more than 300 technical papers in international journals and conferences, and 12 books in the areas of optical communications.

Danshi Wang received the Ph.D. degree in electromagnetic field and microwave technology from the Beijing University of Posts and Telecommunications (BUPT) in 2016. He is currently a lecturer of the Institute of Information Photonics and Optical Communications at BUPT. He has published more than 50 articles. His research interests include wavelength switched optical networks, optical transmission and optical signal processing, artificial intelligence and deep learning, etc.

Yueying Zhan was born in Shandong province, China, in 1987. She received her Ph.D. degree from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China. She is currently working at the Chinese Academy of Science. Her main research interests include high-speed space optical communication and high-speed optical signal processing technologies based on semiconductor optoelectronic devices.

Luyao Guan is currently working towards a Ph.D. degree at the Beijing University of Posts and Telecommunications (BUPT), Beijing, China. Her research includes work on architecture and resource allocation algorithm design based fronthaul networks in 5G, software-defined networks (SDNs), and network virtualization.

Wei Liu is currently pursuing a master's degree at the Beijing University of Posts and Telecommunications (BUPT), Beijing, China. His research focuses on cloud networking, network slicing, and edge cloud computing.

Lin Zhang is currently pursuing a master's degree at the Beijing University of Posts and Telecommunications (BUPT), Beijing, China. His research focuses on 5G network slicing, software-defined networks (SDNs), and network virtualization.

Siya Xu was born in Beijing, China, in 1988. She received the B.S. degree in communication engineering from the University of Science and Technology Beijing and the M.S. and Ph.D. degrees in communication and information systems from the Beijing University of Posts and Telecommunications. She is currently a post-doctor in the State Key Laboratory of Networking and Switching Technology of the Beijing University of Posts and Telecommunications. Her research interests include smart grid communication network management and optimization.