A Survey of Mobility Management as a Service in Real-Time Inter/Intra Slice Control

MOHAMMEDALI, N. A. In-network softwarization, Network Slicing provides scalability and flexibility through various services such as Quality of Service (QoS) and Quality of Experience (QoE) to cover the network demands. For the QoS, a set of policies must be considered in real-time, accompanied by a group of functions and services to guarantee the end-user needs based on network demand. On the other hand, for the QoE, the service's performance needs to be improved to bring an efficient service to cover the demands of the end-user. The 3G Partnership Project (3GPP) defined the slice as a component of resources used to process a set of packets. These resources need to be flexible, which means the resources can be scaled up or down based on the demand. This survey discusses softwarization and virtualization techniques, considering how to implement the slices for future networks. Specifically, we discuss current advances concerning the functionality and architecture of the 5G network. Therefore, the paper critically evaluates recent research and systems related to mobility management as a service in real-time inter/intra slice control by considering the strengths and limitations of these contributions to identify the research gaps and possible research directions for emerging research and development opportunities. Moreover, we extend our review by considering the slice types and their numbers based on the 3GPP Technical Specification (3GPP TS). The study presented in this paper identifies open issues and research directions that reveal that mobility management at a service level with inter/intra slice management techniques has strong potential in future networks and requires further investigation from the research community to exploit its benefits fully., KANAKIS, T, OPOKU AGYEMAN, M., AND AL-SHERBAZ, A.

Abstract: In-network softwarization, Network Slicing provides scalability and flexibility through various services such as Quality of Service (QoS) and Quality of Experience (QoE) to cover the network demands. For the QoS, a set of policies must be considered in real-time, accompanied by a group of functions and services to guarantee the end-user needs based on network demand. On the other hand, for the QoE, the service's performance needs to be improved to bring an efficient service to cover the demands of the end-user. The 3G Partnership Project (3GPP) defined the slice as a component of resources used to process a set of packets. These resources need to be flexible, which means the resources can be scaled up or down based on the demand. This survey discusses softwarization and virtualization techniques, considering how to implement the slices for future networks. Specifically, we discuss current advances concerning the functionality and architecture of the 5G network. Therefore, the paper critically evaluates recent research and systems related to mobility management as a service in real-time inter/intra slice control by considering the strengths and limitations of these contributions to identify the research gaps and possible research directions for emerging research and development opportunities. Moreover, we extend our review by considering the slice types and their numbers based on the 3GPP Technical Specification (3GPP TS). The study presented in this paper identifies open issues and research directions that reveal that mobility management at a service level with inter/intra slice management techniques has strong potential in future networks and requires further investigation from the research community to exploit its benefits fully.

Index Terms: 5G slicing, future network, inter-slice, intra-slice, M-MaaS, NFV, NS, SDN

1. Introduction

Software-Defined Network (SDN) is an intelligent network that contains two parts: 1) Controller based on software and 2) Packet forwarding based on hardware devices programmed by an open interface such as ForCES and OpenFlow [1]. VMware, owned by Martin Casado, bought Nicira for $1.26 million to strengthen its position in the SDN world by virtual programming networks with pools of resources in a cloud environment to support virtualization and softwarization concept to reduce the cost and enhance the virtual network elements [2]. In a traditional mobile network, Mobility Management Entity (MME), Serving Gateway (SGW), Packet Data Network Gateway (PGW), Home Subscriber Server (HSS) and Policy and Charging Rule Function (PCRF) designed as core network elements. These elements could be configured in a virtual platform using the virtualization advantage after the standardization of the Network Function Virtualization (NFV) in 2012 [3].

Network Slicing is a significant part of the future network. It operates based on a set of services that need algorithms to manage all the demanded services within the slice and manages the continuity of the services for the user within the slice according to the Technical Specifications (TS) and Technical Reports (TR) from the 3GPP and the European Telecommunications Standards Institute (ETSI). Therefore, it is not a function.

FIGURE 1. Slice Classification Options.

The SDN, NFV, and Network Slicing have been widely accepted by the research community and industry as fundamental components of the next-generation telecommunication network [4]. With the SDN approach, the control plane is separated from the data plane. SDN controller is the network's brain that dynamically controls the system in an efficient way [5]. On top of that, all control functions can be programmed in the control plane to optimize the data plane dynamically based on the services [6].

Slice is a set of network resources that run on top of physical devices using SDN/NFV [7]. These resources will be selected and isolated for each slice to satisfy the QoS requirements. Additionally, a sub-network could be generated by the slice, which contains Network Functions (NFs) running on top virtual resources [8]. There are two types of slice provisioning [9], one to create slice links, and another to generate slice nodes virtually on the networks. Furthermore, there are three types of services: enhanced Mobile Broadband (eMBB), ultra-Reliable Low Latency Communications (URLLC) and Massive IoT (MIoT). In these services, two types of priorities need to be considered: high and low priority for each application in these services [10].

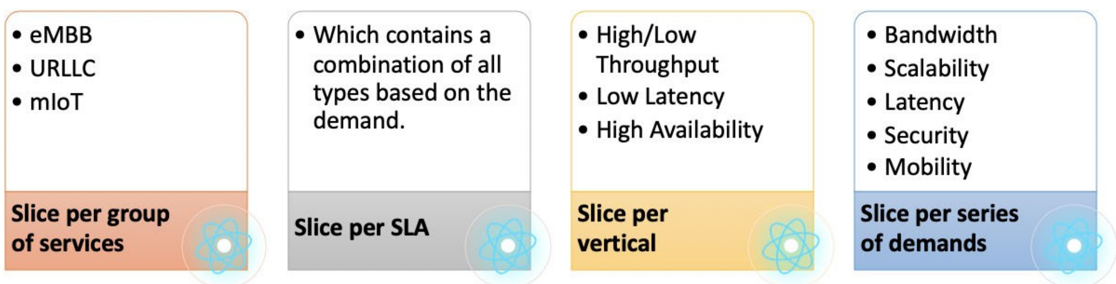
In the future networks, the QoS will be higher when the same amount of resources are required by different slices with different priority values [11]. Based on the priority values, a set of choices will present the details in a group of services within the slice, as shown in Figure 1.

In the 5G Public-Private Partnership (5GPPP) 2017 report [12], they captured almost all the future mobile network's open issues and how these issues could be solved by using softwarization and virtualization mechanisms. Moreover, in the same report, the 5GPPP explained the architecture and the technical directions of the future network and how it could fit all the use cases in future network [12]. The relationship between Network Slice Subnet (NSS) and the network slice service, function, plus the relationship among the NSS and nested network services have been defined in the ETSI report on Network Slicing [13]. In [14] the ETSI discusses the Network Slicing components with their functions in a core and access layer and maps to the 3GPP TR 28.801 to explain the slicing concepts with their usage of the communication services within the slice. Additionally, the concept of Network Slicing and virtualization in 5G system architecture has been explained in details by ETSI [15].

In the 5G networks, a set of functions need to be developed to run in real-time in the core layer to configure the access layer instantaneously according to the networks demand. Afterword, these functions will be managed by service-driven and deployment-driven models. Furthermore, these mechanisms could be programmed easily to satisfy the requirement of network operators. A further important implication is, when the service configuration changes, two parameters need to be considered: bandwidth and delay for each service within the slice. Therefore, we explain a set of functions such as AMF, SMF and PCF. These functions could be achieved with programmable techniques in SDN/5G.

Managing the mobility at a service level is critical in mobile networks for inter/intra slice [16], [17]. Specifically, mobility management affects the real-time QoS delivered to the users based on their demands [10], [18]–[20]. Research has shown that service continuity between slices is an open research area with several challenges and opportunities. It is still under investigation by the telecommunication groups, and researchers in a service level for the real-time application within the core network [19], [21]–[23]. Consequently, this paper presents an overview of current research findings focusing on Network Slicing and network services in the future mobile systems, as highlighted in the red circles in Figure 2.

The paper's structure summarized in Figure 3, and the rest of this survey is organized as follows: Section II, presents a review of the significant trends of softwarization and virtualization technologies such as SND, NFV and Network Slicing. In addition, the 5G functionally and architecture are discussed in terms of their implementation in future networks. Furthermore,



Section II provides a review of slice policy services and the represented slice's resources management as an essential role in the core layer. Also, mobility management is discussed at a service level to manage inter/intra slice services. Section III discusses slice number and their types based on 3GPP TS. Additionally, the Section discusses a typical system model's details according to the ETSI specification and existing related research and projects along with their limitations. Section IV considers the traffic classification for the future network depending on the use cases found in other research papers. Section V highlights the key challenges and open research issues related to internal and external slice management. Also, the Section identifies the critical research gaps and makes recommendations for research directions for emerging researchers and industrial development in future networks. Finally, Section VI concludes the paper.

2. Background about inter and intra slice management techniques

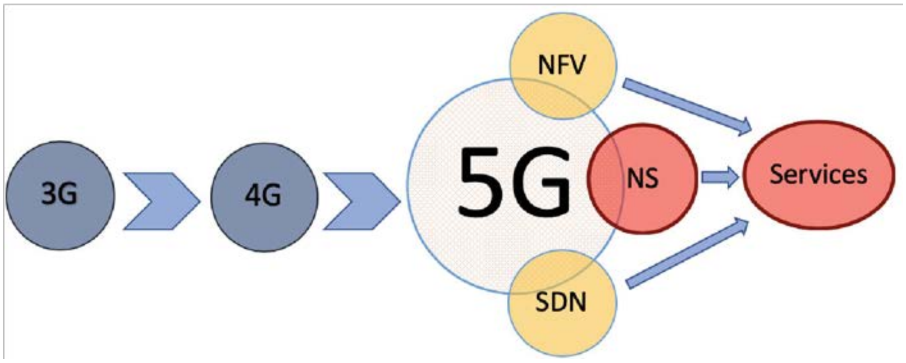
The programmability in future networks means all core network entities will be built in the application layer using programmable software to virtualize the entities. Then, Application Programming Interface (APIs) such as Rest API will connect to the control plane to handle the network recourses from the data centers based on the network demands. Moreover, the virtualization technique will be responsible for mapping the physical resources into the logical recourse to meet the requirements. In the control plane, NFs such as resources allocation and mobility will be managed virtually using SDN/NFV.

This section will explain all the technologies are used to implement the inter/intra slice policies in the future network.

A. VIRTUALIZATION TECHNIQUES

Functions (VNFs) implemented virtually, and the NFVI includes all the resources that the network needs, such as storage and capacity. In addition, these two elements are managed by MANO component.

FIGURE 2. Evolution of 5G Network Slice and Services. Highlighted in red is the Scope of this Paper.



The MANO has three functional components. NVF Orchestration (NVFO) manages the virtual resources, the life cycle for NVF in the slice and the authorization and validation of the NFVI request from the VNF manager. The VNF manager manages the life cycle management of the VNFs component that runs in Network Slicing. Virtual Infrastructure Manager is used to manage and control the resources linked with NFVI to manage the service and the traffic in the network [3].

The NFV provides high volume services that can be allocated in multiple nodes in virtual networks. The NFs implemented in virtual nodes to optimize the networks [28]. Moreover, the main idea of the softwarization concept is providing the flexibility demands at the core network, which means all network functions have to be ready on-demand [29]. Future network devices need to be in high-level configurability and programmability to efficiently support the data flows from different services to visualize this requirement. Scalability is needed to prevent interference between services. Service-oriented management is also required to control the flow of traffic over the network using NFV framework. Also, in the future network, many principles in high-

Network Function Virtualization (NFV) is used to reduce the network infrastructure's capital expenditure by implementing NFs on top of specific hardware platforms by using virtualization techniques. This novel approach makes the telecommunication systems update and migrates from the hardware to software using a softwarization technology [24].

Various factors led developers to use virtualization techniques such as NFV to significantly impact the 5G system when the NFV improves the NFs while the SDN enhances the base system. These advantages are outlined in [25], [26] and in Zhang's book [27], which are summarized in Figure 4.

The ETSI in their framework document [3] defined the architecture of the NFV, which includes a set of NFs, NFV Infrastructure (NFVI) and Management and Orchestration (MANO). In the NFV, all the Virtual Network level should be defined, such as end-user requirement and envisaged service [11].

B. SOFTWARIZATION TECHNIQUES BASED ON SDN

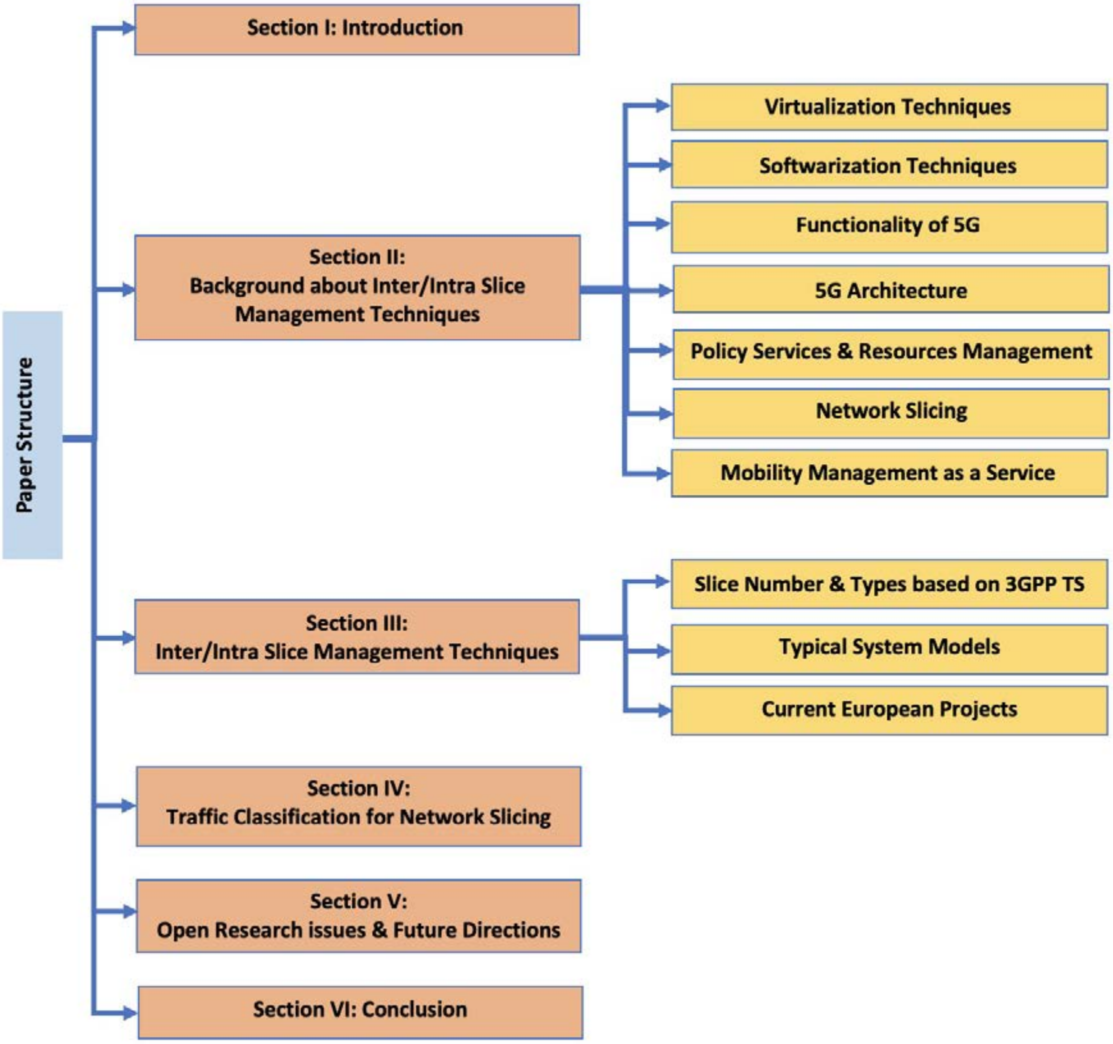
The traditional networks have physical elements such as MME, SGW and PGW. In the future network, SDN/NFV techniques are used to implement these elements in a core network as a pool of virtual entities containing all the **FIGURE**

3. Classification/Taxonomy of this Survey.

functions dedicated to the physical infrastructure [16]. In the core network, the SDN controller will reduce the delay and control data transmission. Moreover, SDN is logically centralized, and it is configured with an intelligent technique to automatically share, manage, and optimize the resources over the network.

Inside the control layer, two types of the controller will be used for the inter/intra slice. First, Software-Defined Mobile Network Coordinator (SDM-X) for inter-slice. In this controller, new services will be configured in the application layer and delivered through Northbound Interface (NBI) to take the actions through Southbound Interface (SBI) based on the demands. Second, Software-Defined Mobile Network Controller (SDM-C) for intra-slice. This controller is worked in SBI for transmission points and sharing resources between controller plane and user plane. On top of, physical resources will be shared the streams of the data entirely [18].

The SDN controllers run specific algorithms to manage and control the state of the network. In the SDN field, several controllers are developed such as Open Daylight [30], OpenFlow [31] [16], Flow Visor [32], NOX, Ryu, Open Virtex and ONOS [33]. Furthermore, with network virtualization, some of these controllers are used to create Network Slicing within the



SDN-based approach.

The OpenFlow [31] is a softwarization protocol in SBI which is used to control the data flow in data plane based on the information that in PDU header, which contains IP address, port number, MAC address and other parameters that added to the header to identify the traffic in the network. OpenFlow controller mange OpenFlow switches in the data plane from the control plane. Typically, the OpenFlow monitor the state of the network with a single controller, but with multiple or different controllers, ONOS will be used for this purpose [34].

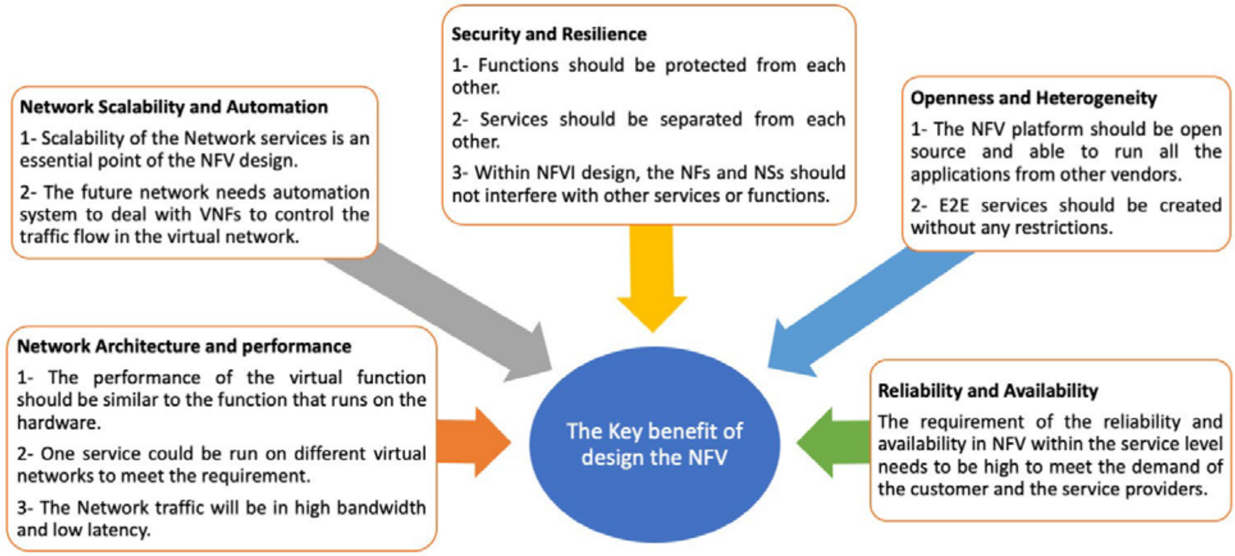


FIGURE 4. The Benefits of NFV for Future Networks [27].

There are three main services that the future networks will cover in a radio and core layer [35]: the eMBB is used with high throughput, which needed in a smart city, the URLLC characterized in terms of high reliability and low latency, which needed in industrial automation use cases, and the mMTC is used with a large density for a connection object in IoT devices.

Mahindra *et al.* [36] investigated an algorithm for slicing the resources dynamically in the radio access control.

Furthermore, the authors mentioned that their system (RadioVisor) built on top of the SoftRAN system [37] to isolate the resources virtually to implement the Radio Access Network (RAN) slices. Moreover, RadioVisor attempted to fix the resources sharing per slice by implementing a heuristic solution. However, the solution ended up with another problem: the slice owner could not decide to allocate resources to his customer within the slice [38].

Several studies, for instance, Peng *et al.* [39] and Zhang *et al.* [40] have discussed some of the challenges in the RAN, core network and some of the issues that related to Network Slicing that need to be addressed in the new mobile network to be able to connect 5G networks and other technologies to create one system fit all. One of these issues is network reconstruction, which used to rebuild the radio and core layers to support the slices for End-to-End (E2E) mobile communication.

C. ARCHITECTURE OF 5G TECHNOLOGY

FIGURE 5. Point-to-Point Architecture in 5G Core Network [42].

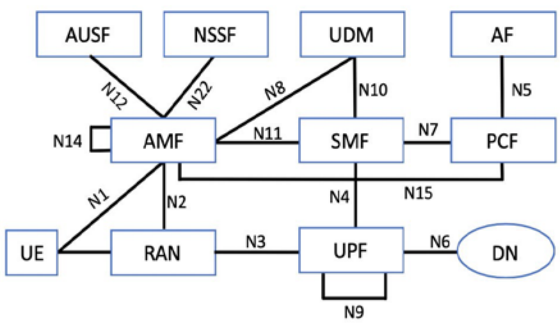


FIGURE 5. Point-to-Point Architecture in 5G Core Network [42].

5G core network contains two phases in their architecture based on 3GPP TS 23.501 release 15. In the First phase, point-to-point connected between the NFs [41]. Also, these links connected virtually and physically. Moreover, this point-to-point reached the access layer by connecting the core functions with the access layer functions, as shown in Figure 5. Furthermore, this phase will bring complexity to future networks because it requires an E2E configuration in multiple nodes [42].

In the second phase, service-based has the same connection links in the point-to-point phase, but just in the access layer (user plane) [41]. Additionally, there is a message bus that connected 5G core layer elements in the control plane with NFs, and the relationship between these functions be one-to-one or one-to-many based on the event provided by the APIs as shown in Figure 6 [42].

D. FUNCTIONALITY OF 5G

The architecture design for 5G research project in [43], it is defined the NFs and inter-function interfaces with several benefits that keep the networks away from the overhead optimize the NFs within specific implementations. The 5G architecture [44] contains several elements to control the NFs and build control layer used to manage the data plane upon the system demands. In the control plane, next-generation core system entities will be implemented to manage the system infrastructure to bring scalability and flexibility to the network.

1) Access and Mobility Management Function (AMF): This element is similar to the MME in today networks. It is used to handle the registration, connection and mobility management. As well, it is used to handle the access control function. Plus, this function will be responsible for providing priority to mobility management to prioritize the services.
2) Network Slice Selection Function (NSSF): NSSF is used to provide the functionality of Network Slicing to virtualize the network to slices. The selection function for each container will be done based on the slice's behaviour and service. For instance, a new system will have the first slice for the video stream, the second slice for emergency service and the last one for the IoT service.
3) Authentication Service Function (AUSF): This function is used to identify the user authentication and authorization, which it looks like the HSS work in the traditional network.
4) Session Management Function (SMF): This element is used to handle the traffic session across the network. It is similar to the S/PGW, but this time in the control layer.
5) Policy Control Function (PCF): This element is used to identify the policy rules for the network to handle and prioritize traffic types.
6) User Plane Function (UPF): It is used to handle the traffic in a user plane. This function is also similar to the S/PGW when it is mange the service type.
7) Unified Data Management (UDM): This function is used to integrate and manage the subscriber's data and provide security management to the network.
8) NF Repository Function (NPF): NPF allows other functions to connect and discover each other. Furthermore, this function used by Network function instance to store all functions used in the network.

E. NETWORK SLICING

A group of telecommunications companies has identified the concept of the Network Slicing [12], [45], [46] which have divided the slices into three layers: service layer, network slice layer and resource layer to isolate the logical components from the physical infrastructure. Moreover, telecommunications societies divided the networks into sub-networks inside the slice to provide specific slice priorities.

Several principles build the shape of Network Slicing, such as automation, isolation, customization, programmability and elasticity. In automation, the network will be re-configured without rewrite the SLA again. In addition,

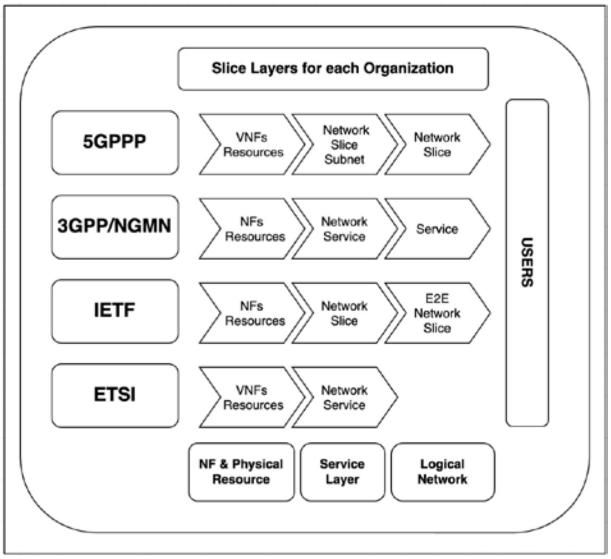


FIGURE 7. Slice Layers for each Organization [50].

isolation represents an essential feature in Network Slicing, which is used with different resources on a virtual network based on specific policy for each tenant [47]. On top of that, customizing the resources to a specific tenant meets the network level requirement. Therefore, data flow and service tailored network function can be customized on the data plane and introducing policies and protocols can be customized on the control plane after programming the controller and control the network via open APIs [48]. Moreover, slices will be scaled up/down after adding new functions to modify the VNFs [49].

In Figure 7, the first layer contains network function and infrastructure resources. The second layer contains all the services that the slice needs to provide its users, followed by the last layer, which includes the logical network characteristic that meets the customer's requirements [50].

According to the TS and TR from the 3GPP and ETSI, the Communication Services (CS) relationships with Network Slice Services Management Function (NSSMF) will be proposed in Figure 8. Besides, through Communication Service Management Function (CSMF), the system demands could be converted to Network Slice components which contain the Network Slice Instance (NSI). In addition, Network Slice Services Management Function (NSSMF) could communicate with Network Slice Management Function (NSMF), and managed by Network Slice Subnet Instance (NSSI) to provide batter services to the users based on their requirement. Moreover, the Network Slice could be used and managed by the Nested Network Services (NNS) and the VNFs with the Core Network Function (CNF) and Access Network Function (ANF), as shown in Figure In future, Network Slicing will be responsible for providing flexibility and efficiency to the mobile system. There are two ways of slicing the network: Vertical Slicing and Horizontal Slicing [51]. The Vertical Slicing has shared the resources between 8.

the services and applications. Also, it classified the traffic according to the application type to enhance the QoS in the network domain. On the other hand, Horizontal Slicing has shared the resources between network nodes and devices to increase its capability. In this type, new functions could be added dynamically when a specific slice needs a specific support [52].

Through Network Slicing, physical infrastructure will be sliced into several logical infrastructures that support different QoS demands at the same time. On top of that, slices will be isolated from the control plane to the data plane. Besides, the QoE for the user in a logical network will be the same as in the physical network.

The SDN has supported Network Slicing with different application scenarios, which can reconfigure the network automatically. Additionally, there are several types of controller used in Network Slicing. For example, FlowVisor, OpenVirtex and ONOS.

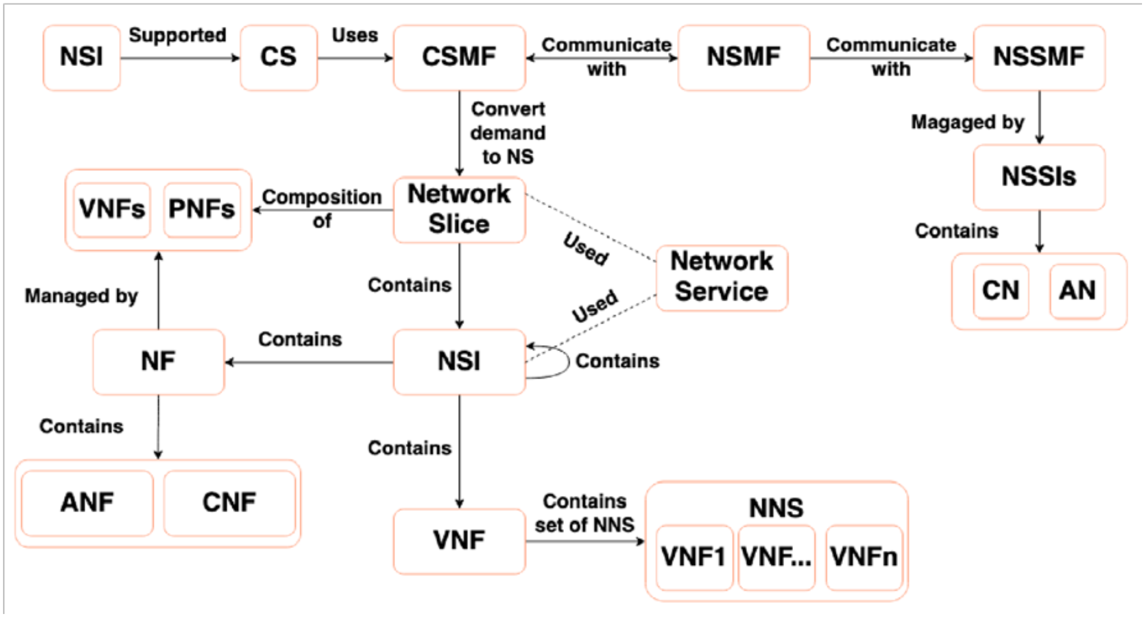


FIGURE 8. Relationship Between Network Slice and Other 5G Components.

The FlowVisor is represented as a proxy in Network Slicing techniques. Each slice in the network connected with its controllers using this representative. On the other hand, this controller has a limitation in packet header space because all slices in the network shared the flow and address space. Also, it does not support a virtual network on top of the physical network, and the development state of this controller stopped in 2013 [32]. The OpenVirtex is used to re-write the IP/MAC address to split up the traffic among multi-tenants. Inside this controller, any technology could be implemented in the virtual network. This controller's limitation, full rages of IP/MAC, is not available, and this controller's development state stopped in 2014. ONVisor is designed as a platform based on the ONOS controller. It is used to implement the slices in the virtual network because it is supported virtualization techniques. Additionally, In this platform, virtual networks were isolated using VLAN tags with a limited tag size [53]. None of the above controllers is designed for Network Slicing; even these controllers work in a virtual network. Consequently, Open Network Foundation proposed a model based on SDN platform with Network Slicing in 5G network to control and manage the slices and data flows in the network from the control plane to the data plane [5].

F. RESOURCES AND POLICY MANAGEMENT

In the slice, many services could be implemented at the network level, such as policy management, the QoS, setting the load balance, and isolating the network resources shared in inter-slice and intra-slice. In addition, the priority of the QoS flows will be provided by the SMF based on the 5G network architecture [42]. Further, Network Slicing could manage the creation, activation, and reconfiguration for the slice in a specific life cycle based on the system needs at the service level. Hu *et al.* [54] introduced Network Slicing and scheduling scheme on downlink transmission in the LTE. The scheduling scheme gave a specific number of sub-channels according to each virtual network's requirements to provide services to its users. Shurman *et al.* [55] enhanced the network's performance and decreased the delay by modifying TCP packet using a simulation tool. In this case, the highest priority packet will have the chance to be first in the queue and discard the other packet to reduce the congestion in the network. Moreover, their implementation is not based on SDN/NFV.

5G core network implementation was proposed to guarantee the QoS that the user will get based on the 3GPP architecture for 5G system. In Fattore *et al.* approach [56], the data plane will be programmed in SDN for future work to handle the user traffic based on a set of policies.

Slice selection mechanisms for the 5G era was explained in [57] which developed the core layer to define all the necessary functions to implement services in the core and run it over RAN to improve the network performance. Moreover, each slice could have suited functions in the core network that support specific applications to serve the slice's service.

The slice's weight prioritizes based on the service's cost was explained in [10], which means high premium users get their service using a short path, and other users get their service on a long path. Two types of services have been considered in [58] to improve the networks' flexibility. By focusing on the implementation of the RAN in the cloud by using the SCA problem but in [59] authors tackled the problem by using machine learning as in [60], and SNR in [59], but not with the 5G entities. While Markov Decision process was applied for resources allocation in RAN [61] to enhance the flexibility of the 5G systems.

Q-learning was applied in [60] to manage the slice's resources dynamically for each tenant to provide QoS to the end-user according to their demands. After the learning process, a set of policies added to the slices using a greedy algorithm, as shown in [22]. Besides, their learning approach was based on historical data, not online data [20]. The authors in [22] are used historical data for learning, which is offline learning, not real-time learning with online data. Moreover, in online learning, data collected in real-time and the learning decision occurred in real-time. In addition, in [20], different learning mechanisms were proposed based on supervised learning. Therefore, their learning is done offline; then, their prediction will address future actions or suggest actions for the current situation. Over and above, decision-tree provided different service to several slices, and their supervised learning was executed offline.

Thantharate *et al.* [62] trained the 5G model using a deep learning algorithm to train and predict a slice for a device based on the information calculated from previous connections. The dataflow was treated with a high-level API to build and train the data using TensorFlow. Their machine learning model was also done based on supervised learning because their dataset was big and structured. As a result, a Random Forest (RF) algorithm and Convolution Neural Network classifier (CNN) have been used for traffic classification. Moreover, the authors utilized various parameters to determine the Network Slicing, such as slice type, bandwidth, throughput, latency, equipment type, mobility, reliability, isolation and power. Their focus was to select a slice for a device and select enough resources for the slice based on the traffic prediction. Furthermore, their model did not deal with realtime prediction or with other slice management scenarios. In [61], their prediction for the resources scheduling done after the offline learning. On the other hand, in [63] nonconvex problems have been proposed to control the Network Slicing in the radio access network with admission control to satisfy the QoS inter-slice.

The inter-slice algorithm was proposed in [64], their method mapped a new user to a specific slice according to their demands and predicted the target slice for their new users the historical behaviour of their resources. Further, they reduced the slice creation cost and enhanced the QoS for the users within the slice. Another research done by Han *et al.* [65] applied the genetic algorithm to optimize the resources for inter-slice. The slicing strategy is done by using a binary sequence for accepting or declining the slice. From the decision entry presented in [65], it is not clear what will happen if different types require the accepted slice of resources. In [66] Han *et al.* proposed an architecture due to the 3GPP standardization and their Cross-slice admission and congestion controller implemented inside NSMF to provide E2E resources allocation to the different slices. The Q-learning was used to train a model with different priority services for the slice.

Jiang *et al.* [67] and Hu *et al.* [54] proposed both inter/intra slice in their systems to optimize the radio access network using heuristic and slice selection schemes. Furthermore, the inter/intra slice priority was discussed in [67] to enhance the QoE for the user by providing a dynamic slicing approach for the virtual resources on top of the physical infrastructure using a heuristic-based scheme. Another multi-tenant architecture was presented in [68] to manage the resources over RAN using reinforcement learning which is similar to the work in [69], [70] for the resource allocation scenarios in the RAN and priority for slice using deep reinforcement learning.

In the Lightweight network, the number of resources will be managed using a scheduler for the inter-slice to map virtual resources to physical infrastructure and for the intra-slice will allocate these resources to the users within the slice [81]. Adding to that, policy control for OFDMA-based over RAN in virtual networks studied in [76] to support the requirements of slices by using slice provisioning algorithm for intra-slice. Moreover, the isolation between these slices has not been mentioned in their work. Another paper focused on slice isolation done by Oliveira and Vazão in [78] which provided the slice isolation among different verticals by using a priority adaptation slice scheduler algorithms for the uplink phase in a RAN.

Slice allocation for intra-slice isolation and E2E delay explained and used Mixed-Liner Integer programming (MILP) to optimize the model in the core layer for the future network but not with inter-slice allocation. Additionally, the results show that the minimum requirement has been met within the E2E delay, and more bandwidth was required within the restriction in intra-slice isolation to get lower CPU utilization [80]. Hossain and Hasan [82] highlighted the significant issues related to network performance in terms of measurement and testing of the resources to cover all the requirements in the network virtualization. As mentioned before, the resource demand in a virtual environment has covered a set of resources such as a CPU, storage capacity, bandwidth and switching network. Moreover, dynamic resources allocation is still a challenge in the service continuity in future network [23].

From a business perspective, in [71], [77], the resources allocation was provided for different verticals such as providers, tenants and end-users depending on the admission policies to increase the gain of the network providers. Also, in [72], their heuristic approach reduced the virtual link and node cost. Furthermore, the auction mechanism was designed in [73] to enhance the QoE for the user within the slice in RAN by maximizing the network income as it was purposed in [75], but by using reinforcement learning. On the reverse, in [74] authors proposed an algorithm to minimize the service reward by using Stackelberg schema. Furthermore, An *et al.* [10] proposed a slice management algorithm based on the price of a service indicating the service priority, divided into a high or low category. Moreover, according to the telecommunications companies and researchers' point of view, working on the 5G core based on virtualization technologies SDN/NFV to provide a dynamic service chain will improve network flexibility, scalability and reliability. Besides, reducing the systems' cost by controlling the number of working servers in their data centre.

As a result, a significant amount of research has worked to enhance the service demands presented in Table 1, such as the enhancement of the QoS over the wireless interface for the 5G networks. On the other hand, few of these papers concentrate on the core network to control and optimize the network's traffic flow using machine learning or other programming criteria with SDN-based. Moreover, Table 1 discussed the slice resources and policies that the slice needs when the user moves from one slice to another. Furthermore, Mobility as a service is managed based on a set of parameters, such as priority, cost, QoS and the connection type: TCP or UDP to map the user to specific slice according to their demands.

G. MOBILITY MANAGEMENT AS A SERVICE (MMaaS) Seamless handover for 5G in Network Slicing needs to be managed by developing a new mobility mechanism to enhance the QoE for the user and the service's continuity. In the future, network operators will share network infrastructures with Internet Service Providers (ISPs) flexibly and efficiently, which leads to the concept of ‘‘Slice as a Service’’ [65]. Furthermore, Network Slice as a Service (NSaaS) was introduced in [83] to provide a customized E2E network as a service for their customers. With NSaaS, there are three different service scenarios: Industrial slice, Monopolized slice and Event slice. In addition, NSaaS is a service auto-mapping that is automatically mapping service requirements according to their customer demands to various parameters of Network Slicing to do auto-configuration for network functions and services.

Future cellular systems will be virtualized and programmed using SDN/NFV. Moreover, the SDN/NFV integrated into future networks as enablers to do auto-configuration, autoscaling for network functions and services. Adding up, the AMF is responsible for handling the mobility, reachability of the UE. This entity is also in charge of managing the Next-Generation Application Protocol (NGAP) signalling to control the handover, carried between the AMF and radio access network, described in 3GPP TS 38.413.

Mobility management depends on several parameters such as network, terminal, user and service. Handover provides uninterrupted connection when the subscriber moves away from one cell and to another cell. In addition, network interface handover classified to two types: Horizontal and Vertical handover [84]. Handover remains an issue for both heterogeneous and non-heterogeneous networks. In addition, several problems arise in the heterogeneous network, such as Handover Failure, Cell Association Issue and Ping Pong effect. In handover Failure, it can be reduced by adequately managing interference in the cell with some interference management techniques to reduce handover failures. In cell association issue, It can be solved by the Cell Range Expansion Technique. On the other hand, the ping pong effect can be reduced by smartly managing the Handover phenomenon based on cell association area [85].

In LTE, the MME is responsible for handling the users who are attached to the network. In real life, without mobility management, consumers would need to get a new SIM card while travelling from one location to another; otherwise, the users will lose their services. On the other hand, exponential growth in network traffic plus the number of users has led to the meaning of MMaaS to the wireless nodes attached to the 5G new radio networks. The application of mobility management could be implemented on the SDN controller using SDN/NFV. In addition, using NFV, several slices introduced to make it works for 5G networks to accommodate various types of services [86]. Furthermore, these slices will be isolated from each other. Besides, each slice has different demands such as latency, speed and interference, and these requirements are controlled by the slice configuration and service characteristics. After the mobility management implementation, itwill be possible to control the handover requests and procedures with the Network Slicing. Moreover, the essential points of using mobility management reduce the tunnelling overhead and minimizing the handover latency [87].

Gharsallah *et al.* [88] designed a new approach called Software-Defined Handover Management Engine (SDHME) to solve the mobility management challenges such as the handover delay and failure in 5G ultra-dense networks. Their approach for the Software-Defined Handover Management consisted of four phases: data gathering, data processing, virtual cell (V-Cell) creation, and handover execution. Afterword, their mechanism has been implemented in the control plane, defined in the application plane and executed in the data plane based on the SDN architecture. Moreover, their simulation result reduces the handover delay and failure.

Ravindran *et al.* [89] focused on managing the mobility services within the slice dynamically over Information Centric Networking (ICN) in 5G using mobility as a service (MaaS). Amadeo *et al.* [90] discussed the roles and the

TABLE 1. A Summary of the slice resources and policies reported in literature.

References	Inter Slice	Intra Slice	QoS	QoE	Priority	RAN	CN	Optimization Algorithm
Tang et al. [61]	✔	✖	✔	✖	✔	✔	✖	Constrained Markov Decision Process (CMDP)
An et al. [10]	✔	✖	✔	✖	✔	✔	✖	K-hop Greedy routing
Tang et al. [58]	✔	✖	✔	✖	✖	✔	✖	SCA, SDR Algorithms
Addad et al. [22]	✔	✖	✔	✔	✖	✔	✖	Greedy Distributed Multi-layer Knapsack (GDMK)
Kim et al. [60]	✔	✖	✔	✖	✖	✔	✖	Q-learning
Zhu et al. [20]	✖	✖	✔	✖	✖	✔	✔	Decision Tree
Soliman and Leon-Garcia [63]	✔	✖	✔	✖	✖	✔	✖	Heuristic
Bega et al. [71]	✔	✖	✔	✖	✔	✔	✖	N3AC
Han et al. [66]	✔	✖	✔	✖	✔	✔	✔	Q-Learning
Jiang et al. [67]	✔	✔	✔	✔	✔	✔	✖	Heuristic
Yi et al. [72]	✖	✖	✔	✖	✖	✔	✖	Heuristic
Natalino et al. [68]	✔	✖	✖	✖	✔	✔	✖	Reinforcement Learning
Hu et al. [54]	✔	✔	✔	✖	✖	✔	✖	Heuristic
Jiang et al. [73]	✔	✖	✔	✔	✖	✔	✖	Auction-based mechanism
Hu et al. [74]	✖	✖	✔	✖	✖	✔	✖	Stackelberg Game
Kammoun et al. [64]	✔	✖	✔	✖	✔	✔	✔	Overload Cost and Requirements based algorithm (OvC&R)
Raza et al. [75]	✖	✖	✔	✖	✔	✔	✖	Reinforcement Learning
Li et al. [69]	✖	✖	✖	✔	✔	✔	✖	Deep Reinforcement Learning
Sun et al. [70]	✖	✔	✔	✖	✖	✔	✖	Heuristic and DQN Algorithm
Jian et al. [59]	✔	✔	✖	✔	✔	✔	✖	Greedy Largest Utility Algorithm (GLUA)
Parsaeefard et al. [76]	✖	✔	✔	✖	✔	✔	✖	Slice Provisioning Algorithm
Shurman et al. [55]	✖	✖	✔	✖	✔	✔	✔	✖
Bega et al. [77]	✖	✔	✔	✖	✔	✖	✖	Q-Learning & Markov Decision
Han et al. [65]	✔	✖	✔	✖	✔	✖	✖	Genetic algorithms
Fattore et al. [56]	✔	✖	✔	✖	✖	✖	✔	✖
Oliveira and Vazão [78]	✔	✖	✔	✖	✔	✔	✖	PASS Scheduler Algorithm
Yousaf et al. [18]	✔	✖	✔	✔	✖	✖	✔	Dijkstra’s algorithm
Ma et al. [79]	✖	✖	✔	✖	✖	✖	✔	Workload allocation Algorithm
Sattar and Matrawy [80]	✖	✔	✔	✖	✖	✔	✔	Mixed-Liner Integer programming
Choi and Park [57]	✔	✔	✔	✖	✖	✔	✔	NAS & slice selection Algorithm

principles of the ICN-5G in wire and wireless scenario. Here, the ICN distributed approach highlighted the mobility services and the virtualized network functions in the 5G new radio network based on the fixability of the services offered by the Network Slicing using the SDN/NFV features.

The mobility management mechanism is classified and selected according to the service context, defined as a stateful service [18]. New mobility mechanism needs to be developed to manage seamless handover for 5G in Network Slicing to enhance the service’s continuity and scalability for the user experience. The railway network will have many handovers because of the high-speed of the train[91], but the IoT devices will have handover in a low latency service.

Different types of slices have been identified to cover the user requirements. Therefore, mobility happened between different base stations and between the slices of diverse service types such as eMBB, URLLC and MIoT [92]. Furthermore, even when the service type determined by the use case on the application layer, this service should be continuing when the user moves from one location to another. For that reason, Addad *et al.* [92] grouped the slice mobility into different patterns: user subscription type, network slice type, accessing a type, service area, access characteristic and geographical location. Moreover, slice mobility divided into different types: full slice mobility, slice breathing, slice splitting and slice merging. Therefore, a slice will be modified according to service consumption. Furthermore, the number of use cases introduced in [92] to cover the slice mobility types are listed below.

1) First use case: When a swarm of drones move as one group in the same direction outside the initial service range, this leads to full slice mobility. On the other hand, when the swarm move one by one or as small groups from the current service range, the dedicated slice is divided into two service areas. In both cases, drones need to be controlled accurately with low latency.
2) Second use case: Service continuity is significant for autonomous vehicles because messing a connected car signal could cause physical harm or even death. In this case, the backup slice should be activated in a reallocation time to ensure the service's availability during the migration of the primary slice then the backup slice do the reallocation at a different time.
3) Third use case: The audience attends video streaming, and some of these audiences on the train and the area around the train, including tunnels, has less capacity and service. In this case, the new slice needs to be created temporally to serve the audience. When these audiences arrive in the area with better service, the temporal slice will be released dynamically, and the original slice will serve the audience. This case called slice breathing which is a compensation of the slice splitting and slice merging.

For a different approach, in the next-generation mobile network new schema for managing the mobility in the service level needs to be created to enhance the networks [19]. Furthermore, based on the knowledge of the authors, there is no existing research that works on the mobility in service level in a real-time by providing requirement resources for the slice after generating the service request and accepting this request by the controller and give a specific VNFs that dynamically related to these services.

III. INTER/INTRA SLICE MANAGEMENT TECHNIQUES In the inter/intra slice, the network resources will be managed virtually and dynamically using virtualization techniques. This section will categorize the research's main dimensions by identifying the service types within the slices number based on 3GPP TS, the existing driven models, and the current systems.

A. SLICE TYPES AND THE NUMBER OF THE SLICES BASED ON 3GPP TS

Most studies have only focused on Network Slicing's main parts, which is categorized into three essential types: eMBB, URLLC and mIoT. On the other hand, some of these studies are added new types for V2X or emergence situation, and if the connection failed with the main slice, the subscribers would be forwarded to the default slice. None of the research papers discovers the actual number of slices that the future network capable of. Consequently, this section will explain these types based on the 3GPP TSs.

The 3GPP defined Network Slicing as a logical network built on top of the hardware devices that provide unique capabilities and components. Additionally, the 3GPP defined NSI as a set of functions and resources that will be used

In [93], the MME has been implemented in the application layer inside the control plane to manage the mobility in the user plane to provide a low latency service for the users to optimize the handover procedure within 5G systems, which is similar to the solution in [94] for the handover optimization using the SDN/NFV with distributed mobility management in a control plane. Moreover, MMaaS needs more attention for the Next generation of the mobile telecommunication network. Furthermore, Zhang *et al.* [40] classified mobility management into two procedure based on the location registration and handover management. The mobility management is not just about the handover between different networks; it is the handover between different services.

In a recent review of the literature on this area in [17], the authors reviewed the mobility management signalling steps based on the 5G architecture to match the requirement of the QoS by addressing different scenarios in terms of latency, scalability and throughput with different service types eMBB, URLLC and mMTC. Moreover, in [22], a set of policies attached to the slices without dealing with mobility among the slices. In comparison, mobility is the primary key in real-time scenarios to provide a better result for the networks' user experience.

to expand Network Slicing, which is deployed on top of the physical infrastructure. In 2016, the 3GPP start working on the Network Slicing specification for the future network according to the 3GPP TR 23.799. The Network Slice Function and Architecture presented in TR 23.501. The Network Slice Service Function and Management proposed in TS 28.531. After that, the management specification identified in TS 28.530. All the previous work for the Network Slicing in [42], [95]–[97] are still in progress.

Within the 3GPP TR 28.801, the NSSI implementation concept was explained. The NSMF send a request to the NSSMF to generate the NSSI. The NSSI depends on the physical and logical resources in term of availability and reliability, and it is making progress on run-time operations [45].

In 3GPP TS 23.501 version 15.2.0 Release 15, Single Network Slice Selection Assistance Information (S-NSSAI) identifies the Network Slicing. Moreover, the Network Slice Selection Assistance Information (NSSAI) is a set of S-NSSAIs. From the 3GPP regulation, each user could be served eight slices simultaneously when each request goes to a specific NSI. The S-NSSAI comprises two portions: Slice/Service Type (SST) and Slice Differentiator (SD). The
SST standardized values are shown in Table 2. In addition,
the SST referred to the predicted Network Slicing performance concerning services and features. On the other hand, the SD represented as the optional information complementing the SST(s) to distinguish between multiple Slice/Service network slices of the same type.

The S-NSSAI type was defined by the 3GPP TS 29.571 v15.6.0 Release 15. The SST was defined as an unsigned integer with a range from 0 to 255, and the SD was defined as an integer with a range from (0-9) or (A-F), as shown in Table 3.

Are cent survey of the literature on this topic done by Sajjad *et al.* [98] which reviewed the 3GPP TS 23.501 and 23.502 in term of the mobility management in 5G within the inter-slice handover in a service level. The authors of [98] recommended research directions and mentioned the continuity between the slices is not given in the 3GPP TS yet. Furthermore, these slices will be registered to the Allowed (NSSAI) list, which contains all the subscribed slices to the UE. Moreover, the UE could access all resources with in the slices. The number of the slices in the Allowed list and the S-NSSAIs portions identified by the 3GPP TS 29.571 v15.6.0 Release 15. Moreover, the network administrator or dynamically list modification could be done by the controller by adding some configuration based on the UE request within the PDU session management [98].

TABLE 2.

Standardized SST value.

Slice/Service type	SST value
eMBB	SST 1
URLLC	SST 2
MIoT	SST 3

TABLE 3. SST and SD values.

Name	Data type	Range	Description
SST	Unsigned Integer	0 - 255	From 0 to 127 representing to the SST range. From 128 to 255 representing the Operator Specific Range (QCI). According to 3GPP TS 23.203, the QCI means packet forwarding treatment which is used to ensure the traffic bearers and each bearer required a QoS and QCI. Also, mapping the access layer to the network layer.
SD	Integer	(0-9) or (A-F)	Consists of 3 octets, each octet contains 1 byte which equal to 8 bits. The First 4 bits for 0 to 9, and the second 4 bits for A-F, as represented in Figure 9.

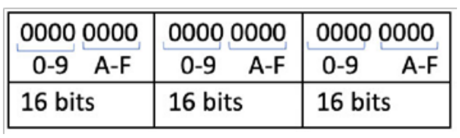


FIGURE 9. Slice differentiator (SD) bits.

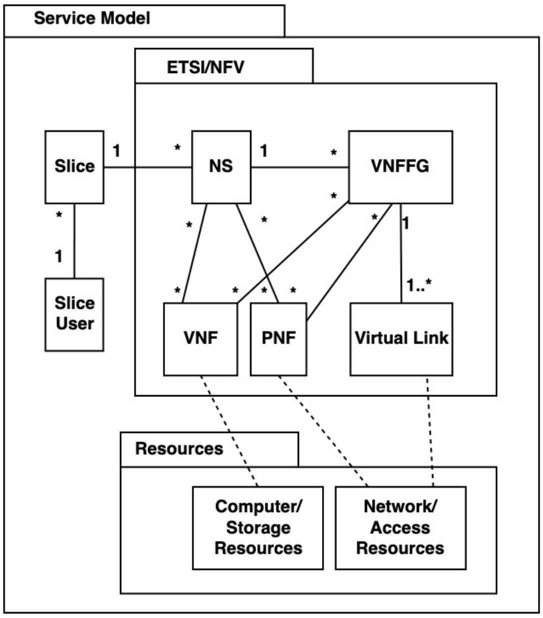


FIGURE 10. Service Model [41].

In Figure 9, the 16 bits represent the numbers from zero to nine in a binary plus the hexadecimal number from A to F in binary to get 16 bits for each octet to represent the SD.

By combining these results, it could be concluded as: the SD => 16 * 16 * 16 = 4096. The S-NSSAI comes from the SST * SD => 256 * 4096 = 1,048,576 number of slices that could be used in the future network according to the 3GPP TSs.

B. TYPICAL SYSTEM MODELS

Slice service model in Network Slicing has been explained in ETSI 2019 [93], the basic concepts for modelling network slices in the core network built around the Network Services (NSs). These services are identified to service-driven model deals with network services and deployment-driven model, which deals with network resources, services and mobile services.

The service model consists of sets of resources such as a computer, storage and network. These resources will belong to the slice to provide a set of services that the slice needs. In Figure 10, Slice consists of one or more NSs consisting of VNFs, PNF, Virtual Links (VLs), and VNF Forwarding Graphs (VNFFG). This model structure with the relationship defined by the ETSI NFV standard. Network resources contained physical and virtual infrastructure, which will exist for specific slice within the run time of NS [41].

Deployment model [41] consists of sets of physical resources linked to network services like mobile Edge application instance, VNFs instances or directly connected to the

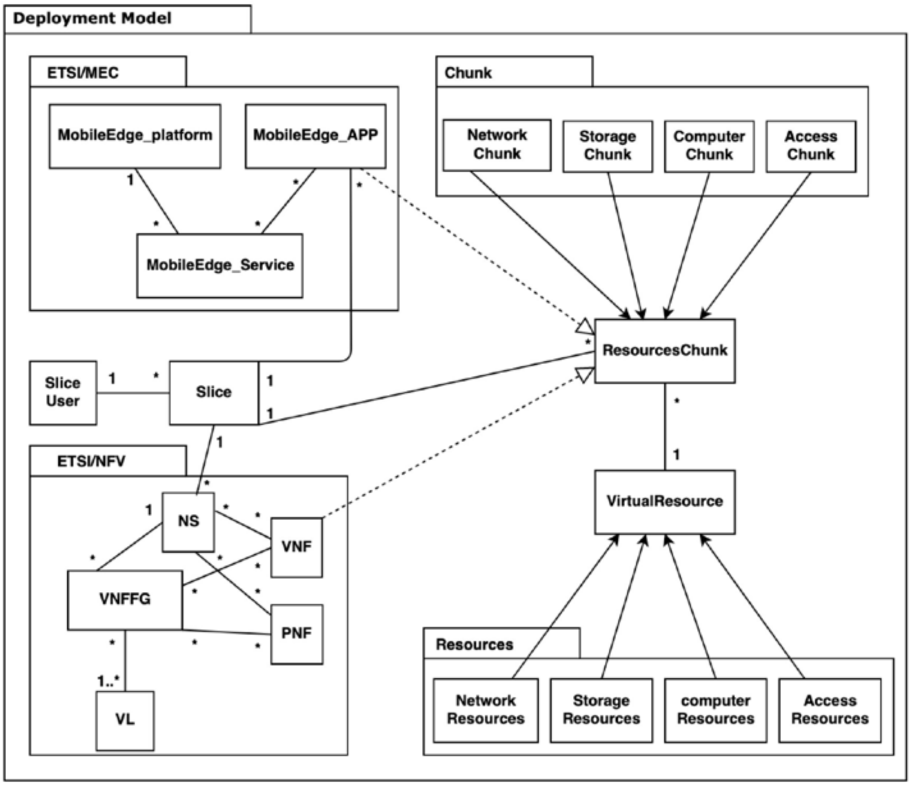


FIGURE 11. Deployment Model [41].

slice model. The model structure shows in Figure 11, where slices are connected to the NSs in the same way in the service model. Also, slices are connected to infrastructure resources. This model represents a hybrid approach. Furthermore, the VNFs instances will run on the resources chunk. Besides, physical resources could exist without NSs. Moreover, an internal model could be added to the slice model to run on a resources chunk, such as a mobile Edge application with a similar deployment to the NSs.

C. CURRENT EUROPEAN PROJECTS

the authors proposed system architecture and explained each layer's benefit in their system. Moreover, there is no simulation result. They extended the work in [101] to manage and control the network over multi-administrative domains to maximize the network resources that share across the network and reduce the capital expenditure. The main goal of the 5GPPP in the SliceNet platform builds E2E Network Slicing in multi-domain to optimize the capability of the QoS/QoE by using network controller style in a friendly softwarization environment. Moreover, this project's configuration capability will be used with a wide range of vertical

The 5G City project's architecture has been addressed in [99], the Network Slicing concepts have been implemented in different layers, including access, core and cloud. The aim of this platform allocated new service dynamically for each slice. Moreover, this project will run multiple logical networks over multiple-tenant infrastructure, responsible for managing the service's lifecycle dynamically based on the slice policy function. Furthermore, the 5G City project achieved a few seconds within the slice creation. On the other hand, internal services mean the service between the customer and the service provider has not been implemented in this prototype.

SliceNet project in [100] applied software-based to develop the control plane's functionality to enhance mobility management and QoS support. They reduced the signalling cost comparing to the LTE system, but Network Slicing has not been implemented in the presented model. In 2018 [114], use cases such as automation, eHealth,and smart city. In addition, the data plane has not been developed in the SliceNet prototype.

5G-EmPower platform [102] built upon an open-source project to provide flexibility and scalability to the RAN, but not for the core network. The 5G-EmPower platform worked on three use cases: radio access network slicing, mobility management and load balancing. The system's scalability is based on open-source code, which has a limitation in the system support and the CPU utilization reached between 60% to 70% for two users. Furthermore, the 5G-EmPower developed an eNodeB simulator to connect 50 eNBs to 100 UEs to evaluate the 5G-EmPower operating system's scalability at the SBI; the CPU utilization reached between 70% for 5000 UEs. This project's limitations are slice specification schedulers and the admission control to set policies for the slice.

Limitations of the existing European projects.

System Name	Deployment layer	Network Slicing	Model Type	NBI/SBI	Limitation/ Further Developments
SliceNet [100]	Control/Data plane	Not included	Simulation	Not specified	5G entities need to be used not just depending on SDN-based to reduce the cost compering to LTE.
SliceNet [101]	Control plane	Included	Simulation for the core layer	Not specified	Data plane implementation need to be developed to enhance the system performance.
5G city [99]	Access/Core layer	Included	Simulation	NBI/SBI	Internal services require more research to be implemented in 5G city system.
5G-EmPower [102]	RAN	Included	Simulation	SBI	Core Network, set polices and slice schedulers represent as a fundamental demand that needs to be involved in future network accomplishment.
CellSDN [103]	Control/Data plane	Included	Prototype	SBI	Set of police services have not been developed in the CellSDN and needs more investigation and research to make it work with real mobile system.
SoftMow [104]	Control/Data plane	Included	Simulation/Prototype	SBI	The SoftMow prototype is not suitable for real-time environment.
SESAME [105], [106]	Core Network/RAN	Included slices for the small cell	Simulation	NBI	The SESAME prototype is not suitable for real-time environment. Moreover, they need to enhance the QoE for the end-user.
5G NORMA [107], [108]	Core Network/RAN	Included	Architecture design for the system	NBI/ EBI/ WBI	Still under development to deal with inter/intra slice control.
5GEX [109]–[111]	Control/Data plane	Not included	Architecture framework/ Prototype	SBI/NBI	E2E multi-domain Network Slicing needs more consideration, investigation about the security and provide a decentralize controller. Also, add more effort on the service continuity.
CogNet [112]	Core network	Not included	Architecture	Not specified	Prototype or simulation need to be provided to proof the proposed machine learning algorithm when it is work in real-time learning to enhance the QoS/QoE for the services and the end-users.
COHERENT [113]	SDN controller	Included	Architecture design	SBI	Prototype or simulation model will be required to proof the COHERENT aims for controlling the traffic and managing the resources.

5GINFIRE is a European project that aims to be finished in 2020 [115]. The 5GINFIRE project proposed many challenges within the management and orchestration (MANO) platform to enable network services on top of the VNFs and verticals in the 5G ecosystem. The 5GINFIRE project targets different verticals, mainly when the Internet of Things (IoT) and programmability network are applied in different scenarios such as eHealth, automotive and smart cities that need to be connected in the future network.

Some of these systems are developed to manage mobility dynamically with different services at a service level. In addition, other projects are enhanced mobility management by implementing a set of functions in a control plane to enhance the mobility between different use cases in the user plane. Moreover, the system's limitations have been identified in Table 4 with the recommended solutions to optimize the system performance.

IV. TRAFFIC CLASSIFICATION FOR THE FUTURE NETWORK

The priorities of the slice and user per service within the slice will be changed based on the state of the network in an unexpected situation such as disasters time that needs an emergency contact[116]. Furthermore, most of the published work in literature identifies three types of slices and many types of services for each slice [17], [40], [42], [81]. Thantharate *et al.* defined an alternative type of the slice called master slice to adapt the user request until the deep learning model arranges a specific slice for the UE based on the demands after the network failures [62]. On the other hand, other researcher added a new slice type called High-Performance Machine Type Communications (HMTC), which works with applications in a high data rate, low latency and high availability, as shown in the ATIS report.

When the user mobile from one slice to another will be based on a set of parameters, one of these parameters will be the traffic type. Therefore, we need to classify the traffic to manage the inaccurate mobility way. This section will classify the traffic for future networks according to the requirement, as shown in Table 5. Moreover, we could analyse the classification principles to different ranks based on the variable's priority:

1) First rank for route (same route or different route).
2) Second rank concerning with TCP or UDP.
3) Third rank which includes bandwidth and delay. 4) Forth rank premium (low, medium and high).

TABLE 6. Limitations of existing work along with possible solutions.

V. OPEN RESEARCH ISSUES AND FUTURE DIRECTIONS

Most of the research papers identified sharing resources per slice, controlling the network slice dynamically by using a set of policies, managing the priority of a slice and service that the user belongs to, and other challenges related to slice management still open issues. Aspects related to the security of Network Slicing is outside the scope of this paper hence not considered. However, future networks should consider security issues regarding the users' potential in the slice to introduce any risk to the system. Consequently, security will be another challenge for the 5G system that needs robust network policies to protect the network from any risk or threats [48].

Mobility is one of the critical issues at a service level for future network. Moreover, mobility management is defined in two-mode modes: idle mode for the user reachability and connected mode for the handover [117]. Moreover, the 3GPP discussed how to control the level of mobility in their TS document [95]. Additionally, Addad *et al.* [118] implemented mixed-integer linear programming to solve this problem with a fixed number of nodes. The authors also recommended heuristic algorithms that could solve the problem by distributing the nodes in polynomial time. Further, mobility needs in depth research to be elaborated for real-world applications to satisfy user demands [118].

Zhang *et al.* [40] implemented a handover mechanism and resource allocation to enhance the scalability and flexibility in 5G networks under Network Slicing for macrocells and small cells. Furthermore, the authors identified the open issues that need to be solved in the future communication system, such as resources sharing within inter/intra slice, QoS/QoE, polices and load balancing at the network level. Besides, cooperating with a service level in Network Slicing. Plus, considering the union between multiple RATs to provide seamless mobility and high throughput. Although slice management is still an issue significantly when the complexity of services and applications increase [21]. In addition, Hossain and Hasan [82] mentioned numerous research issues, and some of these issues remain a challenge until now. These include inter-slice isolation, intra-slice resource allocation that need sufficient design and functionality, including how modern networks make these issues work together. In addition, some of the existing research with their limitations have been identified in Table 6 with the recommended solutions to improve these issues for further developments.

TABLE 5. Traffic classification.

Traffic Types	Explanation
Live Streaming	In this type, all packets need to be transmitted in the same route in high bandwidth and small delays.
Video	Receiving packets need to be in sequence order because if any delay happens in one packet, this packet will be discarded. Therefore, it is required the same route for sending and receiving the packet and demand bandwidth. Dealing with video is represented as the biggest challenge for the telecommunications companies in the future network because the video has two fundamental requirements.
Social Media	Normally, it is one-way communication for all platform. It needs a high bandwidth because all users upload many pictures and videos that others need to see, and the delay is not an issue.
IoMT	The Internet of Medical Things (IoMT), IoT sensors are used to send emergency data to the hospital or to the doctor to track the patient health condition, which needs a bandwidth with significantly no delay because if the patient loses heartbeat for one second, the patient will be dead. Few bits with no delay on the path or path have no congestion at all. Plus, it is expected no loses by noise. Moreover, this traffic needs to be sent in mile seconds and get an acknowledgement message to confirm that other parties receive the message.
IoT	It is essential for the 5G era, dealing with IoT, which have millions of sensors, but they produce a minimal amount of data. Accordingly, it does not require extra bandwidth, long or short route, high bandwidth or delay. Therefore, these data necessitate being delivered without an acknowledgement message.
Email	The maximum amount of the file attached in it is 250 MB. It is not a large amount of data. Hence, no needs for an acknowledgement message or high bandwidth or small delays.

The network resources optimization will be based on QoS/QoE. These two issues set the rule for managing the network in term of mobility management, resources allocation and Network Slicing [16], [19]. Additionally, the QoS will be higher within the future network when the same amount of the resource is required by different slices with different priority values. Furthermore, changing the amount of the recourse allocation for each slice dynamically based on the slice priority without effecting the QoE for the user within the slice is a significant point to enhance the performance and minimize the latency for future systems. In addition to the above research issues, future directions will be listed below for further exploration:

BLE 6. Limitations of existing work along with possible solutions.

References	Existing Research	Recommendation approach for future improvement
[10]	An et al. proposed a slice management algorithm based on a service's price indicating the service priority into high and low categories.	Alternative approaches will prioritise the situation of the service and the user needs, especially for the emergency. On the other hand, for the typical situation, the controller will choose a batter path to send the user's request with a minimum amount of delay.
[38]	Gudipati et al. implemented an algorithm to fix the resources sharing per slice. Nevertheless, the slice owner cannot decide to provide better allocate resources to his customer within the slice.	It is better to optimize the resource allocation by providing better QoS to cover all the demands without effecting the QoE for the users.
[119]	Bordel et al. proposed Inter-slice management based on graph theory and saving the slice in a queue based on the priority that happened based on the probability event. The authors provided predictive solutions to evaluate the network and improve the QoS in future networks.	A more practical solution for this problem is developing an algorithm that deals with inter/intra slice management in real-time to control and manage the slice and the service within the slice dynamically.
[19]	Barakabitze et al. mentioned that inter and intra slice management and NFs placement need significant efforts to let the Network Slicing approach be used in the next-generation mobile networks.	For the future, researchers need to do more investigation to implement and manage the inter/intra slice in an efficient way.
[17]	Akkari et al. focused on the 5G specification to review the mobility signalling steps to satisfy the QoS demands with different service types in terms of latency, scalability and throughput.	Seamless handover decision algorithm needs to be built in the SDN controller to satisfy the QoS demands within a wide range of IoT applications as a recommended solution to enhance the signalling phase and reduce the delay after investigating the service continuity.

- Multiple controllers need to be implemented in the control plane to improve the scalability and the availability for the 5G services when the SDN controller manages these services.

- NBI connects between the SDN controller and the 5G entities when implemented in the application plane. Therefore, the REST API in the NBI needs more attention. Moreover, the West/Eastbound Interface need further consideration to improve the scalability of the controller.
- With mobility management as a service, real-time learning needs more attention to enhance the QoS/QoE in the 5G network by implementing machine learning algorithms for this purpose.
- In real-time, policies and services should be provided to the slice according to the demands. On top of that, Network Slicing should be managed dynamically in this operation.
- Classification for the services and traffic in slice needs to be considered by using, for example, a machine learning approach.
- Inter-slice resources allocation and the scalability for the resources from one slice to another need attention.
- Isolation among the slices in terms of services and functionality at the beginning of the slice implantation represents an essential research direction to improve the networks' security part.
- Slice selection discussed, which needs more investigation in the slice selection type (selective and sequential). If the system applies the particular type, it needs to identify when the slice will be selected and why. If this system utilises the sequential type, then it needs to find the benefit of this type. Also, why future systems need to use slice selection?. Furthermore, what is the interest of the telecommunication companies in terms of cost and energy?

To sum up, these open issues and research directions are critical and require more investigations to optimally use the capabilities of the SDN/NFV in future networks without any limitations.

VI. CONCLUSION

Network Slicing technology is a significant component for the future network, enhancing the data transmission's reliability and reducing the transmission delays and minimizing small fluctuations and small errors in the network. This paper reviews the trending technologies and summarises the existing research and systems with the current systems' limitations. Besides, according to our review, slice management is essential in inter/intra slice scenarios. This is the first survey paper to discuss and summarize the slice management in service continuity in terms of policies, services, and functionality to the best of the authors' knowledge.

Studies into existing literature presented in this paper identify that managing the Network Slicing dynamically in realtime operation with softwarization techniques are missing in the current systems. Furthermore, network providers currently investigating how softwarization techniques will be evolved with cellular systems, besides how the virtualized functions and services from one provider will cooperate with other providers to move to their services dynamically without affecting the QoE for the subscribers.

In future, mobile networks for inter/intra-slice and mobility management at a service level represents crucial features that influence the QoS in real-time as it is provided to the users based on the user's expectations. Based on the research that has been done in this paper, mobility management as a service is still an open issue and under investigation by telecommunication companies and researchers, especially at a service level for the real-time application within the core and radio access layer. Consequently, future work includes optimizing resource management and service continuity to meet the QoS standard and maximize the QoE for the user.

ACKNOWLEDGMENT

This research was funded by the Ministry of Higher Education and Scientific Research in the Republic of Iraq in (06/09/2017) to sponsor Noor A. Mohammedali to pursue her PhD research.

REFERENCES

[1] B. A. A. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka, and T. Turletti, ‘‘A survey of software-defined networking: Past, present, and future of programmable networks,’’ *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1617–1634, 3rd Quart., 2014.

[2] M. Pérez, N. Losada, E. Sánchez, and G. Gaona, ‘‘State of the art in software defined networking (SDN),’’ *Visión Electrónica*, vol. 13, no. 1, pp. 178–194, 2019.

[3] *Network Function Virtualisation (NFV); Management and Orchestration*, European Telecommunications Standards Institute, Management and Orchestration, Boston, MA, USA, 2014, p. 51, vol. 1.

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

[5] M. Paul, S. Schallen, M. Betts, D. Hood, M. Shirazipor, D. Lopes, and J. Kaippallimalit, ‘‘Applying SDN architecture to 5G slicing,’’ *Open Netw. Fundation*, no. 1, pp. 1–19, 2016.

[6] W. Xia, Y. Wen, C. H. Foh, D. Niyato, and H. Xie, ‘‘A survey on software-defined networking,’’ *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 27–51, 1st Quart., 2014.

[7] S. Kazmi, L. Khan, N. Tran, and C. Hong, ‘‘Network slicing: The concept,’’ in *Network Slicing for 5G and Beyond Networks*. Cham, Switzerland: Springer, 2019, pp. 13–24.

[8] F. Fitzek, F. Granelli, and P. Seeling, *Computing in Communication Networks: From Theory to Practice*. New York, NY, USA: Academic, 2020.

[9] X. Li, C. Guo, J. Xu, L. Gupta, and R. Jain, ‘‘Towards efficiently provisioning 5G core network slice based on resource and topology attributes,’’ *Appl. Sci.*, vol. 9, no. 20, p. 4361, Oct. 2019.

[10] N. An, Y. Kim, J. Park, D.-H. Kwon, and H. Lim, ‘‘Slice management for quality of service differentiation in wireless network slicing,’’ *Sensors*, vol. 19, no. 12, p. 2745, Jun. 2019.

[11] V. Wong, *Key Technologies for 5G Wireless Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[12] O. Queseth, O. Bulakci, P. Spapis, P. Bisson, P. Marsch, P. Arnold, P. Rost, Q. Wang, R. Blom, S. Salsano, T. Chen, B. Teodora Sandra, U. Herzog, V. Frascolla, X. Li, and Z. Yousaf, ‘‘5G PPP architecture working group: View on 5G architecture,’’ 5G Infrastruct. Public Private Partnership, Brussels, Belgium, Tech. Rep. 2.0, 2017.

[13] *Network Functions Virtualisation (NFV) Release 3;Evolution and Ecosystem; Report on Network Slicing Support With ETSI NFV Architecture Framework*, document ETSI GR NFVEVE 012, V3.1.1, European Telecommunications Standards Institute, Dec. 2017.

[14] *Network Functions Virtualisation (NFV) Release 2; Management and Orchestration; Network Service Templates Specification*, Standard ETSI GS NFV-IFA 014, V2.3.1, European Telecommunications Standards Institute, Aug. 2017.

[15] *Network Functions Virtualisation (NFV) Release 3; Reliability; Report on NFV Resiliency for the Support of Network Slicing*, Standard ETSI GR NFV-REL 010, V3.1.1, European Telecommunications Standards Institute, 2019- 06.

[16] V.-G. Nguyen, A. Brunstrom, K.-J. Grinnemo, and J. Taheri, ‘‘SDN/NFVbased mobile packet core network architectures: A survey,’’ *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1567–1602, 3rd Quart., 2017.

[17] N. Akkari and N. Dimitriou, ‘‘Mobility management solutions for 5G networks: Architecture and services,’’ *Comput. Netw.*, vol. 169, Mar. 2020, Art. no. 107082.

[18] F. Z. Yousaf, M. Gramaglia, V. Friderikos, B. Gajic, D. von Hugo, B. Sayadi, V. Sciancalepore, and M. R. Crippa, ‘‘Network slicing with flexible mobility and QoS/QoE support for 5G networks,’’ in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2017, pp. 1195–1201.

[19] A. A. Barakabitze, A. Ahmad, R. Mijumbi, and A. Hines, ‘‘5G network slicing using SDN and NFV: A survey of taxonomy, architectures and future challenges,’’ *Comput. Netw.*, vol. 167, Feb. 2020, Art. no. 106984.

[20] G. Zhu, J. Zan, Y. Yang, and X. Qi, ‘‘A supervised learning based QoS assurance architecture for 5G networks,’’ *IEEE Access*, vol. 7, pp. 43598–43606, 2019.

[21] G. Nencioni, R. G. Garroppo, A. J. Gonzalez, B. E. Helvik, and G. Procissi, ‘‘Orchestration and control in software-defined 5G networks: Research challenges,’’ *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–18, Aug. 2018.

[22] R. A. Addad, M. Bagaa, T. Taleb, D. L. C. Dutra, and H. Flinck, ‘‘Optimization model for cross-domain network slices in 5G networks,’’ *IEEE Trans. Mobile Comput.*, vol. 19, no. 5, pp. 1156–1169, May 2020.

[23] Y. Kim, J. Gil, and D. Kim, ‘‘A location-aware network virtualization and reconfiguration for 5G core network based on SDN and NFV,’’ *Int. J. Commun. Syst.*, vol. 34, no. 2, pp. 1–15, Jan. 2021.

[24] *Network Functions Virtualisation (NFV) Release 3; Virtualised Network Function; Specification of the Classification of Cloud Native VNF Implementations*, Standard ETSI GS NFVEVE 011, vol. V3.1.1, European Telecommunications Standards Institute, 2018-10.

[25] S. Sun, M. Kadoch, L. Gong, and B. Rong, ‘‘Integrating network function virtualization with SDR and SDN for 4G/5G networks,’’ *IEEE Netw.*, vol. 29, no. 3, pp. 54–59, May 2015.

[26] F. Z. Yousaf, M. Bredel, S. Schaller, and F. Schneider, ‘‘NFV and SDN,’’ Key technology enablers for 5G networks,’’ *IEEE J. Sel. Areas Commun.*, vol. 35, no. 11, pp. 2468–2478, Oct. 2017.

[27] Y. Zhang, *Network Function Virtualization: Concepts and Applicability in 5G Networks*. Hoboken, NJ, USA: Wiley, 2018.

[28] S. Asif, *5G Mobile Communications: Concepts and Technologies*. Boca Raton, FL, USA: CRC Press, 2018.

[29] European Telecommunications Standards Institute, *Network Functions Virtualisation (NFV) Virtualised Network Function*, Standard ETSI GS NFV-MAN 001, V1.1.1, 2014.

[30] Z. K. Khattak, M. Awais, and A. Iqbal, ‘‘Performance evaluation of OpenDaylight SDN controller,’’ in *Proc. 20th IEEE Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Dec. 2014, pp. 671–676.

[31] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, ‘‘OpenFlow: Enabling innovation in campus networks,’’ *ACM SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, pp. 69–74, Mar. 2008.

[32] R. Sherwood, G. Gibb, K.-K. Yap, G. Appenzeller, M. Casado, N. McKeown, and G. Parulkar, ‘‘Flowvisor: A network virtualization layer,’’ *OpenFlow Switch Consortium*, vol. 1, p. 132, Oct. 2009.

[33] M. N. A. Sheikh, ‘‘SDN-based approach to evaluate the best controller: Internal controller NOX and external controllers POX, ONOS, RYU,’’ *Global J. Comput. Sci. Technol.*, vol. 19, no. 1, 2019.

[34] P. Berde, M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O'Connor, P. Radoslavov, W. Snow, and G. Parulkar, ‘‘ONOS: Towards an open, distributed SDN OS,’’ in *Proc. 3rd Workshop Hot Topics Softw. Defined Netw.*, 2014, pp. 1–6.

[35] S. E. Elayoubi, S. B. Jemaa, Z. Altman, and A. Galindo-Serrano, ‘‘5G RAN slicing for verticals: Enablers and challenges,’’ *IEEE Commun. Mag.*, vol. 57, no. 1, pp. 28–34, Jan. 2019.

[36] R. Mahindra, M. A. Khojastepour, H. Zhang, and S. Rangarajan, ‘‘Radio access network sharing in cellular networks,’’ in *Proc. 21st IEEE Int. Conf. Netw. Protocols (ICNP)*, Oct. 2013, pp. 1–10.

[37] A. Gudipati, D. Perry, L. E. Li, and S. Katti, ‘‘SoftRAN: Software defined radio access network,’’ in *Proc. 2nd ACM SIGCOMM Workshop Hot Topics Softw. Defined Netw. (HotSDN)*, 2013, pp. 25–30.

[38] A. Gudipati, L. E. Li, and S. Katti, ‘‘RadioVisor: A slicing plane for radio access networks,’’ in *Proc. 3rd Workshop Hot Topics Softw. Defined Netw.*, Aug. 2014, pp. 237–238.

[39] M. Peng, C. Wang, V. Lau, and H. V. Poor, ‘‘Fronthaul-constrained cloud radio access networks: Insights and challenges,’’ *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 152–160, Apr. 2015.

[40] H. Zhang, N. Liu, X. Chu, K. Long, A.-H. Aghvami, and V. C. M. Leung, ‘‘Network slicing based 5G and future mobile networks: Mobility, resource management, and challenges,’’ *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 138–145, Aug. 2017.

[41] A. Papageorgiou, A. Fernández-Fernández, S. Siddiqui, and G. Carrozzo, ‘‘On 5G network slice modelling: Service-, resource-, or deploymentdriven?’’ *Comput. Commun.*, vol. 149, pp. 232–240, Jan. 2020.

[42] *System Architecture for the 5G System*, Standard V15.3.0, no. Release 15, The 3rd Generation Partnership Project, 2018.

[43] J. F. Monserrat, G. Mange, V. Braun, H. Tullberg, G. Zimmermann, and Ö. Bulakci, ‘‘METIS research advances towards the 5G mobile and wireless system definition,’’ *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, p. 53, Dec. 2015.

[44] J. Zhang, W. Xie, and F. Yang, ‘‘An architecture for 5G mobile network based on SDN and NFV,’’ in *Proc. 6th Int. Conf. Wireless, Mobile MultiMedia (ICWMMN)*, Nov. 2015, pp. 87–92.

[45] *Telecommunication Management; Study on Management and Orchestration of Network Slicing for Next Generation Network*, Standard Release 15.1.0, The 3rd Generation Partnership Project, 2018, pp. 1–75.

[46] *Description of Network Slicing Concept*, Alliance, NGMN, Geneva, Switzerland, 2016, pp. 1–11, vol. 1.

[47] T. Yoo, ‘‘Network slicing architecture for 5G network,’’ in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2016, pp. 1010–1014.

[48] I. Afolabi, T. Taleb, K. Samdanis, A. Ksentini, and H. Flinck, ‘‘Network slicing and softwarization: A survey on principles, enabling technologies, and solutions,’’ *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2429–2453, 3rd Quart., 2018.

[49] K. N. Sivarajan, ‘‘Network slicing and SDN: New opportunities for telecom operators,’’ *CSI Trans. ICT*, vol. 8, no. 1, pp. 15–20, Mar. 2020.

[50] A. J. Gonzalez, M. Xie, and P. Grønsund, ‘‘Network slicing architecture and dependability,’’ in *Proc. Int. Conf. Mobile, Secure, Program. Netw.* Cham, Switzerland: Springer, 2018, pp. 207–223.

[51] M. Asif Habibi, B. Han, and H. D. Schotten, ‘‘Network slicing in 5G mobile communication architecture, profit modeling, and challenges,’’ 2017, *arXiv:1707.00852*. [Online]. Available: http://arxiv.
org/abs/1707.00852

[52] Z. Kotulski, T. W. Nowak, M. Sepczuk, and M. A. Tunia, ‘‘5G networks: Types of isolation and their parameters in RAN and CN slices,’’ *Comput. Netw.*, vol. 171, Apr. 2020, Art. no. 107135.

[53] Y. Han, T. Vachuska, A. Al-Shabibi, J. Li, H. Huang, W. Snow, and J. W. Hong, ‘‘ONVisor: Towards a scalable and flexible SDN-based network virtualization platform on ONOS,’’ *Int. J. Netw. Manage.*, vol. 28, no. 2, p. 20, 2018.

[54] M. Hu, Y. Chang, Y. Sun, and H. Li, ‘‘Dynamic slicing and scheduling for wireless network virtualization in downlink LTE system,’’ in *Proc. Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Nov. 2016, pp. 153–158.

[55] M.Shurman,E.Taqieddin,O.Oudat,R.Al-Qurran,andA.A.A.Nounou, ‘‘Performance enhancement in 5G cellular networks using priorities in network slicing,’’ in *Proc. IEEE Jordan Int. Joint Conf. Electr. Eng. Inf. Technol. (JEEIT)*, Apr. 2019, pp. 822–826.

[56] U. Fattore, F. Giust, and M. Liebsch, ‘‘5GC +: An experimental proof of a programmable mobile core for 5G,’’ in *Proc. IEEE 23rd Int. Workshop Comput. Aided Modeling Design Commun. Links Netw. (CAMAD)*, Sep. 2018, pp. 1–6.

[57] Y.-I. Choi and N. Park, ‘‘Slice architecture for 5G core network,’’ in *Proc. 9th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2017, pp. 571–575.

[58] J. Tang, B. Shim, and T. Q. S. Quek, ‘‘Service multiplexing and revenue maximization in sliced C-RAN incorporated with URLLC and multicast eMBB,’’ *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 881–895, Apr. 2019.

[59] Z. Jian, W. Muqing, M. Ruiqiang, and W. Xiusheng, ‘‘Dynamic resource sharing scheme across network slicing for multi-tenant C-RANs,’’ in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC Workshops)*, Aug. 2018, pp. 172–177.

[60] Y. Kim, S. Kim, and H. Lim, ‘‘Reinforcement learning based resource management for network slicing,’’ *Appl. Sci.*, vol. 9, no. 11, p. 2361, Jun. 2019.

[61] L. Tang, Q. Tan, Y. Shi, C. Wang, and Q. Chen, ‘‘Adaptive virtual resource allocation in 5G network slicing using constrained Markov decision process,’’ *IEEE Access*, vol. 6, pp. 61184–61195, 2018.

[62] A. Thantharate, R. Paropkari, V. Walunj, and C. Beard, ‘‘DeepSlice: A deep learning approach towards an efficient and reliable network slicing in 5G networks,’’ in *Proc. IEEE 10th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Oct. 2019, pp. 0762–0767.

[63] H. M. Soliman and A. Leon-Garcia, ‘‘QoS-aware frequency-space network slicing and admission control for virtual wireless networks,’’ in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[64] A. Kammoun, N. Tabbane, G. Diaz, and N. Achir, ‘‘Admission control algorithm for network slicing management in SDN-NFV environment,’’ in *Proc. 6th Int. Conf. Multimedia Comput. Syst. (ICMCS)*, May 2018, pp. 1–6.

[65] B. Han, J. Lianghai, and H. D. Schotten, ‘‘Slice as an evolutionary service: Genetic optimization for inter-slice resource management in 5G networks,’’ *IEEE Access*, vol. 6, pp. 33137–33147, 2018.

[66] B. Han, A. Dedomenico, G. Dandachi, A. Drosou, D. Tzovaras, R. Querio, F. Moggio, O. Bulakci, and H. D. Schotten, ‘‘Admission and congestion control for 5G network slicing,’’ in *Proc. IEEE Conf. Standards for Commun. Netw. (CSCN)*, Oct. 2018, pp. 1–6.

[67] M. Jiang, M. Condoluci, and T. Mahmoodi, ‘‘Network slicing management & prioritization in 5G mobile systems,’’ in *Proc. 22th Eur. Wireless Conf.*, May 2016, pp. 1–6.

[68] C. Natalino, M. R. Raza, A. Rostami, P. Öhlen, L. Wosinska, and P. Monti, ‘‘Machine learning aided orchestration in multi-tenant networks,’’ in *Proc. IEEE Photon. Soc. Summer Topical Meeting Ser. (SUM)*, Jul. 2018, pp. 125–126.

[69] R. Li, Z. Zhao, Q. Sun, C.-L. I, C. Yang, X. Chen, M. Zhao, and H. Zhang, ‘‘Deep reinforcement learning for resource management in network slicing,’’ *IEEE Access*, vol. 6, pp. 74429–74441, 2018.

[70] G. Sun, K. Xiong, G. O. Boateng, D. Ayepah-Mensah, G. Liu, and W. Jiang, ‘‘Autonomous resource provisioning and resource customization for mixed traffics in virtualized radio access network,’’ *IEEE Syst. J.*, vol. 13, no. 3, pp. 2454–2465, Sep. 2019.

[71] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, and X. Costa-Pérez, ‘‘A machine learning approach to 5G infrastructure market optimization,’’ *IEEE Trans. Mobile Comput.*, vol. 19, no. 3, pp. 498–512, Mar. 2020.

[72] B. Yi, X. Wang, and M. Huang, ‘‘Optimised approach for VNF embedding in NFV,’’ *IET Commun.*, vol. 12, no. 20, pp. 2630–2638, Dec. 2018. [73] M. Jiang, M. Condoluci, and T. Mahmoodi, ‘‘Network slicing in 5G: An auction-based model,’’ in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[74] J. Hu, Z. Zheng, B. Di, and L. Song, ‘‘Tri-level Stackelberg game for resourceallocationinradioaccessnetworkslicing,’’in*Proc.IEEEGlobal Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.
[75] M. R. Raza, C. Natalino, P. Öhlen, L. Wosinska, and P. Monti, ‘‘A slice admission policy based on reinforcement learning for a 5G flexible RAN,’’ in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2018, pp. 1–3.
[76] S. Parsaeefard, V. Jumba, M. Derakhshani, and T. Le-Ngoc, ‘‘Joint resource provisioning and admission control in wireless virtualized networks,’’ in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 2020–2025.
[77] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, K. Samdanis, and X. Costa-Perez, ‘‘Optimising 5G infrastructure markets: The business of network slicing,’’ in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, May 2017, pp. 1–9.
[78] A. Oliveira and T. Vazão, ‘‘Adapting priority schemes to achieve network slice isolation,’’ in *Proc. 35th Annu. ACM Symp. Appl. Comput.*, Mar. 2020, pp. 1164–1171.
[79] L. Ma, X. Wen, L. Wang, Z. Lu, and R. Knopp, ‘‘An SDN/NFV based framework for management and deployment of service based 5G core network,’’ *China Commun.*, vol. 15, no. 10, pp. 86–98, Oct. 2018.
[80] D. Sattar and A. Matrawy, ‘‘Optimal slice allocation in 5G core networks,’’ *IEEE Netw. Lett.*, vol. 1, no. 2, pp. 48–51, Jun. 2019.
[81] D. Marabissi and R. Fantacci, ‘‘Highly flexible RAN slicing approach to manage isolation, priority, efficiency,’’ *IEEE Access*, vol. 7, pp. 97130–97142, 2019.
[82] E. Hossain and M. Hasan, ‘‘5G cellular: Key enabling technologies and research challenges,’’ *IEEE Instrum. Meas. Mag.*, vol. 18, no. 3, pp. 11–21, Jun. 2015.
[83] 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, Jul. 2016.
[84] B. R. Chandavarkar and G. R. M. Reddy, ‘‘Survey paper: Mobility management in heterogeneous wireless networks,’’ *Procedia Eng.*, vol. 30, pp. 113–123, Jan. 2012.
[85] U. Mian, A. Mudassir, and S. Akthar, ‘‘A survey on handover issues and management techniques for LTE heterogeneous networks,’’ *Procedia Eng.*, vol. 7, no. 4, pp. 99–106, 2016.
[86] A. Jain, E. López-Aguilera, and I. Demirkol, ‘‘Mobility management as a service for 5G networks,’’ 2017, *arXiv:1705.09101*. [Online]. Available: http://arxiv.org/abs/1705.09101
[87] W. Hucheng, C. Shanzhi, A. Ming, and S. Yan, ‘‘Mobility driven network slicing: An enabler of on demand mobility management for 5G,’’ *J. China Universities Posts Telecommun.*, vol. 24, no. 4, pp. 16–26, Aug. 2017.
[88] A. Gharsallah, F. Zarai, and M. Neji, ‘‘SDN/NFV-based handover management approach for ultradense 5G mobile networks,’’ *Int. J. Commun. Syst.*, vol. 32, no. 17, p. e3831, Nov. 2019.
[89] R. Ravindran, A. Chakraborti, S. O. Amin, A. Azgin, and G. Wang, ‘‘5G-ICN: Delivering ICN services over 5G using network slicing,’’ *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 101–107, May 2017.
[90] M. Amadeo, C. Campolo, and A. Molinaro, ‘‘Empowering 5G network softwarization through information centric networking,’’ *Internet Technol. Lett.*, vol. 1, no. 2, p. e30, Mar. 2018.
[91] H. Song, X. Fang, and L. Yan, ‘‘Handover scheme for 5G C/U plane split heterogeneous network in high-speed railway,’’ *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4633–4646, Nov. 2014.
[92] R. A. Addad, T. Taleb, H. Flinck, M. Bagaa, and D. Dutra, ‘‘Network slice mobility in next generation mobile systems: Challenges and potential solutions,’’ *IEEE Netw.*, vol. 34, no. 1, pp. 84–93, Jan. 2020.
[93] J. Heinonen, P. Korja, T. Partti, H. Flinck, and P. Pöyhönen, ‘‘Mobility management enhancements for 5G low latency services,’’ in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, May 2016, pp. 68–73.
[94] T.-T. Nguyen, C. Bonnet, and J. Harri, ‘‘SDN-based distributed mobility management for 5G networks,’’ in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2016, pp. 1–7.
[95] *Study on Architecture for Next Generation System (Network Slicing)*, Standard V0.5.0, Release 10, The 3rd Generation Partnership Project, 2011, pp. 1–522.
[96] *Management and Orchestration; Provisioning, Concepts, Use Cases and Requirements*, Standard 15.4.0, Release 15, The 3rd Generation Partnership Project, 2018, pp. 1–32.
[97] *5G;Management and Orchestration; Provisioning*, Standard 15.5.0, no. Release 15, The 3rd Generation Partnership Project, 2020, pp. 1–70.
[98] M. M. Sajjad, C. J. Bernardos, D. Jayalath, and Y.-C. Tian, ‘‘Inter-slice mobility management in 5G: Motivations, standard principles, challenges and research directions,’’ 2020, *arXiv:2003.11343*. [Online]. Available: http://arxiv.org/abs/2003.11343
[99] H. Khalili, A. Papageorgiou, S. Siddiqui, C. Colman-Meixner, G. Carrozzo, R. Nejabati, and D. Simeonidou, ‘‘Network slicingaware NFV orchestration for 5G service platforms,’’ in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2019, pp. 25–30.
[100] H. Wang, S. Chen, H. Xu, M. Ai, and Y. Shi, ‘‘SoftNet: A software defined decentralized mobile network architecture toward 5G,’’ *IEEE Netw.*, vol. 29, no. 2, pp. 16–22, Mar. 2015.
[101] L. Baldini, Q. Wang, J. A. Calero, M. B. Weiss, A. Gavras, G. Bernini,
P. G. Giardina, C. Angelo, X. Vasilakos, C.-Y. Chang, N. Nikaein, S. Spadaro, A. Pages, F. Agraz, G. Agapiou, T. Truong, K. Koutsopoulos, J. Cabaca, and R. Figueiredo, ‘‘SliceNet control plane for 5G network slicing in evolving future networks,’’ in *Proc. IEEE Conf. Netw. Softwarization (NetSoft)*, Jun. 2019, pp. 450–457.
[102] E. Coronado, S. N. Khan, and R. Riggio, ‘‘5G-EmPOWER: A softwaredefined networking platform for 5G radio access networks,’’ *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 2, pp. 715–728, Jun. 2019.
[103] L. E. Li, Z. M. Mao, and J. Rexford, ‘‘Toward software-defined cellular networks,’’ in *Proc. Eur. Workshop Softw. Defined Netw.*, Oct. 2012, pp. 7–12.
[104] M. Moradi, L. E. Li, and Z. M. Mao, ‘‘SoftMoW: A dynamic and scalable software defined architecture for cellular WANs,’’ in *Proc. 3rd Workshop Hot Topics Softw. Defined Netw.*, Aug. 2014, pp. 201–202.
[105] L. I. B. López, J. M. Vidal, and L. J. García Villalba, ‘‘Orchestration of use-case driven analytics in 5G scenarios,’’ *J. Ambient Intell. Humanized Comput.*, vol. 9, no. 4, pp. 1097–1117, Aug. 2018.
[106] L. Goratti, C. E. Costa, J. Perez-Romano, O. Sallent, C. Ruiz, A. Betzler, P. S. Khodashenas, S. Vahid, K. M. Nasr, B. Abubakar, and A. Whitehead, ‘‘Network architecture and essential features for 5G: The SESAME project approach,’’ in *Proc. IFIP Int. Conf. Artif. Intell. Appl. Innov.* Springer, 2016, pp. 676–685.
[107] C.Mannweiler,M.Breitbach,H.Droste,I.L.Pavon,I.Ucar,P.Schneider, M. Doll, and J. R. Sanchez, ‘‘5G NORMA: System architecture for programmable & multi-tenant 5G mobile networks,’’ in *Proc. Eur. Conf.*
Netw. Commun. (EuCNC), Jun. 2017, pp. 1–6.
[108] M. Gramaglia, I. Digon, V. Friderikos, D. von Hugo, C. Mannweiler, M. A. Puente, K. Samdanis, and B. Sayadi, ‘‘Flexible connectivity and QoE/QoS management for 5G networks: The 5G NORMA view,’’ in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, May 2016, pp. 373–379.
[109] C. J. Bernardos, B. P. Gerö, M. Di Girolamo, A. Kern, B. Martini, and I. Vaishnavi, ‘‘5GEx: Realising a europe-wide multi-domain framework for software-defined infrastructures,’’ *Trans. Emerg. Telecommun. Technol.*, vol. 27, no. 9, pp. 1271–1280, Sep. 2016, doi: 10.1002/ett.3085.
[110] C. Bernardos, O. Dugeon, A. Galis, D. Morris, C. Simon, and R. Szabó, ‘‘5G exchange (5Gex)-multi-domain orchestration for software defined infrastructures,’’ *Focus*, vol. 4, no. 5, p. 2, 2015.

[111] A. Sgambelluri, F. Tusa, M. Gharbaoui, E. Maini, L. Toka, J. M. Perez, F. Paolucci, B. Martini, W. Y. Poe, J. Melian Hernandes, A. Muhammed, A. Ramos, O. G. de Dios, B. Sonkoly, P. Monti, I. Vaishnavi, C. J. Bernardos, and R. Szabo, ''Orchestration of network services acrossmultiple operators: The 5G exchange prototype,'' in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2017, pp. 1–5.

[112] L. Xu, H. Assem, I. G. Ben Yahia, T. S. Buda, A. Martin, D. Gallico, M. Biancani, A. Pastor, P. A. Aranda, M. Smirnov, D. Raz, O. Uryupina, A. Mozo, B. Ordozgoiti, M.-I. Corici, P. O'Sullivan, and R. Mullins, ''CogNet: A network management architecture featuring cognitive capabilities,'' in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2016, pp. 325–329.

[113] A. Kostopoulos, G. Agapiou, F.-C. Kuo, K. Pentikousis, A. Cipriano, D. Panaitopol, D. Marandin, K. Kowalik, K. Alexandris, C.-Y. Chang,N. Nikaein, M. Goldhamer, A. Kliks, R. Steinert, A. Mämmelä, and T. Chen, ''Scenarios for 5G networks: The COHERENT approach,'' in *Proc. 23rd Int. Conf. Telecommun. (ICT)*, May 2016, pp. 1–6.

[114] Q.Wang,J.Alcaraz-Calero,M.B.Weiss,A.Gavras,P.M.Neves,R.Cale, G. Bernini, G. Carrozzo, N. Ciulli, G. Celozzi, and A. Ciriaco, ''SliceNet: End-to-end cognitive network slicing and slice management framework in virtualised multi-domain, multi-tenant 5G networks,'' in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, Jun. 2018, pp. 1–5.

[115] F. Silva, P. Rosa, H. Hrasnica, and A. Gravas, ''5GINFIRE: Enabling an NFV based experimentation of vertical industries in the 5G context,'' in *Proc. Anais do X Workshop de Pesquisa Experim. da Internet do Futuro (WPEIF)*, May 2019, pp. 64–69.

[116] Q. Wang *et al.*, ''Enable advanced QoS-aware network slicing in 5G networks for slice-based media use cases,'' *IEEE Trans. Broadcast.*, vol. 65, no. 2, pp. 444–453, Jun. 2019.

[117] J. Song, T. Yoo, and P. J. Song, ''Mobility level management for 5G network,'' in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2016, pp. 940–943.

[118] R. A. Addad, T. Taleb, M. Bagaa, D. L. C. Dutra, and H. Flinck, ''Towards modeling cross-domain network slices for 5G,'' in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–7.

[119] B. Bordel, R. Alcarria, D. Sánchez-de Rivera, and A. Sánchez, ''An interslice management solution for future virtualization-based 5G systems,'' in *Proc. Int. Conf. Adv. Inf. Netw. Appl.* Abu Dhabi, United Arab Emirates: Springer, 2019, pp. 1059–1070.