Software-Defined Control of the Virtualized Mobile Packet Core

Malla Reddy Sama, Luis M. Contreras, John Kaippallimalil, Ippei Akiyoshi, Haiyang Qian, and Hui Ni

Malla Reddy Sama is with Orange Labs.

Luis M. Contreras is with Telefónica I+D.

John Kaippallimalil and Hui Ni are with Huawei.

Ippei Akiyoshi is with NEC.

Haiyang Qian is with China Mobile.

ABSTRACT

Mobile packet core networks are undergoing major changes to meet the requirements of the future data tsunami, enhance network flexibility, and reduce both CAPEX and OPEX. In this regard, SDN and NFV technologies have gained great momentum among the telcos, with the promise of interoperability, programmability, and on-demand dynamic provisioning. In this article, we present work being developed in the Mobile Packet Core project within the ONF Wireless & Mobile Working Group regarding SDN architecture for the Mobile Packet Core. In addition, we propose an SDN-based MPC in the NFV context in order to facilitate dynamic provisioning of MPC network functions. Finally, a potential control architecture considering both SDN and NFV is proposed.

INTRODUCTION

Over the last few years, cloud and virtualization technologies have gained great momentum among mobile operators and network equipment vendors. This is due to the flexibility these technologies introduce in the network, offering significant economic potential to reduce both capital expenditure (CAPEX) and operational expenditure (OPEX) costs. In addition, operators are also given the privilege to program the network functions (e.g., gateways, routers, load balancer) independent of the proprietary hardware substrate. This has not been the case over recent decades. For instance, operators could not change any software function or implement their own software function on vendor-specific hardware. Instead, they had to wait for vendors' developments to apply changes.

Innovations in the mobile field such as powerful new terminals (e.g., smartphone and tablets) and the proliferation of data-hungry mobile applications (mobile health, mobile education, context-awareness applications, cloud communication etc.) have dramatically increased data traffic usage because of the demand for rich and sophisticated services. It is estimated that mobile traffic will grow at a compound annual growth rate (CAGR) of 61 percent from 2013 to 2018 [1]. To sustain this data tsunami, operators are required to deploy a wide variety of hardware appliances such as gateways, routers, servers, intrusion detection systems and firewalls. These appliances run on dedicated proprietary hardware placed in different locations in the network. In the traditional mode of operation, telcos have populated the network with these monolithic physical devices, resulting in difficulties in rolling out new devices and allocating the required resources in a cost-effective manner. Under these circumstances, operators have to repeatedly invest to cope with severe data usage increments, resulting in increasing the network CAPEX in a context where the average revenue per user (ARPU) is declining.

In addition, the mobile network entities in the Evolved Packet Core (EPC), that is, the mobility management entity (MME), serving gateway (SGW), packet data network gateway (PGW), and home subscriber server (HSS), follow the same model, which is based on customized hardware that requires static deployment, provisioning, and configuring. As a consequence, the network architecture does not inherently define any flexibility, dynamicity, or on-demand features. Those network entities are tightly coupled in two dimensions: (i) software and hardware; and (ii) control and user planes. For example, to provision network capacity, operators typically dimension both the control and user plane capacity based on the load foreseen in peak hours. User plane processing is I/O bound and requires high capacity, while control plane processing is CPU bound. Significantly, static provisioning of both planes is not an optimal solution, causing waste of truly expensive resources.

On these ground, network function virtualization (NFV) and software defined networking (SDN) emerge as the latest incarnation of technological promises for reaching the necessary cost efficiency. SDN decouples the control and user planes, and logically centralizes the network intelligence (i.e., control plane), while the underlying network infrastructure (i.e., user plane) is abstracted for external applications requesting services through that control plane. This mechanism facilitates the programmability of the network resources by automatically and dynamically allowing the control of underlay network switches. The OpenFlow protocol is a key component of the SDN concept to model flow abstractions. In this context, last year, the Open Networking Foundation (ONF) chartered the Wireless &

0163-6804/15/$25.00 © 2015 IEEE

The LTE/EPC architecture has been widely accepted. However, the architecture was not designed with elasticity in mind. Functionality is deployed on specific network entities, which execute specific functions. In addition, the current network entities are too prone to vendor locking, too complicated to manage, and too hard to change in behavior.

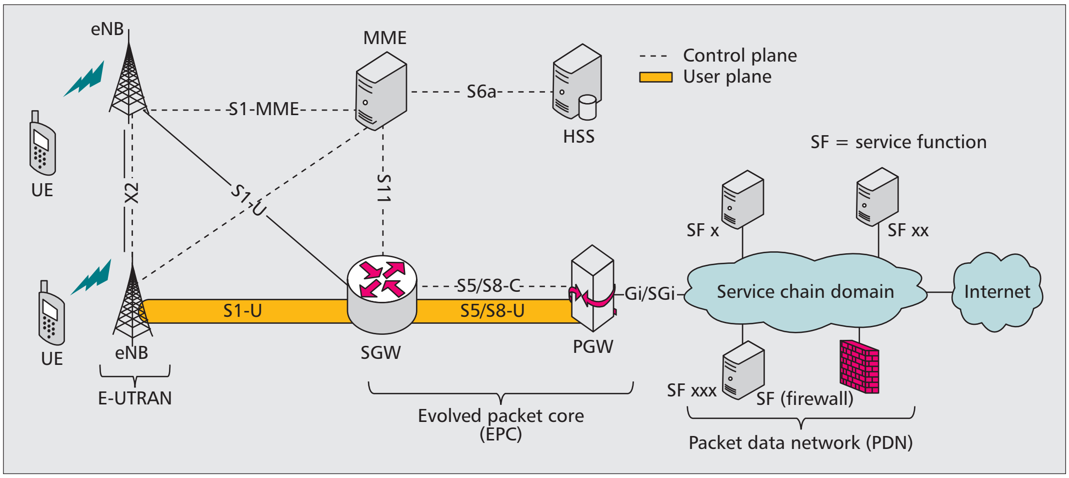

Figure 1. A mobile operator network.

Mobile Working Group (WMWG)[1] to foster the adoption of OpenFlow-based SDN technology in mobile and wireless networks. This group studies and proposes simplified reference mobile architectures for transport by leveraging OpenFlow-based SDN. In addition, this group aims to simplify the interaction between wireless networks and fixed networks (i.e., physical layer to upper layers interaction and vice versa). Moreover, it is also responsible for proposing extensions to the OpenFlow protocol specification, such as extensions required to transfer the general packet radio service (GPRS) Tunneling Protocol (GTP) tunnel-ids to the switches, and the parameters for interaction between the control and user planes.

On the other hand, the NFV initiative was created by a group of tier-1 telcos within the European Telecommunications Standards Institute's (ETSI's) Industry Specification Group and released a white paper in October 2012 [3]. The scope of NFV is to propose consistent and common architecture to decouple mobile network function (e.g., firewalling, NAT, gateways) from proprietary hardware, making it possible to run such functions in general-purpose commodity servers, switches, and storage units, which could be deployed in an operator's data center. The main expected benefits with NFV are lower OPEX, greater flexibility (i.e., easily scaling network resources up and down), and easier network management.

This article addresses the evolution of the existing mobile packet core (MPC) architecture toward a new architecture where network functions are deployed in a virtualized manner with the separation of the control and user planes using a programmable infrastructure.

SOFTWARE DEFINED MOBILE PACKET CORE

LTE/EPC ARCHITECTURE

The Third Generation Partnership Project (3GPP) Long Term Evolution/Evolved Packet Core (LTE/EPC) architecture has been designed to provide seamless IP connectivity between user equipment (UE) and external packet data networks (PDNs) as shown in Fig. 1. The EPC has a flat all-IP architecture [4] composed of a number of entities. The MME acts as the manager of network connectivity, and also for UE authentication and authorization, UE session setup, and intra-3GPP mobility management. The SGW routes data packets through the access network and is the local mobility anchor point for inter-eNB handover. The PGW is responsible for user plane quality of service (QoS) management and is the anchor point for external PDN Networks. Both GTP and proxy mobile IP (PMIP) are the main communication protocols within the LTE/EPC architecture.

The LTE/EPC architecture has been widely accepted. However, the architecture was not designed with elasticity in mind. Functionality is deployed on specific network entities, which execute specific functions. In addition, the current network entities are too prone to vendor locking, too complicated to manage, and too hard to change in behavior. With the introduction of SDN and NFV, greater flexibility can be achieved by network operators. In the first case, by usage of SDN capabilities, dynamic control of traffic flows can be performed, redirecting the traffic to gateways according to, say, workloads. In the second case, the introduction of NFV permits the separation of service functionalities from the capacity constraints limited boxes could suffer by allowing dynamic instantiation in commodity and powerful servers.

THE SDN APPROACH TO THE LTE MPC

The Mobile Packet Core (MPC) project within the ONF is identifying the OpenFlow (OF) protocol extensions needed to support handling of packet data network (PDN) connections. Such extensions should be able to provision and manage the establishment, modification, and deletion of PDN connections while providing support for QoS, accounting, and online charging, which translates into programmable rate limiting, metering, and rating functionality configuration in the supporting switch. Forwarding abstractions for PDN connections in an SGW, a PGW, and a combined gateway (SGW and PGW combination) are being defined. These abstractions model the PDN connection segments using

---

[1] ONF Wireless and Mobile Group, https://www.opennetworking.org/technical-communities/areas/specification

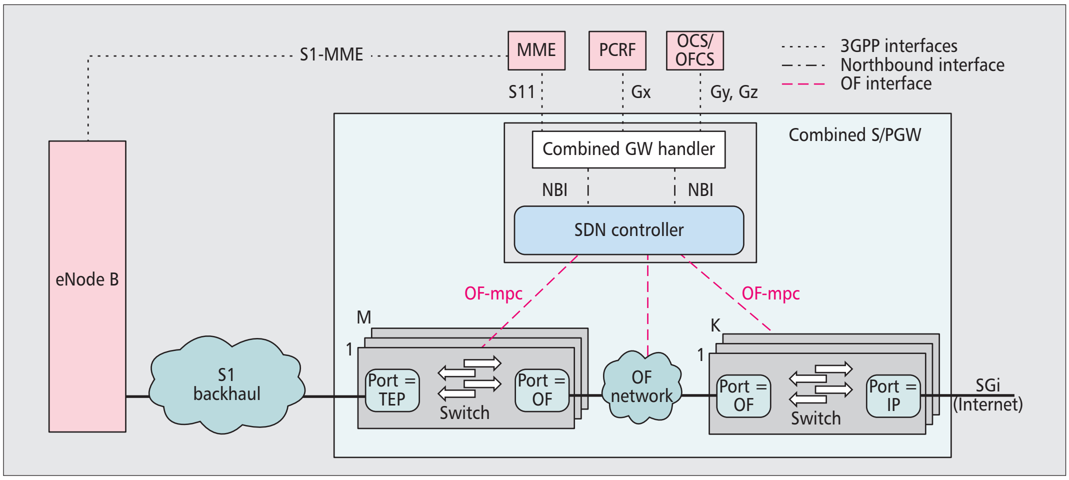

Figure 2. Combined SGW and PGW in the MPC architecture.

3GPP/LTE gateways need to support a number of capabilities, including IP address allocation and management, gating and rate control, event reporting and bearer binding, and support for handovers with end-markers and buffering in the SGW.

ports, tables, and control signaling in OpenFlow. On the other hand, no changes are necessary to standard 3GPP defined interfaces.

The SDN controller in this case may also be implemented as a plug-in in an NFV realization of 3GPP control functions. A description of an OF based combined SGW and PGW gateway and its relation to 3GPP interfaces is shown Fig. 2. Note that the dotted lines refer to control and signaling interface while solid lines represent data interface.

Figure 2 shows a combined SGW and PGW node with an SDN controller and OpenFlow channel (OF-mpc) extensions to control the switch. The necessary OF-mpc extensions are currently being defined by the ONF W&MWG with final specification expected during 2015. The combined gateways are common in current LTE deployments. However, there are scenarios where even in that case the S5 interface has to be implemented (not shown in Fig. 2 for simplicity). For instance, a roaming mobile terminal may attach via the SGW functionality of a combined gateway of the roaming network, and to a PGW in the home network via the S5 interface.

The combined gateway handler is the control software of the combined gateway and interfaces to the other 3GPP functional entities. In an NFV-based deployment with virtual machines (VMs), both the combined gateway handler and SDN controller may be implemented as a VM. The northbound interface (NBI) in Fig. 2 can be realized through controller plugins. Since the control plane scaling is critical in mobile infrastructure, the SDN controller itself may be implemented as a cluster.

Forwarding abstractions for the PDN connection consist of ports that serve as network interface points to connect the S1 interface (toward the eNB) and the SGi (connecting to external networks), as well as flow and group tables for OF switch components. Packets received on the ingress port are processed through a pipeline of flow tables that route them to the egress port. The SDN controller can manage multiple switches. In addition, these switches are able to encapsulate and decapsulate the GTP and generic routing encapsulation (GRE) packets based on fine-grained rules installed by the SDN controller. All the necessary extensions for such packet manipulation are currently being defined in ONF W&MWG.

3GPP/LTE gateways need to support a number of capabilities, including IP address allocation and management, gating and rate control, event reporting and bearer binding, and support for handovers with end-markers and buffering in the SGW. Support for these capabilities will be exposed by the OF-mpc channel extensions including metering, rating, and switch events.

OpenFlow Extension for Mobile Packet Core — In the MPC architecture OF switches need to handle GTP-U for actions like GTP-U encapsulation or decapsulation. In order to do that, the OpenFlow protocol needs to be extended. Possible extensions are:

- *Match field extension*: To distinguish evolved packet system (EPS) bearers and the corresponding GTP-U signal messages, the OF switch shall be able to match "message type" and "TEID" of a GTP-U header.
- *Action type extension*: The OF switch shall encapsulate and decapsulate GTP-U tunnels. The former action is used to encapsulate the user packet into an appropriate GTP-U tunnel, while the latter is used to decapsulate a GTP-U tunnel header from a GTP-U packet.

NFV- AND SDN-BASED MPC

The NFV framework in [3] consists of the NFV infrastructure (NFVI), which logically partitions the resources from underlying physical hardware resources. NFV management and orchestration (MANO) is the life cycle management of network service and the virtual network function (VNF) instance. The VNF is purely a software application (i.e., network functions like MME and PCRF) deployed in NFVI, and VNF is mapped into a VM running on top of NFVI. The MANO orchestrates other specific managers such as the virtual infrastructure manager (VIM) and the VNF Manager (VNFM). The VNFM is in charge of interacting with the VNFs, whereas the VIM is in charge of managing the NFVI, which includes computing, storage, and

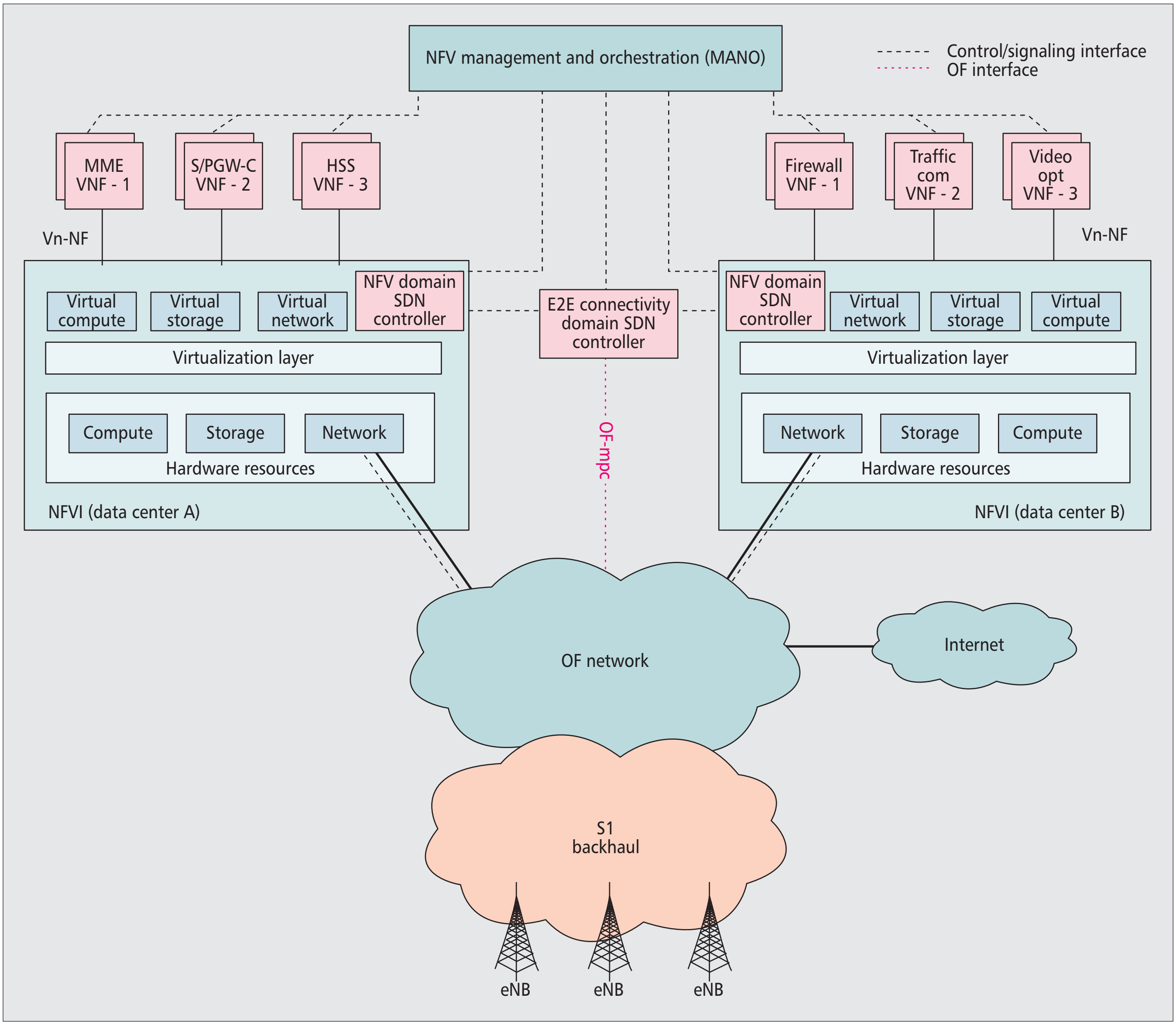

Figure 3. NFV- and SDN-based MPC architecture.

networking. The MANO has a total visibility of all VNFs running inside the NFVI. In addition, it is in-charge of the operation and configuration of VNFs, for example, through the operations support system (OSS)/base station subsystem (BSS).

One of the use cases outlined in the NFV group specifications is the mobile core [5]. A major motivation for enabling virtualization in this part of the network is the different pace of the scaling needs for each of the involved core entities, like the MME, SGW, and PGW. Those needs can be translated into higher or lower necessity for function distribution, or even higher or lower requirements in terms of resource usage (some entities limited by user plane traffic vs. others limited by control plane signaling).

NFV AND SDN COMBINED ARCHITECTURE

To enable flexibility, programmability, and optimization of network resources, we propose a logically centralized EPC control plane (i.e., MME, combined gateway handler, PCRF, etc.) running as a group of VNFs hosted in operators' data centers, as illustrated in Fig. 3.

In order to optimize resources in the user plane, the GTP tunnels are only used between the eNode B (eNB) and the first ingress switch of an OpenFlow-enabled network (i.e., S1 backhaul in Fig. 3). Note that substitutes for GTP tunnels are possible as long as they can manage UEs' localities and identities. Beyond the first ingress switch, the traffic is forwarded by an OpenFlow-enabled network infrastructure considering both the user equipment IP and destination addresses, and the QoS characteristics defined by the control entities through the NBI. In addition, if required, layer 2 (VLAN) tags, for instance, can be used to differentiate the transport bearers to differentiate QoS. Moreover, tags can also be used for overlapping UE addresses going to different access point names (APNs).

The "NFV domain SDN controller" is in charge of providing connectivity to the computing and storage resources locally to the data center. Those resources will attach as endpoints to transport capabilities in the OF-enabled network.

On the other hand, the "E2E connectivity domain SDN controller" is the main entity in our architecture as it manages the forwarding plane. It can consist of a network operating system (NetOS) running a collection of application modules, such as topology discovery, path computation, resource management, and load balancing. It also installs fine-grained packet handling rules in the switch. For instance, forwarding rules for UE flows can be dependent on the UE location, type of service, and network conditions. Note that the centralization of functions can introduce scalability issues (e.g., bottlenecks). For this scalability issue, there can be two possibilities:

- The controller is executed in a clustered environment (i.e., deploying active and backup controllers) using shared databases.
- The control functions (i.e., flow rules installation) are decentralized between the controller and switch.

In the latter case, the switch itself acts as a local controller and installs the rules. If the switch is unable to install the rules, the switch communicates with the centralized controller and gets the rules to improve scalability. In any case, the scalability issue in the SDN network is an extensive field of research, while in most cases that issue can be addressed without losing the benefits of SDN [6].

EPC CONTROL ENTITIES IN AN SDN CONTROLLED NFV ENVIRONMENT

The EPC control plane entities can be instantiated into high-volume servers, while the data plane forwarding tasks can be performed by OpenFlow-enabled switches as shown in Fig. 3.

All these control entities are mapped as software applications in VNF instances running on top of NFVIs in a data center. For instance, the VNFs (e.g., MME, PGW) can run on top of kernel-based VM (KVM) hypervisors in a server. This KVM isolates the server hardware resources between the VNFs and interconnects the VNFs using virtual network ports and bridges inside the server. On the other hand, standard interfaces are maintained toward the other EPC control plane applications such as the MME, HSS, and PCRF, as shown in Fig. 2. In addition, the end-to-end (E2E) data plane is connected by means of SDN controllers by using communication mechanisms such as the REST application programming interface (API).

For instance, the combined control function of SGW and PGW (S/PGW-C in Fig. 3 or combined GW handler in Fig. 2) allocates the tunnel ids for UE flows, acting as a mobility anchor for inter-eNB communication. Furthermore, the S/PGW-C connects to the MME using GTP-C (i.e., the S11 interface, which is GTP-C-based).

The S/PGW-C function will use a northbound interface (e.g., REST API) with the "E2E connectivity domain SDN controller," which translates northbound messages into OpenFlow messages. For instance, the S/PGW-C sends the UE user plane bearer GTP TEIDs to the controller, which translates these IDs into OpenFlow messages (e.g., Packet Out) instructing the switches, in a similar manner as proposed by [14]. To do this, first the OpenFlow protocol needs an extension to transfer the GTP TEIDs; and second, the S/PGW-C function needs to send UE information (e.g., mobility updates) to the controller. The northbound interface needs to be extended to exchange the UE information (GTP TEIDs, eNB address, etc.).

The MANO is in charge of managing the hardware and virtual network resources within the data center as well as configuring the VNFs and their interfaces, defined in Fig. 3. For instance, if a new virtual MME is switched on in a data center, the MANO will notify and configure this new MME with the corresponding association of other virtual functions such as the PGW-C, HSS, and other MMEs. In addition, the MANO will interact with the referred SDN controllers for setting up the required connectivity, both internally and E2E.

On the other hand, the new MMEs are also required to configure and establish S1-AP interfaces with the eNBs, which are physical devices geographically distributed across the network. For instance, when an MME is scaled-in into the network, the MANO gives the pool of eNBs of interest and the information details to be configured. Based on this information, the new MME will start an initial handshake procedure and establish the S1-AP interface as specified in the 3GPP specification [7].

When an MME is scaled-in into the network, the MANO gives the pool of eNBs of interest and the information details to be configured. Based on this information, the new MME will start an initial handshake procedure and establish the S1-AP interface as specified in the 3GPP specification.

REALIZATION OF AN OPERATIONAL CONTROL ARCHITECTURE

The combination of virtual mobile control functions together with an underlying programmable infrastructure requires clear differentiation of actions to be implemented on the network. Some of those actions are intended for deploying and maintaining the network functionality, while others are specific for building pure connectivity. It seems difficult to develop programmatic applications capable of covering such a broad scope of actions in a unique controller.

LAYERED SDN CONTROL ARCHITECTURE

The centralized MPC control functions have to interact with different types of resources to accomplish service delivery. Obviously, control over the underlying transport resources is needed to forward the user plane traffic properly (e.g., the connectivity to the Internet for mobile data access), which relies on the bidirectional transport of user packets between the PGW and SGW. This will require the population of forwarding rules on OpenFlow-enabled switches in a coordinated way and the composition of an E2E connectivity path across the SDN domain. These are resources related to the transport function.

Notwithstanding, the packet core entities need to access databases and storage resources for a variety of functions, including security, logging, and registers, or even processing capacity (in terms of computing CPUs) for running their

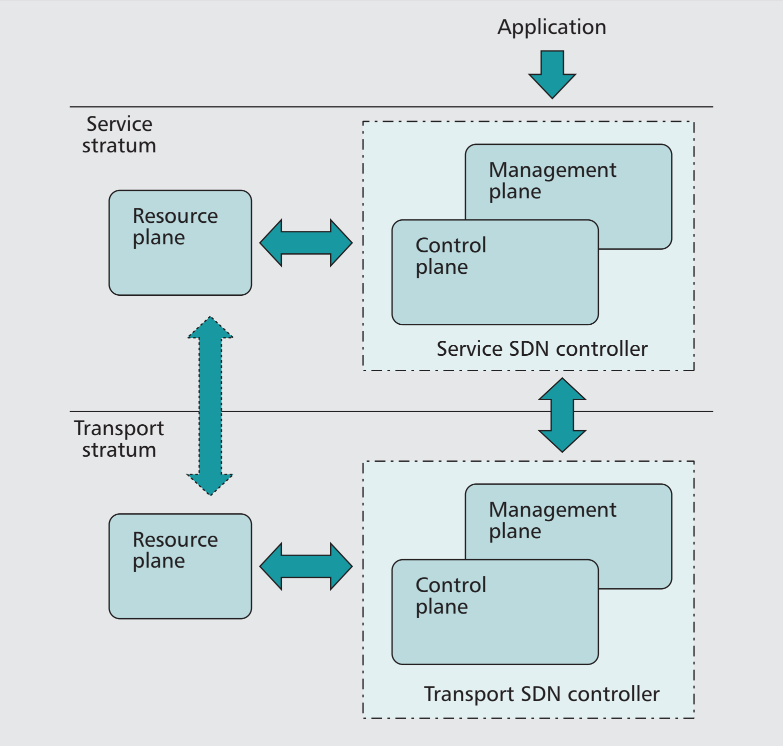

Figure 4. Layered control architecture [8].

own logic and finite state machine. These resources are related to the service itself.

There is then a clear differentiation between the kind of resources (for transport and service) involved in the service provision. However, all of those resources have to be orchestrated in a consistent way to perform the expected service. An approach is then needed that considers both kinds of resources at the same time.

Existing proposals for SDN-based control do not provide a clear separation between service and transport control.

The cooperating layered architecture for SDN (CLAS) [8] proposes a clear separation between both service and transport control, with both layers cooperating in a tight way for the final service provision. It is based on the functional separation in the NGN architecture defined by International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) Recommendation Y.2011.

Basically, in CLAS both the control functions associated with transport and those related to services are differentiated by allocating them in separate control layers, which cooperate to build the final service. Figure 4 presents a schematic view of the layered architecture.

Two strata are defined, the transport stratum and the service stratum. The former includes the functions devoted to handling the data forwarding by building (and managing over time) connectivity between communication endpoints. Apart from dealing with distinct adapters for interacting with the infrastructure (OpenFlow, NETCONF, etc.), this may involve some transport abstraction capabilities.

The latter comprises the functions related to the provision of services as well as the control capabilities exposed to external applications that run on top for providing value added services.

Each stratum consists of three different planes. One of those planes refers to the resources targeted for the given stratum. For example, the service stratum could require storage capacity, while the transport stratum could require disjoint paths for ensuring network resiliency. In some cases the service resources can be connected to the transport ones (e.g., as the terminating points of a transport function), while in other cases they can be totally decoupled.

The control plane consists of the control functions of each stratum and directly implementing control actions (e.g., configuration) on the corresponding resources. Both control planes have to be tightly coordinated in order to provide consistent service.

The management plane is in charge of performing management procedures, including the collection of monitoring data, alarms, and so on for elements that are part of either the resource or control plane.

Both the control and management planes in CLAS can be assimilated to the homonym planes (plus the associated abstraction layer) in [2]. The separation of service and transport functionalities can be seen as a method of recursion. On the other end, the resource plane in CLAS can be homologated to the forwarding and operational planes in [2], including the corresponding abstraction layer thus defined.

COORDINATED SERVICE AND TRANSPORT CONTROL IN THE VIRTUALIZED MOBILE PACKET CORE

The NFV architecture considers a number of components. While the VNFM is in charge of instantiating and controlling the EPC functions (and others distinct from EPC for complementing the service), the VIM controls the computing, storage, and networking resources associated with them.

Porting this idea to the CLAS layering, the VIM will interact with (or incorporate the capability of) the SDN controller in the service stratum when deploying the VNFs for configuring the computing and storage resources for the VNF of interest, but also for the networking part to attach those VNFs to the border of the underlying transport network to make them reachable from outside the data center. Those are the resources required from the service point of view when forming the resource plane of the service stratum.

On the other hand, the VIM has to prepare the paths in the pure transport network to make those VNFs reachable. For instance, looking at Fig. 1, the VIM has to prepare the transport network, for example, for ridding the S1 or S5 interfaces on top of it. Then the SDN controller in the transport stratum takes actions for preparing those paths, attaching the needed resources from the service stratum as endpoints for those paths.

Figure 5 presents a functional mapping of the NFV architecture to the CLAS concept. According to this mapping, the management of the VNFs (via the VNF manager) resides in the service stratum. This component should also config-

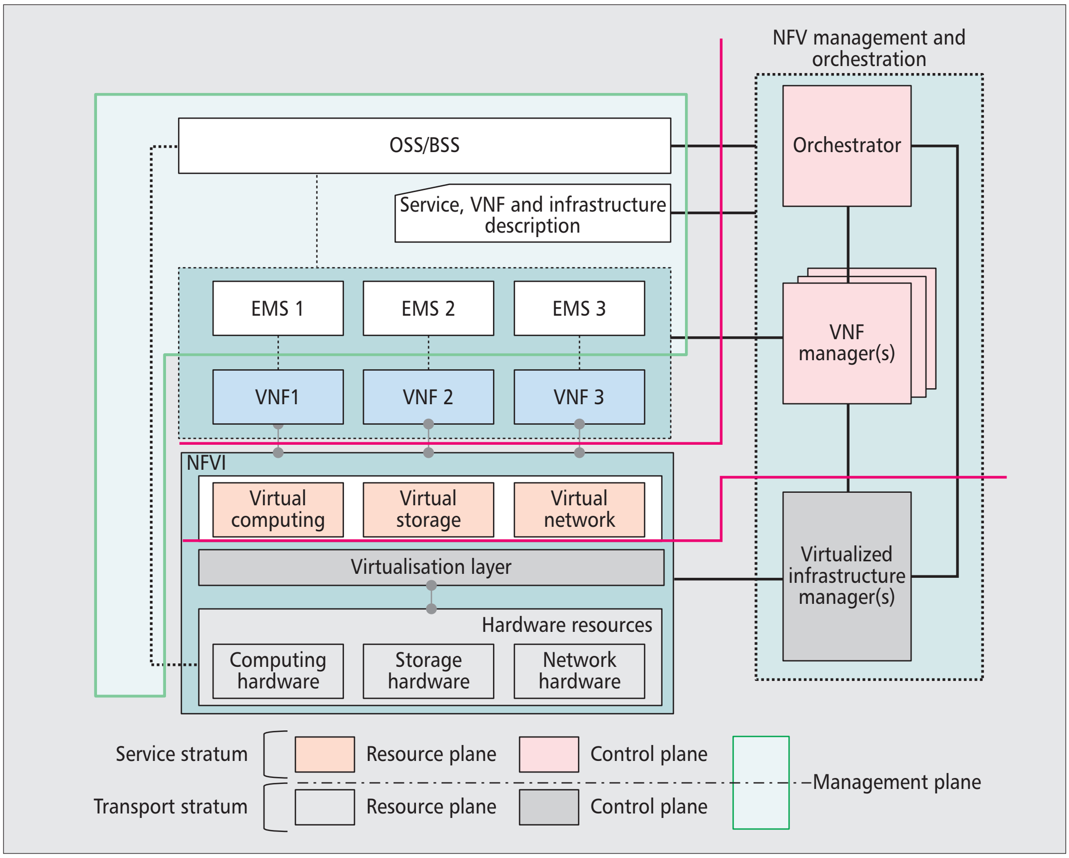

Figure 5. Layered approach to NFV control.

NFV will provide a new dimension in network flexibility by allowing a dynamic instantiation of network functions across the network. To complete the picture, coordination between the service and transport components is required to facilitate an agile composition of E2E mobile services.

ure the required resources where the VNFs are built on top, for either computing, storage, or networking. That control role in the service stratum performed by the service SDN controller in the CLAS proposal corresponds to the NFV domain SDN controller in Fig. 3.

On the other hand, the handling of the hardware resources needed to provide convenient transport of data flows between separated VNFs or VNFs and access nodes will be the responsibility of the transport SDN controller. That role is played by the E2E connectivity domain SDN controller in Fig. 3 (equivalent to the transport SDN controller in CLAS) offering the OF-mpc southbound interface.

RELATED WORK

In recent years, the concept of SDN and NFV technologies in the mobile network has attracted significant attention in the standardization bodies and research fields. In fact, the SDN and NFV topics in the mobile networks arena have addressed several topics such as heterogeneous wireless networks [9], mobile backhauling [11], LTE/EPC [10], and prominent generic use cases in wireless networks [12]. The authors of CellSDN [13] described a generalized concept of an SDN-based cellular core network for low-cost management. This paper considered LTE technology, but lacked detailed modification requirements for LTE/EPC standards. Specifically, the LTE/EPC and their functions are not studied in this paper. This work is more closely related to a non-3GPP access network like WiFi.

On the other hand, the work in [14, 15] can be viewed as complementary to our work, emphasizing the key role that software-ization and virtualization perform in the future mobile network. In [14], the authors present a study on the evolution of cloud-based EPC, where all the control functions of the SGW, PGW, and MME are “lifted up” into the cloud. The user plane is shifted into the OpenFlow switches, and these switches are extended to support GTP. The authors of [15] propose the MobileFlow architecture for future carrier networks and provide a functional evaluation of a demonstrator. In this architecture, the user and control planes of the EPC are split into a MobileFlow forwarding engine (user plane) and MobileFlow controller (control plane). The eNB participates in this functional split, and the entire control plane (EPC and eNB) is centralized. Both aspects are different from our approach. In fact, the eNB split makes processing more complex and control messages between the eNB user and control planes more frequent. For instance, with frequent UE idle/active and handover messages, the eNB needs to communicate with the centralized eNB-C to obtain the actions, which is also another major difference compared to our approach. Indeed, any changes/split in the eNB needs to transform the existing deployed eNBs in the entire network.

Table 1 shows a comparison matrix of the solution proposed in this article with respect to the above mentioned related work. The incremental step of this work is to propose an integrated architecture considering different

Solution	Network segment	Compatibility with 3GPP	Southband interface	Integration with NFV
CROWD [9]	Mobile backhaul and radio access	No	OpenFlow (switching) and others (radio)	No
S.B.H. Said *et al.* [10]	EPC and radio access	No	OpenFlow	Possible
Costa-Requena [11]	Mobile backhaul and EPC	No	OpenFlow	Possible
Bernardos *et al.* [12]	EPC and radio access	Possible	OpenFlow, ForCES, NETCONF	Yes
SoftCell [13]	EPC and radio access	No	OpenFlow	No
Kempf *et al.* [14]	EPC and radio access	Yes (control entities)	OpenFlow	Possible
Mobileflow [15]	EPC and radio access	Yes (control entities)	MobileFlow and Open-Flow	Possible
Described in this article	EPC	Yes	OpenFlow	Yes

Table 1. Comparison to related work.

proposals in ONF, ETSI, and the Internet Engineering/Research Task Force (IETF/IRTF) for SDN control of a virtualized packet core. Compatibility with existing 3GPP networks is ensured by not modifying the interfaces dealing with entities external to the packet core. OpenFlow is the southbound protocol of choice, leveraging on the work currently done in ONF for extending such protocol for support of mobile packet core functionalities in a standard manner.

CONCLUSIONS

In this article, we have presented high-level software defined control of virtualized mobile packet core network architecture, and discussed the configuration of mobile control entities and their interfaces in the virtualization platform. The SDN- and NFV-based MPC will foster the services and rollout of new network features to the network through flexibility and programmability.

Current efforts in ONF will facilitate SDN/Open Flow architecture to abstract PDN connection segments from mobile access. However, further work is needed to enable a transport network that is able to support capabilities such as online charging or deep packet inspection. ONF is making efforts to use hybrid OpenFlow switches and specialized nodes to implement complex functionality in 3GPP. In addition, carrier grade features including scalability with multi-controllers and recovery from failure need to be addressed.

Similarly, NFV will provide a new dimension in network flexibility by allowing dynamic instantiation of network functions across the network. To complete the picture, coordination between the service and transport components is required to facilitate agile composition of E2E mobile services.

The proposed approach in this article is simple and threefold:

- Separation of control and user planes
- Separation of hardware and software
- Separation of services and transport control

In addition, in our proposal no changes are required to existing EPC interfaces specified by the 3GPP standard and interoperability toward existing approaches and deployments of EPC.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers, Serge Manning (ONF WMWG Chair), and Carlos J. Bernardos (Universidad Carlos III) for their helpful and constructive comments that greatly contributed to improving the overall quality of the article.

REFERENCES

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018.

[2] E. Haleplidis *et al.*, "Software-Defined Networking (SDN): Layers and Architecture Terminology," RFC 7426, Jan. 2015.

[3] Network Functions Virtualization Introductory white paper, v. 2, http://portal.etsi.org/NFV/NFV_White_Paper2.pdf.

[4] 3GPP TS 23.401, "Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access," Release 12, v. 12.0.0, 2013.

[5] ETSI GS NFV 001, "Network Functions Virtualisation (NFV); Use Cases," v. 1.1.1, Oct, 2013.

[6] S. H. Yeganeh, A. Tootoonchian, and Y. Ganjali, "On the Scalability of Software-Defined Networking," *IEEE Commun. Mag.*, vol. 51, pp.136, 141, Feb. 2013.

[7] 3GPP TS 36.413, "Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 Application Protocol (S1AP)," Release 12, v. 12.0.0, 2013.

[8] L. M. Contreras, C. J. Bernardos, and D. Lopez, "Cooperating Layered Architecture for SDN," draft-contreras-sdnrg-layered-sdn-01 (work in progress), Aug. 2014.

[9] H. Ali-Ahmad *et al.*, "CROWD: An SDN Approach for DenseNets", *2nd Euro. Wksp. Software Defined Networks*, Oct. 2013.

[10] S. Ben Hadj Said *et al.*, "New Control Plane in 3GPP LTE/EPC Architecture for On-Demand Connectivity Service", *IEEE 2nd Int'l. Conf. Cloud Networking*, Nov. 2013.

[11] J. Costa-Requena, "SDN Integration in LTE Mobile Backhaul Networks," *Int'l. Conf. Information Networking*, 10–12 Feb. 2014, pp. 264–69.

[12] C. J. Bernardos *et al.*, "An Architecture for Software Defined Wireless Networking," *IEEE Wireless Commun.*, vol. 21, no. 3, June 2014, pp. 52–61.

[13] X. Jin *et al.*, "SoftCell: Scalable and Flexible Cellular

Core Network Architecture",," *ACM CoNEXT '13*, New York, NY.

[14] J. Kempf *et al.*, "Moving the Mobile Evolved Packet Core to the Cloud," *IEEE 8th WiMob*, 8–10 Oct. 2012.

[15] K. Pentikousis, Y. Wang, and W. Hu, "Mobileflow: Toward Software-Defined Mobile Networks," *IEEE Commun. Mag.*, vol. 51, no. 7, July 2013.

BIOGRAPHIES

MALLA REDDY SAMA (mallareddy.sama@orange.com) is a research engineer at Orange Labs, France, where he is currently working on the research and development of mobile packet core virtualization using SDN and NFV technologies. He earned his Master's degree in telecommunication engineering from Politecnico di Torino, Italy. During his Master's work he spent a year at Telecom SudParis, France, as part of an Erasmus scholarship. His research interests lie in the field of 3GPP next generation mobile packet core and wireless networks, SDN, network virtualization, mobile cloud networking, and distributed networks. He has been on the Technical Program Committees of international conferences as well as an active contributor to ONF standardization.

LUIS M. CONTRERAS completed a six-year Telecom Engineer degree at the Universidad Politécnica of Madrid (1997), and also holds an M. Sc. in telematics from the Universidad Carlos III of Madrid (2010). In 1997 he joined Alcatel Spain, taking several positions (R&D, standardization, product development, and customer engineering) in both the wireless and fixed network fields. In 2006 he joined the Network Planning department of Orange Spain (France Télécom group) taking on responsibilities in IP backbone planning. Between 2002 and 2010 he was also an adjunct lecturer in the Telematics Department of the Universidad Carlos III, where he is currently a Ph.D. student. Since August 2011 he has been part of Telefónica I+D/Telefónica Global CTO, working on SDN, transport networks, and their interaction with cloud and distributed services. He has participated in the FP7 project GEYSERS and FI-PPP project XIFI. In July 2012 he co-organized the International Workshop on Cross-Stratum Optimization for Cloud Computing and Distributed Networked Applications, co-located with the 10th IEEE International Symposium on Parallel and Distributed Processing with Applications, Leganés, Spain. He is an active contributor to the ONF and IETF, co-authoring two RFCs.

JOHN KAIPPALLIMALIL (John.Kaippallimalil@huawei.com) is a senior standards and research manager at Huawei USA with more than 20 years of experience in IP technologies, mobile networks, and telecom product design. He is currently working on next generation architecture for 5G mobile networks, and his interests include SDN and virtualization of network functions. He is an active contributor to IETF and ONF, and currently leads the ONF MPC project. He received his Master's degree in computer science from Texas Tech University in 1993 with a thesis on real-time distributed computing and monitoring.

IPPEI AKIYOSHI (i-akiyoshi@ah.jp.nec.com) received his degree in communication engineering from Osaka University, Japan, in 2002. He has been with NEC since 2002, mainly working on the research and development of wireless and mobile networks such as 3GPP and WiMAX. He is currently focusing on SDN and network virtualization for wireless and mobile networks, and is involved with ONF standardization.

HAIYANG QIAN received his B.S. degree in communication engineering from Nanjing University of Science and Technology, China; his M.S. degree in telecommunications from the Technical University of Denmark, Lyngby; and his Ph.D. degree from the University of Missouri — Kansas City. He worked at PhoneFactor Inc. (acquired by Microsoft) and Futurewei Technologies in engineering roles. He was at China Mobile U.S. Research Center as a senior researcher. He is the co-founder of Grabtalk. He has published over 10 peer reviewed papers and filed a few patents in the NDN, SDN, mobile Internet, and cloud computing fields. He has served on the Technical Program Committees of a number of international conferences. His current research interests include cloud computing, software-defined networking, and mobile Internet systems and products.

HUI NI (hui.ni@huawei.com) received his Ph.D. degree in computer science from Tianjin University in 2007. In the same year, he joined the Wireless Research Department of Huawei as a researcher and standards delegate. From 2008 to 2011, he actively contributed to the WiMAX Forum and held leadership of the WiMAX-3GPP interworking group. Currently, he focuses on evolution of next-generation wireless networks, and his research interests include SDN, NFV, ICN, and IoT. He is an active contributor to ONF.