Traffic Jam

Handling the Increasing Volume of Mobile Data Traffic

Toktam Mahmoodi
and Srini Seetharaman

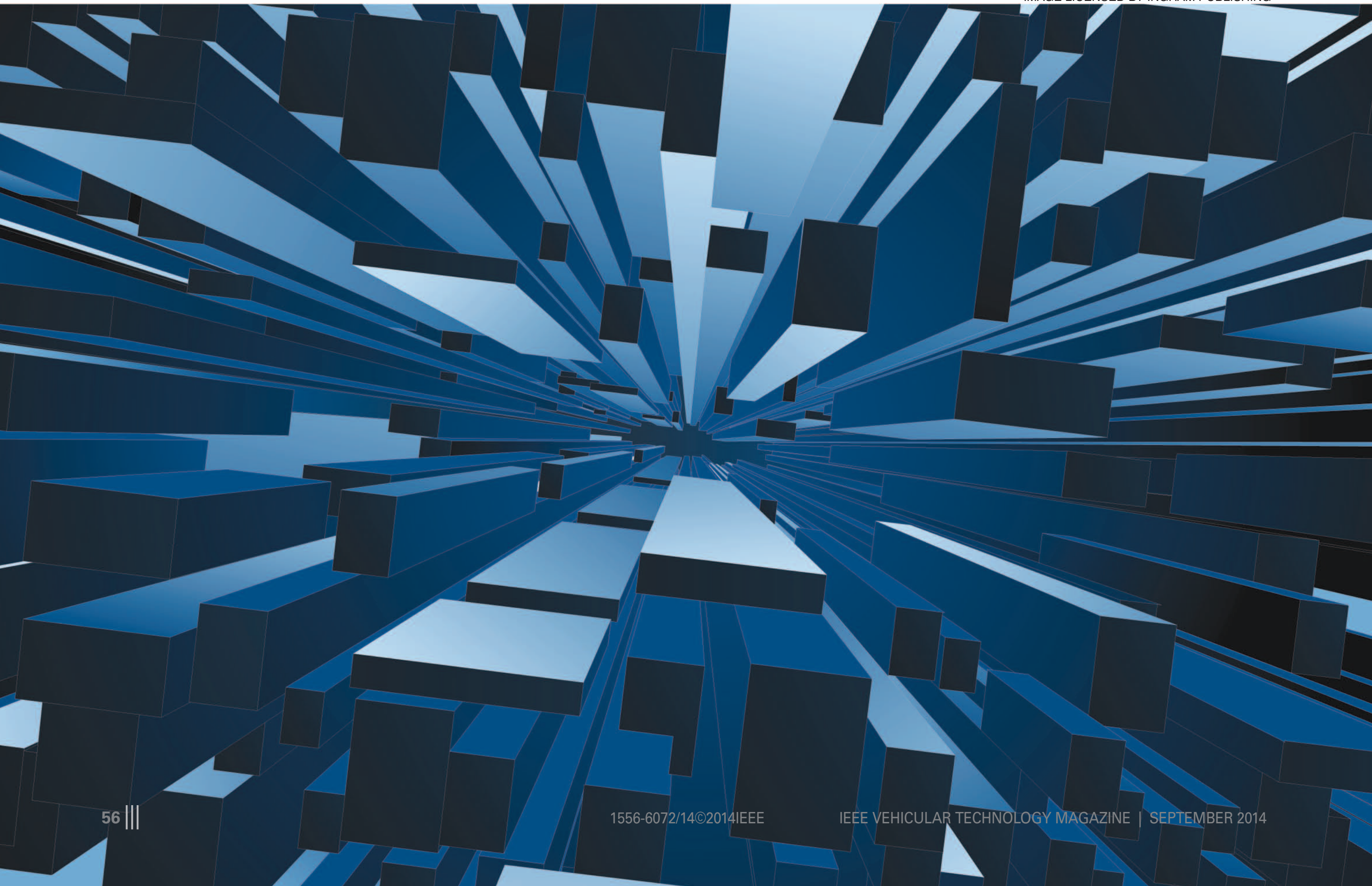
Today's mobile operators face significant challenges with handling the ever-increasing volume of mobile data traffic. With new mobile communication standards, the mobile backhaul architecture has a clear split of a packet-only data plane and a management plane. Although this new backhaul architecture yields to easier management, we find that this architecture can be improved further by applying the principles of software-defined networking (SDN). SDN allows for better evolvability of the data plane without depending on a slew of management or control protocols, allows for centralized control of the overall infrastructure, and allows for a richer feature set based on its programmable nature. This article investigates the redesign and illustrates its potential with mobility management as an example.

Digital Object Identifier 10.1109/MVT.2014.2333765
Date of publication: 10 September 2014

Introduction

To address the demands of increasing mobile data traffic, the network capacity must be increased despite the higher costs to build, maintain, and upgrade radio

IMAGE LICENSED BY INGRAM PUBLISHING



access networks. On the other hand, mobile operators are increasingly turning to small cells (e.g., femtocells) to deal with the capacity crunch in dense urban areas and to add coverage in areas with low signal levels. This will result in an increased rise of mixed mobile networks of Wi-Fi, cellular small cells, and traditional base stations, and opens new challenges in optimizing users' access to the networks.

At the same time, the industry has converged on a new standard for mobile communication. With fourth generation (4G)/long-term evolution (LTE), the mobile backhaul architecture has a clear split into 1) a packet-only data plane comprising a base station (eNodeB), serving gateways (S-GWs), and a packet data network (PDN) gateway (P-GW) and 2) a management plane to manage mobility, policies, and charging rules, comprising a mobility management entity (MME), policy charging and rules function (PCRF), and a home subscriber server. Although this new backhaul architecture yields to easier management, we find that this architecture is still not as evolvable, flexible, or programmable as it can be.

We posit that the mobile communication architecture, in the evolution toward the new generation of mobile networks, i.e., fifth generation (5G), can be further improved by applying the principles of SDN. The main design goals for adopting an SDN-based control plane for the mobile backhaul are the following:

- *Centralized control of the overall infrastructure*: Managing thousands of cell sites and the corresponding infrastructure is an arduous task. Mobile backhaul operators have already mastered this aspect. However, having fine-grain control on the devices is crucial for improving state management and enabling other value propositions. Furthermore, there is an increasing signaling overhead and convergence penalty in the current architecture that can only be avoided through centralized control.
- *Evolvability of the data plane*: Current data plane elements keep significant state and speak a slew of management protocols. Applying SDN principles leads to a flattening of the data plane by simplifying them into pure forwarding elements and exporting the control plane intelligence to a remote controller node. This architecture requires minimal change during feature upgrades and other evolutionary steps.
- *Programmability*: Today's mobile backhaul is capable of offering quality of service (QoS) guarantees and differentiated services to the user, using PCRF nodes. Introducing more new features, however, has become increasingly complicated because the architecture was not designed to allow it. For example, distributed and coordinated network policy control can be programmed in the network and can replace the traditional centralized or preset models of enforcement. This replacement can significantly optimize the network performance and can also reduce the signaling overhead [2]. Programmability can solve these and other similar predicaments.

Although adopting SDN seems like a clear choice, it is yet unclear what the quantitative gains will be. In this article, we investigate the redesign of the mobile backhaul and look deeper into its potential gains with mobility management; our metrics of interest are signaling overhead and handover delay.

SDN-Based Control Plane

Background on Mobile Backhaul

Today's mobile network starts from the user equipment (UE) that connects to an eNodeB. The eNodeB directs traffic through a S-GW over a tunneling protocol called *general packet radio service* (*GPRS*) *tunneling protocol*. The S-GW serves as a local mobility anchor that enables seamless communication when the user moves from one eNodeB to another. Furthermore, the S-GW tunnels traffic to the PDN gateway P-GW. The P-GW enforces QoS policies, monitors traffic, and acts as a firewall. Besides data plane functionalities, the eNodeB, S-GW, and P-GW also participate in several management-plane protocols. In coordination with the MME, they perform signaling to handle session setup, teardown, and reconfiguration, as well as mobility, e.g., location update, paging, and handover. The IP-based core network architecture in the LTE, the evolved packet core (EPC), supports seamless mobility for both intratechnology and intertechnology handovers (e.g., to or from older network technology).

The evolutionary path for the mobile backhaul is riddled with standards and vertical dependencies. Figure 1 shows a summary of the evolution process of the mobile communication standards and the corresponding evolution

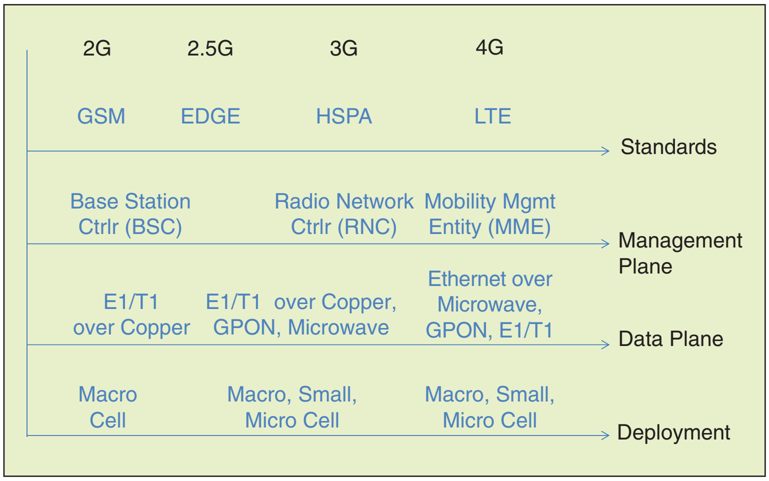


FIGURE 1 An evolution of the mobile backhaul.

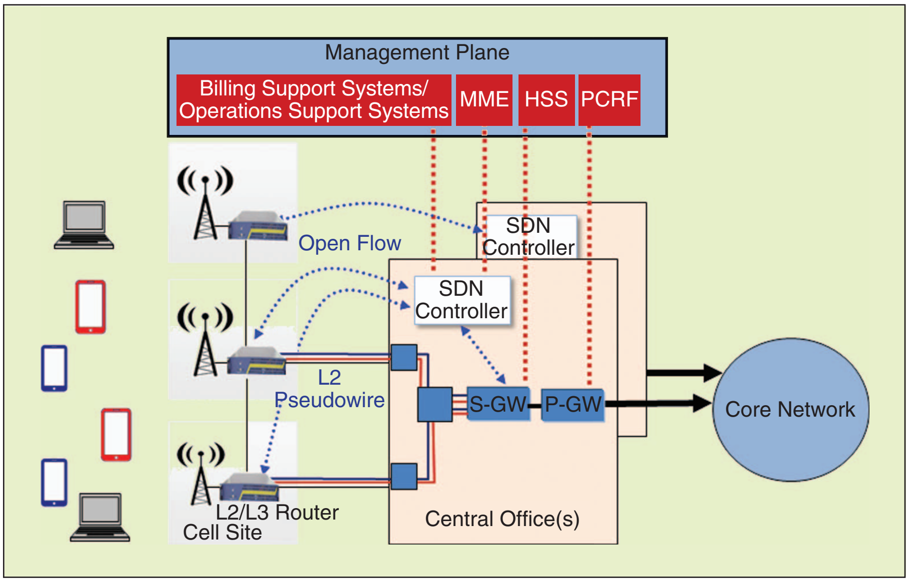


FIGURE 2 The SDN controller located within the mobile backhaul and its communications with other entities in the network, including eNodeB, S-GW, and P-GW.

across the different layers. One aspect that stands out is that the evolution has lacked a cross-layer isolation because the industry has been constantly retrofitting earlier standards with new requirements. We posit that an SDN-based design would allow for evolution of each aspect of the mobile network.

SDN-Based Design

Similar to an earlier proposal on the software-defined cellular network [1], we propose a distributed SDN control plane that is deployed in the central office of the carrier. The SDN controller implements multiple applications, such as mobility management, pseudowire creation, and assignment, and allocation of radio resources, while also interworking with other programmable elements in the protocol stack; examples of programmable media access control and physical layer (PHY) are in [3], [4]. Compared with second generation (2G) and third generation (3G), the newer generation networks are designed in such a way that separation of control plane functionalities fits well within its current architecture of the core and backhaul.

Despite the natural fit, introducing an SDN-based control plane will raise several challenges, including supporting subscriber scale, real-time decision making, and fine-grained measurement and control. Previous work discussed some of these issues. The most important of all is the time scale in which many mobile network decisions are made, which is relatively short compared to the fixed network, and the rapidly changing nature of wireless media is one of the main reasons for that.

The design of the control plane can be summarized in the following steps:

- Separating the control plane functionalities in all network entities starting from the eNodeB up to the P-GW.
- Implementing the control plane functions as SDN controller applications. The traditional functions incorporated in the data plane of the mobile network, including mobility management and resource allocation, are decoupled and ported to the control plane.
- Encapsulating the management plane functions within individual virtual machines (VMs) that interact with the SDN controller. The SDN controller can potentially have control over the number, placement, and workload assignment of instances of each function instantiated as VMs.
- The SDN control plane works in a distributed manner, sharing state among individual controller instances (as can be observed in Figure 2).

Considering that 2G, 3G, and 4G networks are all simultaneously active in today's mobile network, a clean slate approach is not justifiable. Thus, we adopt the described approach to ensure that SDN is introduced within the existing operational mobile network. With our approach, we retain the existing management plane while introducing significant flexibility and programmability through a newly introduced control plane. Since the management plane is also software defined, we foresee that the control plane may, in the long term, subsume the functions offered by the management plane. This carefully staged deployment is crucial in introducing SDN to some of the traditional carrier network. Similar efforts exist in introducing SDN to the access network through OpenFlow-based digital subscriber line access multiplexer hardware and BRAS controller [5].

We require software data plane agents (like Open-vSwitch [6]), installed in all devices, that can be controlled by a logically centralized SDN controller. There are several hints that the industry is headed in this direction: mobile backhaul vendors (such as Ericsson) are in the process of implementing OpenFlow support in their devices; cell-site routers can be implemented with commodity OpenFlow switches managed by a RouteFlow-like control plane application [7]; momentum in the industry to implement OpenFlow and OpenvSwitch in mobile handhelds (such as the Android). Thus, gradually all data plane elements in the mobile backhaul will be SDN-enabled, thereby aiding in the evolution of the whole network.

Note that we are still working on identifying the additional primitives that the data plane elements should support to guarantee all common real-time functions (such as wireless resource allocations).

SDN Controller Implementation

In the SDN-based design, the control plane is decoupled from the data plane and the control aspects are incorporated in a distributed SDN controller. The data plane is reduced to a set of SDN-enabled routers for the primary task of packet forwarding. The control plane, in this decoupled architecture, handles all necessary signaling and also manages the mobility of the UE.

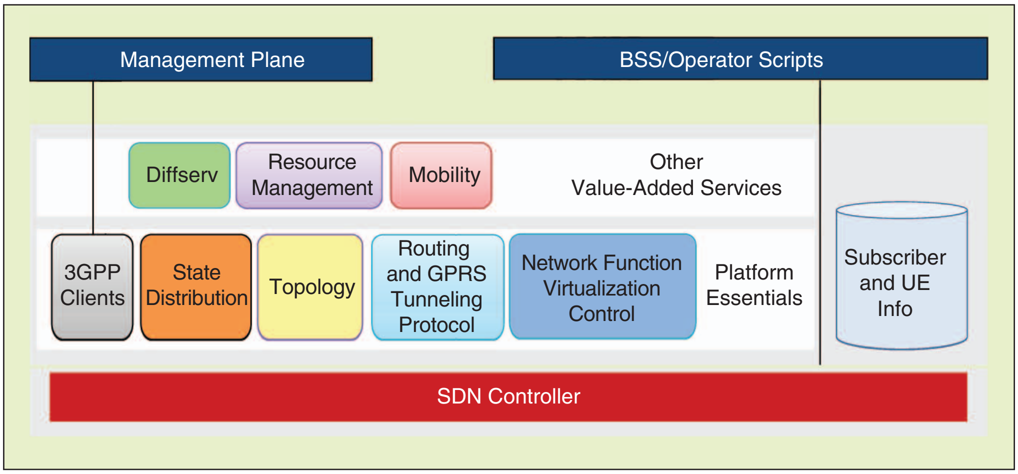


FIGURE 3 The design of the SDN controller for the LTE network. The platform essentials are involved in state dissemination, L3 routing, tunnel setup, and interfacing with management plane. The value-added services handle intelligence for automatically managing resources, mobility, DiffServ/QoS, and network selection for UE.

We envision the SDN controller implementation to be based on OpenFlow [8], with two classes of modules (viz., platform essentials, value-added services), as shown in Figure 3. The SDN controller will push the rules to all devices in a proactive manner to handle the large scale of subscribers.

Advantages of SDN

Earlier work [1] presented several examples of features that can be implemented as SDN applications for today's mobile backhaul; examples include selectively routing traffic through middleboxes, traffic monitoring, and infrastructure slicing to allow virtual operators. Similarly, our design can enable new value-added services, such as location-based policy adaptation, application customization, and location-aware services for the mobile operators.

Adopting SDN in the EPC will help realize the following longer-term benefits:

- It will enable interworking between different mobile standards and allow for cross-technology handoff without incurring too much signaling and setup time.
- It will enable the introduction of new PHY technologies for improving coverage and throughput. For instance, there are new vendors that improve the performance of small cells through advances in PHY technologies, e.g., Kumu Networks and Fast Networks.

As case studies, we discuss the benefits of SDN-based mobility management in the following section.

Case Study: Mobility Management in the LTE Network

In this section, we first elaborate the handover procedure as performed in the LTE network and provide a brief detailing of the signaling involved. We then discuss the signaling involved with the mobility of the UE if the EPC data plane is SDN-based and the handover is managed by an SDN controller. We detail arguments on the gain achieved in terms of signaling overhead, energy, and handover delay.

LTE Handover Procedure

The procedure of handover of a UE from one eNodeB to another consists of three main phases: 1) preparation, 2) execution, and 3) completion. Each of these steps involves sets of signaling between the UE, eNodeB, and EPC. We elaborate each of these three phases next. The exact signaling handshake during the handover is shown in Figure 4.

In the planning phase, the UE measures downlink signal strength and analyzes the measurements. After that, the UE transmits the measurement report to the serving eNodeB, which allows the eNodeB to make a handover decision. The signaling messages 1 and 2 in Figure 4 represent this phase. In the execution phase, the UE detaches from the source eNodeB and synchronizes with the new cell, which corresponds to the signaling messages of three to nine. The source eNodeB will transfer the buffered and in-transit packets to the target eNodeB at this stage. In the completion phase, the S-GW switches the path to the target cell, exchanges messages with the MME and asks the old cell to release resources. As observed from Figure 4, the majority of the signaling messages of this phase are communicated over the backhaul; the bandwidth of the backhaul comprises the most scarce resources in the mobile network.

Handover When Using SDN Controller at the EPC

In the previously described scenario, handover is eNodeB-driven and majority of the signaling messages are communicated between the source and target eNodeBs, or between eNodeBs and the MME. After the execution of handover, the data plane forwards the buffered packets and subsequent data packets to the target eNodeB. Using an SDN controller can provide a common control protocol (e.g., OpenFlow) that manages multiple eNodeBs and works

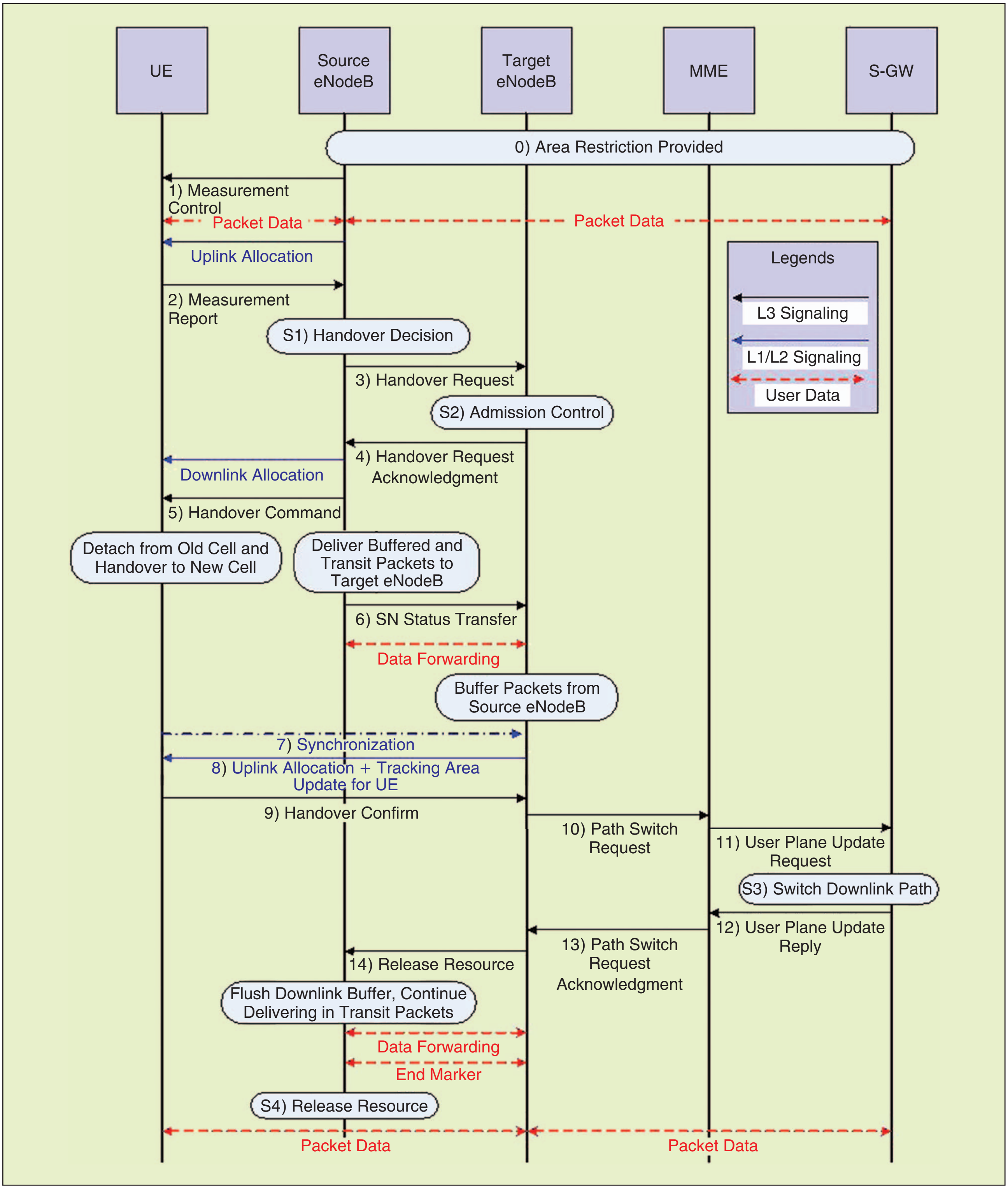


FIGURE 4 Handover signaling in today's LTE network: the red dashed line represents the user data, the blue line represents the L1/L2 signaling, and the black line represents the L3 signaling.

across different technologies. The controller can push new forwarding rules and, thus, various request/confirm messages between eNodeBs (or other wireless access points) can be avoided. The OpenFlow-enabled routers in the data plane will enable incoming packets to be routed according to the flow action corresponding to the handover decision. Finally, the selection of the target eNodeB can be based on a set of predefined policies. Applying these modifications to the procedure of handover affects the sequence of signaling messages that was previously shown in Figure 4.

We elaborate the effect of this new mobility management architecture on the power consumption of the UE and the handover signaling in the following sections.

Gain at the UE Side

At the UE side, one of the main constraints is power, i.e., the battery energy is one of the most precious resources at the UE. The energy is consumed mostly in the planning phase, where the UE scans and measures the received signal strength of all available radio access networks. This phase, however, can be considerably shortened by scanning only a subset of radio accesses. The SDN controller can greatly facilitate this shortening of the planning phase by keeping the record of available access networks and their potential to offer services, and remotely controlling the OpenvSwitch embedded in the UE to select the right network. In the event of no SDN support at the UE, the UE can at least query the central SDN control plane to obtain the necessary information. Thus, in the case of handover, the scanning and measurement phase can be shortened, resulting in considerable savings in power consumption. Quantifying the exact power savings is part of our ongoing research effort.

Gains in the Backhaul

At the backhaul side, the scarcity of communication resources is more crucial and, thus, reducing the signaling message overhead between various entities is one of the major goals. In this regard, we focus on the signaling messages that can be avoided by introducing the SDN controller at the backhaul. We estimate that the number of signaling messages can be reduced to less than 50%. According to a recent report [9], signaling traffic will grow three times faster than mobile data traffic through 2016. Thus, the significant reduction in handover signaling (50%) could greatly contribute to reducing the signaling overhead and the total traffic carried in the backhaul.

In addition to reducing the communication overhead, the signaling in the SDN-based architecture also reduces the handover delay. An Ericsson test report states that the handover delay can be 0.7 s when both planning and execution phases are included [10]. This delay can have a worst-case value of up to 2 min during the handover between the 3G and LTE networks [11]. This handover delay is even more significant in the heterogeneous scenarios and in the handover between multiple technologies. Our hypothesis is that the handover delay can be reduced to two to three times the round-trip time between the UE and the SDN controller. Thus, we gain a faster handover and lower packet loss when using an SDN-based control plane. The signaling messages that can be omitted are detailed in the following.

Handover Request and Handover Request ACK

Once a measurement report is received from the source eNodeB and the target eNodeB, and a handover decision is made, the flow table of the SDN-enabled router can be programmed to route subsequent packets to the target eNodeB. Thus, further signaling messages between the two eNodeBs to complete the handover can be avoided. The SDN router signals the target eNodeB to synchronize with the UE.

SN Status Transfer, Path Switch Request, Path Switch Request ACK, User Plane Update Request, and User Plane Update Reply

With these messages, the MME that has the "track area list" can update the flow table and have the SGW buffer packets in transit. This reduces the unnecessary buffering of packets from the source eNodeB and allows us to release the source eNodeB resources prior to handover.

Release Resource

The source eNodeB can release resources without notification and the flow will be directed to the target eNodeB after the UE synchronizes with it.

Concluding Remarks

In this article, we built upon an earlier work [1] to propose a backward-compatible modular redesign of the backhaul of mobile networks, based on the principles of SDN. We argued that the new design based on the centralized SDN controller not only allows for programmability and evolution of the mobile networks but also, in the long run, reduces the various maintenance and upgrade costs of the network. Moreover, the SDN-based controller can facilitate several value-added services, including policy-based services, for the mobile operators.

In addition to presenting our envisioned design of the SDN-based backhaul for the next generation mobile network, i.e., 5G, we looked deeper into the case study of mobility management and observed the advantages achieved by reducing the 1) energy consumption at the UE, 2) the communication overhead of the signaling over the backhaul, and 3) the handover delay for the UE. Our main future work is toward quantifying the gain through a realistic implementation. Our back-of-the-envelope calculation shows at least a 50% reduction in signaling overhead. Considering that the number of mobile subscribers worldwide has exceeded six billion in 2012 [12], and the high growth of signaling traffic [9], the savings has a tremendous impact on the overall backhaul infrastructure.

Future Work on Use Cases

Besides the quantification of the impact during mobility, we plan to examine two other use cases: 1) the mobile data offloading and 2) the selection of access networks for offloading flows. This is particularly challenging in today's mobile networks because of the increased heterogeneity in the small cells and their coexistence with the macrocells. Such challenges motivated standard bodies to elaborate the interworking between LTE and other technologies [13], and also industries to investigate various offloading strategies [14]. Using SDN-based architecture for data offloading-based network condition and accounting policies is discussed in [15]. Furthermore, increased in the heterogeneity of radio accesses, and mobile devices are equipped with multiple radios, open new challenges in the selection of radio access networks and results in fairly complex algorithms. Cloud-assisted execution, and also selecting the network based on wisdom of the crowd, can be an alternative solution to reduce the complexity of task at the UE [16]. In the architecture presented in [16], an aggregated set of information from a UE's experience in connection to different radio access networks is stored at a local cloud and can be accessed by all neighboring UEs in making the selection decision. One of the best candidate for such a point of reference is the SDN-controller, where properties of the access networks of different technologies are available. Furthermore, various policies, such as charging and accounting policies, and those related to the QoS that affect the selection of network can remotely control the OpenvSwitch embedded in the UE.

Future Work on Platform Challenges

Ensuring compatibility with the current architecture of the mobile network is challenging but crucial because the network infrastructure belongs to various operators and covers a range of technologies. Another challenge that is specific to the mobile network is the time frame in which decisions related to the allocation of wireless and network resources should be made. Our aim is to extend our design and implementation of the SDN controller to address the functionalities of the mobile networks, the policies with which routing and mobility issues are addressed, and the physical location in which various network services can be offered depending on their delay constraint or their running time frame.

Author Information

Toktam Mahmoodi (toktam.mahmoodi@kcl.ac.uk) received her Ph.D. degree in telecommunications engineering from Kings College London in 2009 and her B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2002. From 2006 to 2009, she was a researcher with the Core-4 Efficiency Programme of the Virtual Centre of Excellence in Mobile and Personal Communications (Mobile VCE). From 2010 to 2011, she was a postdoctoral research associate in the Intelligent Systems and Networks group of the Electrical and Electronic Engineering Department of Imperial College London. She is a Member of the IEEE, the Institution of Engineering and Technology, and the Association for Computing Machinery (ACM) and has been on the technical program committee and organizing committee of a number of IEEE and ACM flagship conferences.

Srini Seetharaman (srini.seetharaman@telekom.com) received his Ph.D. degree in computer science from the Georgia Institute of Technology, Atlanta. From 2008 to 2011, he was a member of the OpenFlow/software-defined networking (SDN) team at Stanford University, where he led the SDN deployments in several nationwide campus enterprise networks, including Stanford. He is currently the technical lead for SDN at the Deutsche Telekom Innovation Center. He is a recipient of the Future Internet Design grant from the U.S. National Science Foundation for his proposal on service virtualization.

References

[1] L. E. Li, Z. M. Mao, and J. Rexford, "Toward software-defined cellular networks," in *Proc. European Workshop Software Defined Networking*, Oct. 2012, pp. 7–12.

[2] Sandvine. (2011, Aug.). Distributed decisions: Hierarchical network policy control. White Paper, Sandvine Inc., Waterloo, ON, Canada. [Online]. Available: https://www.sandvine.com/downloads/general/whitepapers/distributeddecisionsinnetworkpolicycontrol.pdf

[3] I. Tinnirello, G. Bianchi, P. Gallo, D. Garlisi, F. Giuliano, and F. Gringoli, "Wireless MAC processors: Programming MAC protocols on commodity hardware," in *Proc. IEEE INFOCOM*, June 2012, pp. 1269–1277.

[4] M. Bansal, J. Mehlman, S. Katti, and P. Levis, "OpenRadio: A programmable wireless dataplane," in *Proc. ACM SIGCOMM Workshops*, Aug. 2012, pp. 109–114.

[5] Floating BRAS: Example of how SDN enables virtualized service nodes over access/aggregation network Open Networking Summit. (2012, Apr.). Open Networking Summit Demo, Santa Clara, CA. [Online]. Available: http://www.opennetsummit.org/archives/apr12/site/pdf/sparc.pdf

[6] OpenVSwitch: An open virtual switch. [Online]. Available: http://openvswitch.org

[7] RouteFlow project. [Online]. Available: https://sites.google.com/site/routeflow/

[8] 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, 2008.

[9] Telecom Review. (2012). Tekelec LTE diameter signaling index: Forecast report and analysis 2011-2016. Tekelec White Paper, Morrisville, NC. [Online]. Available: http://www.telecomreviewna.com/index.php?option=com_content&view=article&id=286:lte-diameter-signaling-traffic-to-triple-mobile-data-traffic-through-2016&catid=44:novembre-decembre-2012&Itemid=88

[10] *Voice Handover in LTE Network*, Ericsson White Paper, Stockholm, Sweden, Oct. 2012.

[11] (2012, Apr.). Modem handoff from 3G to LTE can take two minutes. [Online]. Available: http: //www.cellular-news.com/story/46917.php

[12] Infonetics Research. (2012, Apr.). Total fixed and mobile subscribers pivot. Infonetics Res. [Online]. Available: http://www.infonetics.com/pr/2012/Fixed-and-Mobile-Subscribers-Market-Highlights.asp

[13] (2013, Mar.). 3GPP TS 23.402 V10.7.0: Architecture enhancements for non-3GPP accesses. [Online]. Available: http: //www.cellular-news.com/story/46917.php

[14] Cisco. (2012). Architecture for mobile data offload over WiFi access networks. White Paper. San Jose, CA. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/service-provider-wi-fi/white_paper_c11-701018.pdf

[15] M. Amani, T. Mahmoodi, M. Tatipamula, and H. Aghvami, "Programmable policies for data offloading in LTE network," in *Proc. Int. Conf. Communications*, June 2014.

[16] T. Mahmoodi, " Network selection for autonomous mobile devices," in *Proc. IEEE Wireless Demand Services*, Mar. 2013, pp. 125–127.

VT