



King's Research Portal

DOI:

[10.1109/COMST.2018.2836315](https://doi.org/10.1109/COMST.2018.2836315)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Zaidi, Z., Friderikos, V., Yousaf, Z., Fletcher, S., Dohler, M., & Aghvami, H. (2018). Will SDN be part of 5G? *Ieee Communications Surveys And Tutorials*. <https://doi.org/10.1109/COMST.2018.2836315>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Will SDN be part of 5G?

Zainab Zaidi*, Vasilis Friderikos[†], Zarrar Yousaf[‡], Simon Fletcher[§], Mischa Dohler[†] and Hamid Aghvami[†]

*IEEE Senior Member, Email: zzaidi@ieee.org

[†]Centre for Telecommunications Research

King's Collge London, Strand, London, WC2R 2LS, UK

Email: {vasilis.friderikos, mischa.dohler, hamid.aghvami}@kcl.ac.uk

[‡]NEC Laboratories Europe, Heidelberg, Baden-Württemberg, Germany

Email: Zarrar.Yousaf@neclab.eu

[§] Real Wireless Ltd, Pulborough, RH20 4XB, UK

Email: simon.fletcher@realwireless.biz

Abstract—For many, this is no longer a valid question and the case is considered settled with SDN/NFV (Software Defined Networking/Network Function Virtualization) providing the inevitable innovation enablers towards the realization of a virtualized, flexible, programmable and flexible 5G network. As SDN along with other technology enablers (including NFV) are still in the process of evolution, the first commercial deployment of 5G may take few years. However, some companies are claiming the availability of 5G solutions, but considering the monumental task of softwarization of mobile cellular networks there are genuine concerns that we may only see some point solutions involving SDN technology instead of a fully virtualized SDN-enabled 5G mobile network. In order to determine the technology readiness of SDN solutions in the context of 5G networks, this survey paper attempts to identify all important obstacles in the way, and looks at the state of the art of the relevant research. This survey is different from the previous surveys on SDN-based mobile networks as it focuses on the salient problems and discusses solutions proposed within and outside SDN literature. Our main focus is on fronthaul, backward compatibility, supposedly disruptive nature of SDN deployment, business cases and monetization of SDN related upgrades, latency of general purpose processors (GPP), and additional security vulnerabilities that softwarization brings along to the mobile network. We have also provided a summary of the architectural developments in SDN-based mobile network landscape, including deployment options for SDN within NFV framework, as not all work can be covered under the focused issues. This paper provides a comprehensive survey on the state of the art of SDN-based mobile networks and clearly points out the gaps in the technology.

Index Terms—5G mobile/cellular network, radio access network (RAN), SDN, NFV, control plane, data/forwarding plane, OpenFlow, CloudRAN, LTE, EPC, fronthaul, backward compatibility,

I. INTRODUCTION

While 5G is still not standardized¹, this paper is trying to point out that there is a significant chance that the idea of a

¹The Non-standalone 5G NR (New Radio) specifications are recently approved in Dec. 2017 and the rest of Rel. 15 (Standalone 5G NR) has expected functional freeze date of Sep. 2018. The work on standardization of NFV has moved already to detailed specification and proof-of-concept by ETSI. IEEE is also developing standard for NGFI, a new fronthaul technology.

fully programmable control plane with switches may not be tech-ready for 5G. The emphasis of this paper is only on the mobile or cellular networks (both access and core networks), although, 5G may encompass several other technologies for connectivity, i.e., satellite, cable, optical, WLAN (Wireless LAN) etc.

The major focus of SDN technology is towards decoupling of the software-based control plane from the hardware-based data plane (e.g., packets forwarding) of networking and switching pieces of equipment. SDN brings with it unprecedented ease in innovation, openness, optimum resource utilization, support for virtualization, etc., but it is perhaps the most disruptive idea, the mobile networks have seen since their transition from 1G to 2G, i.e., from analog to digital systems. Even LTE did not come with the similar degree of disruption [1]. The disruption required may possibly push the SDN deployments, and as a consequence the development of SDN-based solutions, to further future. 5G may just see some point solutions with SDN in some places [1] and the focus would largely remain on densification, mm-wave technology, and DAS (Distributed Antenna Systems) to achieve 1000x capacity, 100x data rate, and 100x active connections from the already deployed LTE network. An interesting point here is that all of the salient targets and business opportunities defined for 5G are not strictly dependent on SDN. Nevertheless, it is important to mention that 3GPP's 5G NR (New Radio) is a step closer to SDN-based mobile networks due to complete separation of control plane (CP) from user plane in NGC (New Generation Core), modularizing of CP functions, and defining the interfaces between CP functions as services [2].

In recent years, there has been a lot of interest in SDN-based mobile network and there are a number of papers proposing SDN-based mobile networks architectures and listing the huge benefits they can bring to the mobile industry [3]–[11]. These papers were basically advocating the case for SDN by emphasizing the positives, a very understandable trend to address the need of the time. As the time passes, researchers are also raising concerns about getting the technology ready for deployments and respective obstacles along the road. The state of the art of SDN technology in wired and fixed network is much more advanced than the SDN-based mobile networks developments. However, with 87% of the total Internet users now

in possession of a smart-phone², mobility is very important for global telecommunication network. LTE Evolved Packet Core (EPC) has done a remarkable job in simplifying the core and separating control and user plane to an extent. 3GPP's 5G NGC has further improved this separation. The base-station eNodeB, however, still contains both planes. Moreover, shared wireless medium and interference between neighboring base-station make it more difficult to realize separation of control from forwarding plane.

We have identified six major obstacles and issues which need to be addressed in order for the technology to move forward. There are certainly other issues as well and as the technological development in this area progresses, there will be an even longer list of questions. We have selected the following six issues according to their relevance with research and innovation. Other issues, such as, re-organization of telecommunication industry to suit the SDN model, regulatory issues, etc., are outside the scope of the present paper. Nonetheless, they are equally important issues to be discussed by the SDN community. Moreover, other issues related to optimization, added functionalities, etc., are also outside the scope of present paper. Examples of such issues are energy efficiency of SDN networks, optimized resources consumption, intelligent networking techniques for big data, integration of IoT (Internet of Things), etc. [12]. However, these issues are also important, e.g., the additional energy and resource consumption introduced by SDN and NFV should also be considered and efficient techniques should be developed for greening the future networks [12]. The major obstacles are:

- 1) fronthaul;
- 2) latency of general purpose platforms;
- 3) backward compatibility;
- 4) disruptive deployment;
- 5) SDN specific security vulnerabilities;
- 6) clear and compelling business case.

We explain the issues in detail in Section III and they set the standard according to which we survey the existing proposals of SDN-based mobile networks. It is very interesting to see that although there are many publications regarding the new architecture, very few touch upon the real hurdles. Although, there has been significant efforts to explore solutions for the above mentioned issues but they are sporadic and isolated and a comprehensive solution is still missing. The main motivation behind this survey paper is to understand the state of the art and highlight the gap between what is being done versus what is needed to be done. A big picture or broad perspective is a compass to help us steer towards the goal. Although, it is not trivial to have a big picture for such a complex problem showing all significant dimensions of the problem space. Major obstacles and issues in the path of technology deployment is one way of constructing the big picture which is meaningful and very effective in understanding the state of the art. As per our knowledge, none of the surveys on SDN-based mobile networks has looked into the matter of obstacles as the main focus. In Section V, we have presented a brief

account of available surveys on SDN-based mobile networks proposals and how our paper is different from them.

Once again, we want to emphasize that this paper is not disputing the benefits of SDN rather we consider them substantial and necessary for the future generations of networks. Our question is not "why?", it is "when?". Considering the gigantic task, it is vital to prioritize the research issues in order of their impact over the SDN technology's development and realization. The main contribution of this survey paper is to provide a clearer picture to the readers about what is needed to be done regarding the realization of SDN-based mobile networks. We achieved the objective by:

- 1) identifying the major issues and obstacles which should be resolved so the technology could reach the deployment stage;
- 2) providing the state of the respective solutions, lessons learned, comparison of alternate proposals wherever possible, and highlighting the gaps through summaries, tables and lists;
- 3) pointing out the standardization activities wherever possible;
- 4) surveying the architectural development in SDN-based mobile networks around the popular themes to get an understanding of the outlook of future networks.

The paper is organized as follows: Section II contains a brief background about SDN, NFV, CloudRAN, and current and future LTE architectures. This section is added to make this paper self-contained, readers knowledgeable in the above technologies can easily skip this section and go directly to Section III. As discussed above, Section III discusses the major obstacles and issues in detail and also lists the solutions proposed in the literature. Section IV then summarizes the SDN-based mobile networks literature as the current state of the art depicting specially the achievements unrelated to the obstacles and issues identified in this paper. Section V discusses the SDN-based mobile networks' surveys and what was the need to have the present one. Finally Section VI summarizes the major take-away points made in this paper.

II. BACKGROUND

In this section, we present brief discussions about SDN, NFV, CloudRAN and LTE including some future proposals for mobile network architecture.

A. SDN

SDN mainly focuses towards decoupling of the software-based control plane from the hardware-based data plane (e.g., packets forwarding) of networking and switching pieces of equipment. The basic SDN model is shown in Fig. 1. The logically centralized controllers contain the control logic to translate the application requirements down to the data plane and are responsible in providing an abstract network view to the application plane. The major issue is to create appropriate mapping of the existing network functionalities to the decoupled control and forwarding planes. Conforming to the terminologies used in the research community, the interface

²<http://www.globalwebindex.net/blog/87-of-internet-users-now-have-a-smartphone>

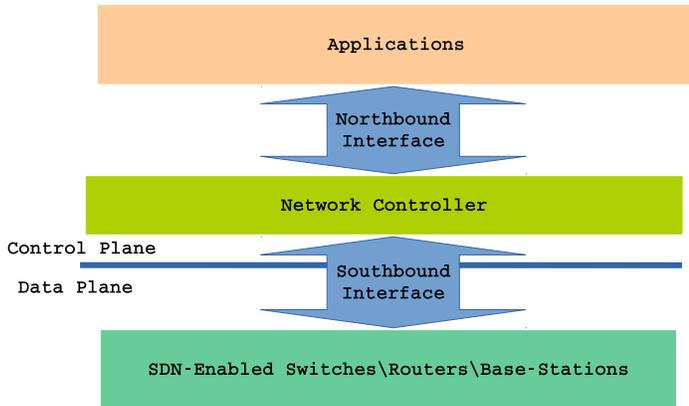


Fig. 1. A generic SDN model of three layers.

south of control plane is called the southbound interface in Fig. 1 and north of control plane is the northbound interface. An important point to note here is that both interfaces define the bi-directional information flow contrary to what may be implied by their names.

The major benefit of softwarization is flexible innovation as other advantages come indirectly, such as, efficient resource utilization is a result of centralization not softwarization. Also, novel business opportunities through RAN-as-a-service (RaaS) is a possibility through cloud computing and virtualization which not necessarily need softwarization. Similarly, NFV is a complementary technology to SDN and does not depend on it [1]. Flexible innovation or fail-fast approach, as shown in Fig. 2, is the important feature which is not possible without SDN. In fail-fast model, ideas are tried through rapid deployment, discarded if failed, and scaled out quickly if taken off. Currently, any upgrade in mobile network takes a long time. With SDN, any mobile technology will have much smaller trial and deployment process.

The LTE architecture also distinguishes between user-plane and control-plane, where the former deals with data packet forwarding, while the latter focuses on signaling and management messages and operations using the same physical infrastructure. Both planes reside in the firmware of the system. This demarcation is much clearly designed into the mobile core network EPC (Evolved Packet Core), but the radio access part of LTE consists of only base-station node, eNodeB, performing data forwarding and control functions. The SDN-based core network requires transporting the control plane into software along with the control logic required for the data forwarding plane, e.g., routing rules, mobility anchoring, etc. We also remark that user plane as defined by LTE is not exactly the same as data forwarding plane of SDN and similarly both control planes also differ slightly [13].

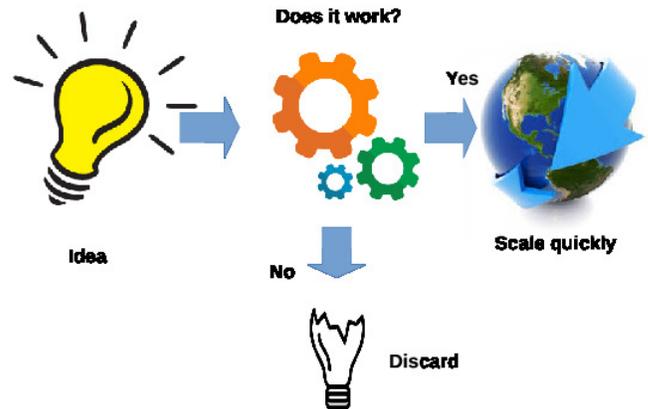


Fig. 2. The fail-fast model, major benefit of SDN technology.

B. NFV

In NFV (Network Function Virtualization), network functions are implemented in software components called virtual network functions (VNFs) [14]. VNFs are deployed on commodity servers or cloud infrastructure instead of dedicated hardware reducing CAPEX (CAPital EXpenditure) and optimizing resource provisioning [14]. The VNFs might not only be linked to a specific service and/or application but they can relate to specific user requirements, network policies and allow for service differentiation based on the user subscription SLAs (Service Level Agreements). The key motivation is that once implemented in software these VNFs are able to run in general-purpose hardware where they can be deployed, configured, scaled, migrated, and updated on demand. Operation of the VNFs is expected to take place in accordance to the ETSI Network Function Virtualization (NFV) management and organization (MANO) framework depicted in Fig. 3 [15].

The MANO framework defines three main logical components, namely the Virtualized Infrastructure Manager (VIM), the VNF Manager (VNFM) and the NFV Orchestrator (NFVO), interconnected over specific reference points. The VIM manages the infrastructure resources (e.g., compute, network, storage, memory) whereas the VNFM performs the life-cycle management of the individual VNFs (such as scaling, migrating, update/upgrade etc). The NFVO is responsible for both the resource/service orchestration of a Network Service (NS), which is composed of one or more VNFs (and its components) interconnected over virtual links. There are additional data repositories that may contain necessary information about NS, VNF, NFV and NFVI (NFV Infrastructure) that will enable the NFVO to perform its tasks. The MANO architecture also defines reference points for interfacing the MANO system with external entities like NFVI, OSS/BSS, VNFs and Element Managers (EM) for delivering a unified management and orchestration of a VNF system.

SDN is a complementary technology to NFV and the realization of NFV does not depend on SDN.

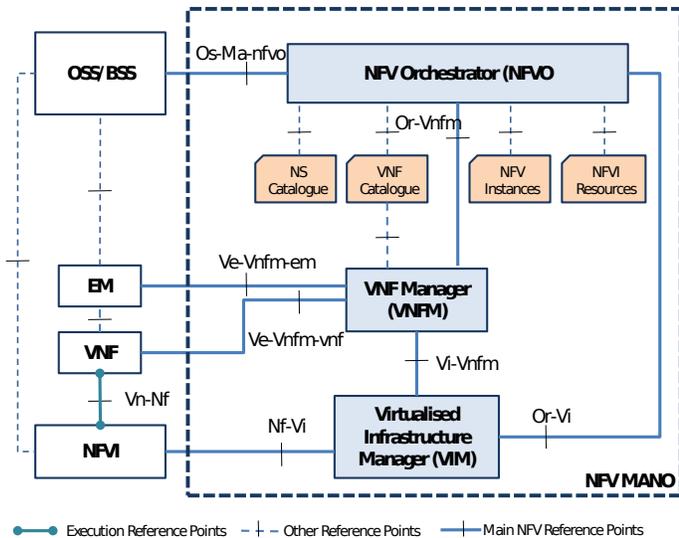


Fig. 3. ETSI NFV Management and Orchestration (MANO) Framework Overview.

C. LTE Architectures- current and future

The LTE architecture [16], shown in Fig. 4, also distinguishes user plane, dealing with the data packet forwarding, and control plane, focusing on signaling and management messages and operations, using the same physical infrastructure. EPC (Evolved Packet Core) has the designated control plane nodes MME (Mobility Management Entity), HSS (Home Subscriber Server), and PCRF (Policy Control and Charging Rules Function) [16]. Serving gateway (S-GW) and packet gateway (P-GW) are used for the data forwarding. From SDN's perspective, the gateways and the base-station also contain control plane functionality along with the designated control nodes of EPC. This is shown by color coding all EPC nodes and eNBs (eNodeB, base-station) partially with green color in Fig. 4 relating them to the control plane in Fig. 1. The base-stations eNBs and S/P-GWs are comprised of the data plane from SDN's point of view and thus shown with partial turquoise coloring in Fig. 4. A possible implementation of EPC control nodes in a SDN environment is as application modules so we color them partially orange. The colors are relating the nodes in Fig. 4 to the layers in SDN model of Fig. 1. As it is clear from Fig. 4, the traditional networks have control and forwarding planes integrated into all elements and a proper logical decomposition, while preserving the performance constraints over the inter-communication between controller and data plane elements, is the biggest challenge.

In this section, we also discuss some of the ideas proposed for future mobile network architecture, i.e., Beyond Cellular Green Generation (BCG2) [17] and Phantom cell concept [18]. These ideas have shown big promises in energy efficiency [19]. The logical decoupling of data transmissions and control signaling paradigm is one of the key directions being explored by GreenTouch³ under the project Beyond Cellular Green

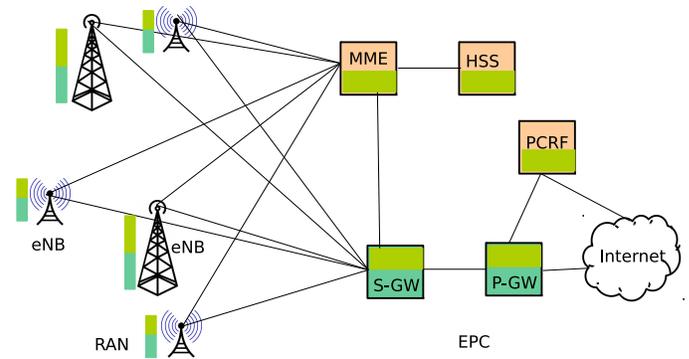


Fig. 4. LTE, RAN and core (EPC) network. The colors indicate the SDN functional decomposition: orange color shows application plane, green refers to control plane functions, and turquoise shows data plane. Solid lines show wired connections.

Generation (BCG2) [17]. In BCG2 architecture, the signaling nodes are responsible for the coverage and are usually assumed to deliver low rate services, such as, random access and paging, over long ranges; whereas the data nodes can be activated and deactivated depending on the traffic demand and it is designed for high rate and small ranges. The decoupling is logical in nature and a single location can host both types of nodes. A set of studies regarding the BCG2 architecture is performed under the EU FP7 IP project EARTH [19]. The study shows that up to 85-90% saving potential is possible with this revolutionary changed architecture compared to the current systems. Coverage is also separated from data processing in a CloudRAN architecture [5], [20], where a centralized BBU pool serves several RRH (Radio Remote Heads) in the area and not-in-service BBUs can be put in sleep mode to save energy.

Phantom cell concept [18] which was introduced for realizing true potential of dense deployment of small cells as suggested for LTE Release 12. In this idea, many small cells, called the phantom cells as they contain only LTE user plane, are overlaid with a normal macro cell which provides interference coordination. However, a macro cell consumes over 100 times more power than a pico/femto cell as described in the EU FP7 EARTH project [19] and keeping it on all the time would lead to severe power inefficiency. [13] proposed to change the macro base-station in Phantom cell architecture with a signaling only large coverage base-station which can also hold some control plane processing for SDN-based mobile network. The architecture can provide interference coordination service to the small base-stations in the area.

Although not yet explored, but the phantom cell or the data cell in BCG2 architecture can become a starting point to realize dump base-stations in a SDN-based mobile network architecture. The dump and small base-stations can provide data forwarding only and their control lies elsewhere, e.g., in the cloud or in the macro base-station. Stripping down signaling capability from dump base-stations opens up bigger challenges as how to determine the best small base-station to connect

³www.greentouch.org

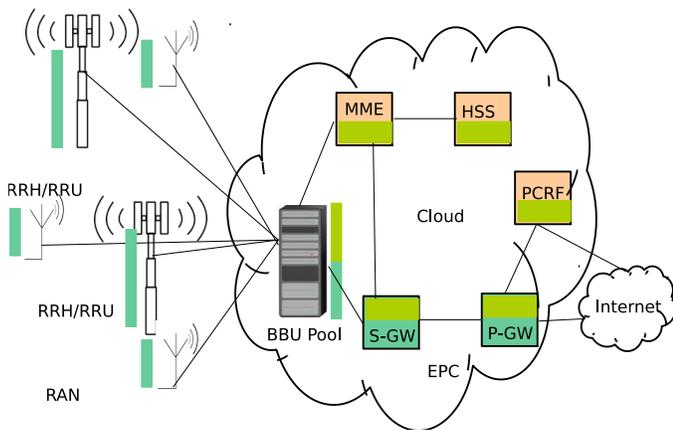


Fig. 5. The CloudRAN architecture, with BBU pool and EPC in the cloud. Colors indicate the SDN functional decomposition: orange shows application plane, green refers to control plane, and turquoise shows data plane. Solid lines show wired connections.

to UE. Moreover, a full SDN-based mobile network can also allow us to choose the base-stations to remain on for signaling instead of a predetermined signaling base-station in BCG2. In a CloudRAN type architecture, the controller can select a set of RRHs to provide signaling functions and the rest RRHs can be used on-demand.

D. CloudRAN

CloudRAN has its roots firmly planted by China Mobile⁴ and it can be viewed as a facilitator for introducing SDN technology. CloudRAN, which is now hitting its stride, centralizes all functionalities, control as well as data plane, into a centralized BBU (Base Band Unit) pool, or a data center/cloud, for easier management and coordination as shown in Fig. 5. In that respect, only antennas and some active RF components, i.e., RRU (Remote Radio Unit) or RRH (Remote Radio Heads), are left on the cell sites [5], [20]. Figure 4 is morphed into CloudRAN architecture by replacing eNB with RRHs/RRUs and moving the processing units, i.e., BBUs and the EPC to the cloud as shown in Fig. 5. CloudRAN also does not decompose control and forwarding layers as shown by the color coding in Fig. 5 to match with SDN layers of Fig. 1. Although, the RRHs/RRUs can be classified as the data forwarding RF devices without any control functionality.

CloudRAN is proposed as a mechanism to realize small cell deployment in LTE through proper coordination for interference management. LTE small cell, however, assumes distributed control with self-organizing (SON) capabilities. The connection from the cell site to the serving gateway (S-GW) in LTE could be through wired or wireless links (<http://scf.io/>). On the other hand, CloudRAN assumes a high bandwidth fiber link between RRU and the centralized data center, which is also the most prohibitively expensive aspect of this proposal.

Centralized processing in the cloud is not necessary for SDN and alternative proposals have been proposed to reduce

the load on the fronthaul by leaving some processing blocks at the base-station [8]. Also, some proposals have left the part of control plane at the base-station in order to perform delay-critical control [3]. These proposals require connections between control elements in the cloud and at the base-stations.

The CloudRAN architecture has the advantage of lower cost, energy efficiency, smaller footprint, and lower maintenance cost [21]. Alcatel Lucent's Light Radio, Ericsson's Antenna Integrated Radio and Nokia Siemens Networks' Liquid Radio are among the earlier CloudRAN products.

The CloudRAN architecture, with its low cost and flexibility of operation, also opens up new possibilities. The architecture is shown to be extremely useful for very high-speed train communications⁵ [22]. Due to low cost, several remote radio heads (RRHs) can be deployed along the track in the framework presented in [22]. Centralized processing allows to use virtualization along with joint beamforming techniques to yield up to 100Mbps at the speed of 450kmph [22].

III. MAJOR OBSTACLES AND ISSUES

We identify six major issues hindering the realization of SDN technology. In no particular order, these issues are

- 1) Fronthaul
- 2) Latency of general purpose platforms
- 3) Backward compatibility
- 4) Disruptive deployment
- 5) SDN specific security vulnerabilities
- 6) Clear and compelling business case

In this section these issues are discussed in detail. We have also looked into the state of the solutions proposed for them. This is one of the most rapidly changing field in digital technology and last few years have seen massive progress in the area. We cannot claim to have looked at each and every publication under the relevant themes but we have certainly covered the major solutions and ideas presented in the literature including industry white papers and technical reports as well as academic research papers.

A. Fronthaul

In centralized RAN architectures, such as, CloudRAN, connections from remote radio units (RRU) to the centralized baseband processing units (BBU) form fronthaul as opposed to backhaul, which connects the base-stations to the core network. RRUs are the wireless transceivers mainly consisting of the antenna heads.

Traditionally, data from the processing cabinets of base-stations, in the form of digitized baseband signal, is carried to the antenna heads located at the rooftop or top of the masts by CPRI (Common Public Radio Interface) protocol over fiber optic connections. CPRI over fiber allowed the antennas to be located away from the base-station cabinets as the legacy co-axial cables caused a lot of power loss and were also expensive and bulky [23] and can only be used for short distances. CPRI over fiber can connect RRUs miles away from

⁴<http://labs.chinamobile.com/cran>

⁵<https://www.globalrailnews.com/2017/12/04/samsung-successfully-tests-5g-on-high-speed-train-in-japan/>

BBUs with minimal losses, however, the limiting factor is the latency constraint to maintain channel state information and synchronization between BBU and RRU [23].

The data to and from the centralized BBUs is digital version of high frequency baseband signal. As an example, a LTE baseband signal with 20MHz channel bandwidth needs to be sampled 30.72 million times per second for proper digitization resulting in data rate of approximately 2.5Gbps if data stream is duplicated for 2×2 MIMO [23]. CPRI rate ranges from 614Mbps to 12Gbps [23]. Such high data rates need fiber optic cables, although, microwave can be used but for lower rate CPRI only. Interestingly, CPRI is very inefficient from data transport point of view, e.g., a 20MHz LTE channel can carry up to 150Mbps in downlink but requires CPRI rate of 2.5Gbps [23]. CPRI was not designed to carry data over very long distances and it was an internal base-station interface to connect the antennas on the rooftop to the base-station cabinet. There is clearly room for introducing efficiency in data transport in fronthaul. Advanced techniques exploiting correlation among RRUs, such as, data compression, quantization, precoding, etc., can be used as well to reduce the data rate for the constraint fronthaul [24].

Other important issues for fronthaul are latency and jitter. Real-time communication, such as, VoLTE (voice over LTE), IoT, tactile applications, etc., and the control signaling between RRH and BBU to maintain channel state information require low delays. To keep one way delay under 75 microseconds, the length of the fiber link between RRH and BBU should not be more than 15Km [23]. According to [25], the delay of CPRI should remain under $100\mu s$ and jitter should be under $65ns$. Jitter in the synchronization signal between BBU and RRU will induce phase noise which in turn will lead to degradation in the transmitted modulated signal [26]. Heterogeneous CloudRAN [27] tried to solve the latency issue by introducing a different high power node (HPN) than RRUs which is responsible for control signaling in a geographical area. However, the latency issue with data transmissions, such as, VoLTE, IoT, and other time-critical applications remains.

In fog/edge computing architectures [28], the nodes are widely spread, away from the cloud or data center, where the users are. They can provide some localized services and can also store some information. Such architecture also reduces the load on fronthaul with some improvement in latency requirements.

An effort to move away from semi-proprietary CPRI has also been made by ETSI through their Open Radio Equipment Interface (ORI) specifications⁶. The ORI interface was built on top of CPRI and it supports line bit rates up to 10.14Gbps. The objective was to make the interface fully inter-operable and improving the data transport efficiency was not the main goal.

Among the various SDN-based mobile network architectures proposed in the literature, some have tried to touch upon fronthaul but not really looking into resolving the transport efficiency or latency issues. As shown in this section, the fronthaul is a big research problem in itself and we should

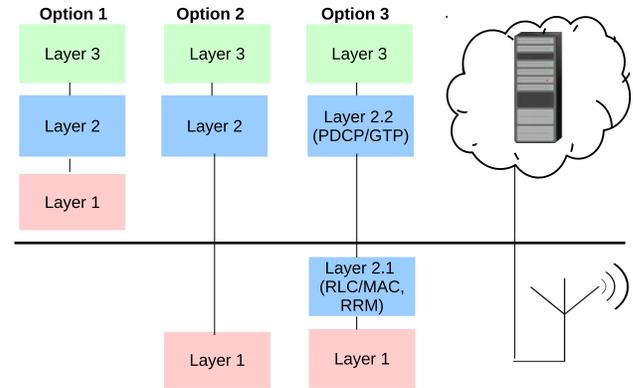


Fig. 6. Fronthaul functional split options (courtesy [20]).

not expect that research efforts towards new architectural organizations for SDN-based mobile networks would also be able to resolve fronthaul issues as well. But since fronthaul is a major element of SDN-based mobile network architectures, it cannot be left behind. Interestingly, we found only two papers [3], [8] which addressed the issue in their architectural design. Details about the design is discussed later in Section III-A1. The only other papers which discussed fronthaul in their architecture is CONCERT [11] and SDCN [9] but the major relevant issues were not their focus. CONCERT and SDCN proposed that fronthaul network can also be a SDN network with optical switches controlled and virtualized by SDN controller. CONCERT claimed virtualization efficiency with SDN fronthaul than using packet switched Ethernet [11].

1) *Functional split between BBU and RRU*: Different centralization options are also considered to reduce the data rate for fronthaul [20], [24] as shown in Fig. 6. When processing for all layers is performed in the centralized BBUs, it is called full centralization as shown by option 1 in Fig. 6. In this case, we have very high data rate waveform to transmit to and from the RRUs. Partial centralization options are also available where layer 1 or PHY was kept at the RRUs as it reduces a lot of fronthaul overhead [24]. This is shown as option 2 in Fig. 6.

Moreover, it sounds desirable to have all programmable OSI layers but there is a strong consensus among network owners and operators not to allow low level network programming by third parties [1] due to network security and management issues. Also, keeping layer 1 on dedicated hardware instead of general purpose processors can also improve the latency issue which is very critical for some IoT, tactical and real time applications. On the other hand, efficient implementation of some advanced features, such as, CoMP (Coordinated Multiple Point) and cooperative processing for massive MIMO will not be possible if layer 1 remains in the hardware. Interaction between layer 1 and 2 will also be complex and one way latency between RRU and BBU should be of the order of 100-150 microseconds [20]. Moreover, the fronthaul should also cater for some control signaling regarding carrier frequency,

⁶<http://www.etsi.org/technologies-clusters/technologies/ori>

transmitter power, etc.

Another option for partial centralization is to split MAC function between the cloud and the antenna site. The one-way latency constraint can be more than a millisecond. This alternative is shown by option 3 in Fig. 6. In this architecture, HARQ of MAC is kept at RRU whereas the MAC scheduler resides in the central location [20]. In another option, only layer 3 resides in the cloud and PHY and MAC are kept at the antenna site [20]. This also reduces a lot of fronthaul overhead and handover overhead is also greatly reduced in case of mobility within the serving area of layer 3 as all cells can be aggregated and provided a single S1 view towards core network [20].

Chinamobile presented NGFI (Next Generation Fronthaul Interface) [29] which decouples the dependency of CPRI on number of antenna elements by putting all antenna related functionality, such as, downlink antenna mapping, fast Fourier transform [FFT], channel estimation, and equalization in RRU. It is shown that an LTE fronthaul bandwidth may decrease on the order of 100Mbps no matter how many antennas are used [29]. Currently, NGFI is being standardized under IEEE 1914.1 project.

Similar ideas of doing cell based processing, such as, FFT/IFFT, subcarrier mapping, etc., of the receiving waveform in RRUs were discussed in [30]. According to the authors, this split will half the fronthaul capacity requirement and further reduction would be possible by moving demodulation and decoding blocks to the RRU. They also suggested to put full PHY for transmitting signal in the RRU to reduce fronthaul capacity.

Similar ideas of performing FFT/IFFT near the users is discussed and evaluated in [31], where LTE and DOCSIS (Data Over Cable Service Interface Specification) share the fibre fronthaul and remote FFT node. The remote FFT/IFFT reduces the fronthaul bit rate to approximately $1/30^{\text{th}}$ of the conventional baseband approach. The system also used frequency domain I/Q symbols instead of conventional time domain I/Q symbols to reduce the required transport bit rate and used caching of the repetitive QAM symbols to further reduce the downstream bit rate requirement. Evaluations show 6.5% (for full load) to 41% (for 10% load) savings for LTE traffic.

SoftAir [8] is one of the two research papers on SDN-based mobile network architecture which also addressed fronthaul transport efficiency issue. In SoftAir modem (modulation and demodulation) is recommended to be kept at RRU in order to reduce the data rate of CPRI supported fronthaul. Such a split is also evaluated through hardware testbed experimentations in [32] and latency and jitter are found to be within acceptable bound. The fronthaul is using Ethernet technology, although, the length of the fronthaul is not reported in [32].

SoftRAN [3] provides another type of functional split, in terms of control functions, between center and remote site. They also realized that the biggest challenge in realizing SDN-based mobile network is the inherent delay between any centralized controller and the radio elements. The delay could typically be 5-10ms. SoftRAN splits the control plane functionalities and leaves the time-critical local control related

to channel state, such as, downlink resource allocation (with transmit power set by the centralized controller), at the individual radio element.

A detailed evaluation of different functional splits is provided in [33] in terms of multiplexing gains. The obvious conclusion is the best multiplexing gain for full centralization solution with all processing in the BBU pool. This is specially true for high load. When the traffic is low, the fronthaul requirements are shown to be relaxed with only higher layers in the BBU pool and lower layer processing near the RRU [33]. The impact of functional split of fronthaul on its performance is studied with packetization and scheduling in [34] and [35] and it is shown that there is a strong interplay between packetization overhead and latency that changes the fronthaul performance. For each split, the analysis finds the maximum number of supported RRUs for the best packetization method satisfying fronthaul capacity and latency constraints [34]. This work exploits packetization latency and options of putting more processing at RRU to support more RRUs in the system for specific fronthaul constraints. The follow-up paper [35] also includes different scheduling mechanisms in the study to improve the number of supported RRUs.

A summary of split functionality proposals are presented in the Table I, where it is clear that transport efficiency can be improved with proper distributions of functions between BBUs and RRUs. Although, no study has been done to compare the different proposals and we have used the generic terms 'Heavy' and 'Light' to show that all the split function proposals have lesser load than CloudRAN. It is not known how they compete with each other. In Table I, RH stands for radio heads or transceivers and RRC is Radio Resource Control. Also, SoftAir [8] and NGFI [29] did not mention any delay figure explicitly and we have taken $\leq 75\mu\text{s}$ as it is considered as a standard for centralized RAN [23]. The split architectures of SoftAir and NGFI focus on only reducing the load of fronthaul. In SoftAir, delay is considered as a constraint while forming clusters of RRUs.

2) *Ethernet vs. CPRI/ORI*: The use of Ethernet for fronthaul is also been looked at with various options of placing CPRI data or RF data directly into Ethernet frames [26]. Ethernet is a mature technology with standardized OAM (Operations, Administration, and Management) capabilities but by putting the data in Ethernet frame, the synchronization information could be lost [26]. Synchronous Ethernet, where clock information is extracted from the received data, and Precision Time Protocol, where time-stamped packets are exchanged, can be used to achieve synchronization in fronthaul Ethernet operation. Moreover, the traditional Ethernet switching cannot fulfill the strict timing requirements of fronthaul traffic and substantial optimization to the switching fabric is required [36].

IEEE 802.1CM Time-Sensitive Network task group is looking into the set of standards focusing on time-synchronized low latency high bandwidth services over Ethernet networks for fronthaul transport network. In a centralized RAN, strict requirements are defined between the RRU and the BBU, such

TABLE I
SUMMARY OF SPLIT FUNCTIONS.

Proposals	Cloud	RRU	Fronthaul load	Latency
CloudRAN [20]	All layers	RH	Heavy	$\leq 75\mu s$
NGFI [29]	Rest	Antenna related	Light	$\leq 75\mu$
Partial/layer1 [20]	Layer 2+	PHY/RH	Light	100-150 μs
Split MAC [20]	RLC/Layer3+	HARQ/PHY/RH	Light	1ms
SoftAir [8]	PHY(ex. modulation)/Layer2+	Mod./Demod./RH	Light	$\leq 75\mu s$
SoftRAN [3]	All layers (ex. RRC)	RRC/RH	Heavy	5-10ms

as, less than 0.1ms one-way latency at 614.4 to 24330.24Mbps [37]. The standardization activity looks into techniques to reduce latency and jitter, such as prioritizing the time-sensitive frames, time-synchronization, end-to-end resource reservation, etc.

In LTE small cell network, a massive coverage of backhaul is needed instead of fronthaul in centralized RAN and it is still considered one of the major issues [38]. Both wired and wireless backhuls are being used according to the operating conditions. However, with excessive cost of laying fibre for dense networks and technological and regulatory issues with millimeter-wave backhuls, it is still an open problem in small cell network similar to fronthaul in CloudRAN type access networks. Smart backhaul/fronthaul solutions where interworking is exploited and joint optimization is performed with the access network are also recently explored [38].

A futuristic idea of reconfigurable backhaul/fronthaul, namely *Xhaul*, is presented in [36]. Here functional split between remote radio site and centralized processors depends on the network conditions and operational scenarios. A software defined xhaul employs different protocols, such as, CPRI/ORI, Ethernet over heterogeneous physical links, e.g., fibre, microwave, etc., with appropriate resource allocation to address latency, jitter, and throughput constraints for each type of data.

Another proposal for software defined backhaul is presented in [39]. Here a SDN orchestrator is responsible to manage multiple smart-gateways, a new element suggested in the scheme, in order to provide demand-based uplink capacity to the connected small cells. The smart-gateways are connected to the S/P-GW of multiple operators and allow resource sharing through SDN orchestrator.

A significant advantage, in the case of backhaul, is the use of Ethernet instead of CPRI which has much better data transport efficiency. Moreover, commodity, or low-cost industry standard equipment can be used lowering the overall CAPEX where fiber is unavailable.

The most important development in this regard is IEEE 1914 working group (NGFI). The working group has two active projects: IEEE 1914.1 is studying NGFI of Chinamobile and IEEE 1914.3 is looking at encapsulation of digitized radio signal (I/Q samples) into Ethernet frames for fronthaul transmission. Also called as 'Radio over Ethernet', IEEE 1914.3 consolidates the efforts of a previous working group IEEE 1904. The Ethernet frame structure and MAC remains unchanged under this project and the focus is on encapsulation

and mapping of digitized payload as well as control and management messages.

An mmWave based fronthaul using Ethernet is presented in [40] for heterogeneous CloudRANs. The system contains small range RRUs and macro RRUs. An Ethernet switch is added with macro RRU to direct the traffic to the UE or to a connected small RRU. In this case, the small RRU does not need to decapsulate the Ethernet frame as it is done at macro RRU. Preliminary results show that the latency is within the limits for fronthaul operation.

In a recent study, an Ethernet based fronthaul is tested using hardware implementation [32] and latency and jitter are found to be within acceptable bound. The length of the fronthaul is not reported in [32]. Another detailed study in [25] claimed that the encapsulation delay would compromise only few kms in fronthaul and it is also possible to reduce jitter, even eliminate it, with some scheduling techniques.

3) *Hardware for fronthaul*: The cheapest option for fronthaul is considered to be WDM (Wave Division Multiplexing) over dark fiber [1], [23] where the fiber is leased per mile not by the bandwidth. Places where dark fiber is not available, duct sharing or Physical Infrastructure Access (PIA) can also be considered as a solution [1], where cables can be placed in access ducts using a license agreement. It is very much clear that countries with fiber rich city infrastructure will benefit far more from these savings than those with low fiber availability.

Although, fiber deployment is gaining momentum around the world but as shown in Fig. 7⁷, the growth is mainly in China and Hong Kong with Latin America, Africa, Middle East, and South Asia lagging far behind. According to the new statistics released in June 2013, fiber-optic links are used for 15.8 percent of broadband connections in the OECD area with Japan and South Korea leading the list with 68.5% and 62.8%⁸. Organization for Economic Cooperation and Development (OECD) monitors economic trends in North America, Europe, Australia, Japan, South Korea, Turkey, and some other developed countries and regions. It is very clear that fiber-rich countries will gain the real benefit of the SDN-based and centralized RAN but the region with low economic incentive to deploy optical network will be left behind. It is a serious concern that the advancement in telecommunication technology will further widen the digital divide.

⁷Source: <http://wireandcablenews.crugroup.com/wireandcablenews/insights/free/2015/10/4522446/>

⁸<https://www.cnet.com/uk/news/fast-fiber-optic-broadband-spreads-across-developed-world/>

Optical cable CAGRs: 2005-2015 = 17%; 2009-2014 = 12%

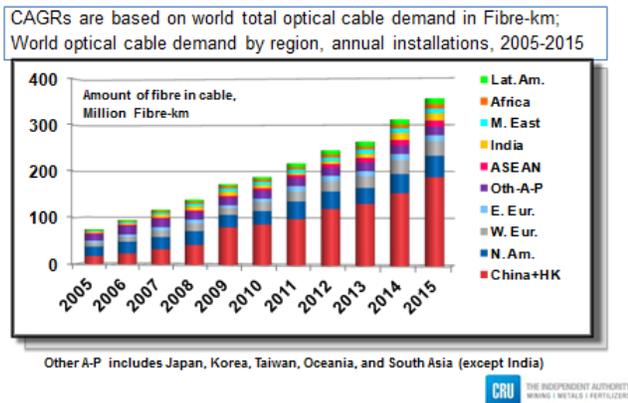


Fig. 7. Compound annual growth rate (CAGR) for optical fiber around the world.

Other technologies, such as co-axial or twisted pair copper cables, microwave, mmWave, are also explored to provide fronthaul transport network. In [40], a mmWave fronthaul is used between small RRUs and macro RRUs. Two different scenarios are tested in preliminary study, i.e., sub-3 GHz and Eband (7176, 8186, and 9295 GHz). Eband is showed to have approximately 11% reduced RRUs for the same traffic volume due to availability of broader spectrum.

4) *Software Defined Optical Networks*: Another important issue in relevance to optical fronthaul is the developments in software defined optical networks so a complete SDN-based networking solution can be realized. There has been a number of developments regarding software defined optical infrastructure, i.e., transceivers, switches, medium access, performance monitoring, etc. [41]. The survey in [41] provides a comprehensive account of research in control layer framework, access control, security, QoS monitoring, virtualization mechanisms, etc. Despite the developments, important challenges are still remaining specially in integration of products from different vendors and heterogeneous physical layer technologies, e.g., radio and fibre [41]. Security vulnerabilities and scalability issues in centralization of control layer are also needed be explored in further details besides other technical challenges [41].

B. Latency of General Purpose Platforms (GPP)

The history of base-station softwarization efforts go back beyond SDN. Potential cost savings and flexibility in implementing and changing protocols were the main drivers behind the efforts. Within SDN ecosystem, base-stations are integral which can have programmable functionalities, i.e., protocols and transmission parameters, either centralized or distributed, implemented on a general purpose computing platform and generic radio heads. By softwarization, the BS and core network functions can be virtualized thus leveraging on the advantages offered by virtualization such as flexible and agile deployment and management of network functions and agile

composition of new network functions. In this context, the performance of GPP became a genuine concern.

GPPs were not designed for real-time signal processing where the delays should be of the order of tens of microseconds [42]. Although, GPPs are also improving their speed by finding and eliminating the bottlenecks, but it is yet to be seen if the 1000x capacity and 100x data rate promise of 5G can be supported by a base-station in GPP. Moreover, the promise of saving CAPEX in SDN through GPP may not be as huge as the predictions, as the cost of specialized ASIC hardware is also coming down with time [1]. The actual cost saving should also consider this downward trend.

There has been significant efforts in implementing LTE eNodeB in software with promising performance results [30], [42], [43]. The latency values were shown to be decreasing with processor speed. Special techniques are used in the software to improve latency using parallelism, replacing computations with lookup tables, kernel patches for time guarantees, etc. So far, data rates up to 43.8Mbps on a 20MHz channel are shown to be supported by a LTE base-station in software. It is still a long way to go as far as 5G goals are concerned. In this section, we will discuss prominent efforts in this respect and assess the gaps in the current state of the art.

[42] presents, Sora, a fully programmable software radio platform on GPP from Microsoft. It was the first effort to program LTE on a commodity general purpose computer. Sora implemented LTE uplink Rx PHY that supports up to 43.8Mbps data rate on a 20MHz channel. Sora, however, uses some hardware components for high throughput and low latency data transfer between radio front-end and host memories [42]. They bridged the synchronous transfer of high data rate RF digitized signal and asynchronous processing of a GPP through FPGA FIFOs (buffers) and on-board memory with the support of PCIe bus standard for high throughput and low latency. They have also provided a separate path for low latency control data. They have used new software techniques, such as, use of lookup tables instead of computations, use of special kernel service for core dedication in multi-core processor environment, etc., to improve the latency for real-time PHY processing. Sora does provide a good starting point but it does not include softwarization of downlink LTE which is also important and requires more complex FFT processing. Also, realizing the projected data rates for 5G, i.e., in the order of tens of Gbps, it is still a long way to go. Moreover, the real-time processing of CoMP (Coordinated Multi Point) interference management scheme to reap the benefits of centralization would be a big challenge as well on software platforms.

Virtual WiMax base-station pool is developed and evaluated in [44] which can meet system requirements including synchronization, latency and jitter. The implementation uses a FPGA based adapter board between radio head and virtual base-station pool implemented on GPP and can support up to 10MHz signal bandwidth.

In another effort, reported in [43], LTE eNodeB prototype is implemented using single core GPP. The prototype was tested for latency in receiving uplink signals and transmitting

downlink signals and the results did not meet LTE standards [43]. The authors, though, expressed the hope that a multi-core platform would bring the delays within LTE specifications.

OpenRadio [6] explored the softwarization of PHY layer of a base-station. OpenRadio uses hardware accelerators for highly computational blocks, such as convolutional coding, Turbo coding, etc. and separates the decision rules from the protocols to be put in the software in order to meet strict deadlines. OpenRadio is also concerned about the software based decision plane not to become very heavy as it would incur prohibitive inter-core communication overheads implying that decision/processing plane separation might not be the best design choice [6].

A notable effort in this regard is from EURECOM called OpenAir Interface⁹. OpenAir Interface (OAI) is an open source implementation of LTE Release 8.6 with parts of Release 10. The maximum expected throughput is reported to be 36Mbps on a 10Mhz downlink channel. Experiments on OpenAir Interface show that the BBU processing delay gets lower with processor speed [30]. A 3GHz single core processor is shown to comply with the HARQ delay requirements of uplink and downlink processing [30]. However, an independent performance study [45] shows that the downlink execution time increases with the load and could reach over 200ms for PDCP payload size of 1500bytes. OpenAir Interface is an ongoing project, it is yet to see how the base-station software fare for 5G throughputs and data speeds.

In a recent effort [46], both EPC and eNodeB are deployed in one mini Intel commercial PC (i7-5557U). All protocol layers of eNodeB are implemented on the GPP platform with X86 CPU instruction set. The communication between RF block and BBU block is through high speed USB 3.0 interface. The software based eNodeB has USRP B210 software defined radio for transmission and reception. Currently, this implementation can support stable real-time signal transmissions with a bandwidth of 10MHz and a stable data rate of 32Mbps to multiple commercial mobile terminals. The authors also commented about the power consumption of their LTE implementation which is comparatively more than the dedicated hardware.

srsLTE is an open source implementation of LTE PHY only on GPP [47] with extensive use of data and instruction parallelism and use of different kernels of SIMD (Single Instruction Multiple Data) code for different architectures. One processing core showed to have met the delay deadlines of LTE but multiple cores were recommended to be used for processing of the rest of the layers. A commercial version of srsLTE is also available.

A LTE eNodeB in software is commercialized by Amarisoft¹⁰, where they claim to have a fully LTE release 13 complaint software which can be used with USRP¹¹ radio front ends. The software seems to be meeting all the timing requirements but no documentation is available to explain how did they achieve it. The company also claims that 500 UEs

can be supported by a single PC but the maximum achievable throughput is not mentioned.

A proprietary virtualized access (vAccess) development platform from Freescale [48] is created to facilitate the development of SDN/NFV products. It is built over OpenStack cloud computing platform with hardware accelerators for L1 real-time processing and Linux patch to support bounded guarantees for application start time. A LTE smart base-station built over the vAccess platform is also presented in the market¹².

In 2016, NEC announced NFV CloudRAN¹³ software which provides flexible functional split between centralized digital unit and radio unit of a CloudRAN. The software application runs on GPP platforms but requires hardware accelerator for L1 processing. Similarly, ASCOS's virtual base-station¹⁴, which came out in 2017, also uses hardware accelerators for L1 signal processing.

There are other efforts to implement LTE in open-source software for research testbeds. OpenLTE¹⁵ and gr-lte¹⁶ are notable examples but they are not for real-time transmission and reception of LTE signals. gr-lte is only a receiver which is based on GNU radio.

A reconfigurable and programmable SDR platform is proposed in [49] which uses hybrid FPGA to meet all timing requirements instead of GPP. However, it is noted in the paper that reconfiguration of hybrid FPGA is a very slow process and it is not straightforward to do it on the fly.

1) Lessons Learned: A summary of notable software implementation of LTE eNodeB is provided in Table II, where it is very clear to see the current state of the art. All of the LTE eNodeB software implementations, open and closed, are relatively recent and unless independent and comprehensive trials are done, it would not be clear that they have resolved the latency issue of GPP or not. Although, they definitely show a lot of promise and potential of the technology. We expect that like every technology, GPPs will also improve and multi-core processors will resolve the bottlenecks slowing down their speed.

Moreover, as we have observed while discussing fronthaul, that leaving PHY at the RRUs could save the capacity of fronthaul and it is very difficult to imagine that network operators would open such low-layer networking to the third party users. We really do not loose much if PHY, which is also very computationally intensive and latency critical specially in 5G, remains in special hardware at the base-stations or RRUs, at least in the initial phase of SDN roll-out.

C. Backward Compatibility

Not long ago, the carrier networks have deployed their 4G LTE systems. The systems are performing very well and

¹²<http://sooktha.com/sbs.html>

¹³<https://www.sdxcentral.com/products/nfv-c-ran/>

¹⁴<https://www.sdxcentral.com/products/asocs-virtual-base-station-vbs/>

¹⁵<https://sourceforge.net/projects/openlte/>

¹⁶<https://github.com/kit-cel/gr-lte>

⁹<http://www.openairinterface.org/>

¹⁰<http://www.amarisoft.com/?p=amarilte>

¹¹<https://www.ettus.com/product/category/USRP-Networked-Series>

TABLE II
SUMMARY OF NOTABLE LTE ENODEB SOFTWARE IMPLEMENTATIONS.

Proposals	In software	Dedicated hardware used or not	Performance
Sora [42]	Uplink Rx PHY	between high data-rate RF front-end and host memories	supports upto 43.8Mbps on 20MHz channel
[43]	Release 8	Same as Sora	Single core processing did not meet timing requirements
OpenRadio [6]	PHY	hardware accelerators for coding blocks	No reported results for LTE
OpenAir [30]	Release 8.6 (part of 10)	None	36Mbps on 10MHz downlink channel
[46]	LTE	None	supports upto 32Mbps on 10MHz
srsLTE [47]	PHY	None	Single core met the timing requirements
Amarisoft (Commercial) ¹⁷	Release 13	None	supports 500UEs per PC
NEC NFV C-RAN ¹⁸	All layers	hardware accelerator	Carrier-grade C-RAN
ASCOS vBase-station ¹⁹	All layers	hardware accelerator	Indoor base-station

getting the most out of them by enhancing their capacity and support for number of users is a straightforward business opportunity. SDN technology comes with big promises but we cannot expect that the carrier operators and owners will decommission the recently-erected network. Most probably, 4G-LTE equipment will remain in function for some number of coming years. In view of this, any new technology should have the essential feature of co-existing and working with the legacy systems. In this section, we will review the work which has been done on the compatibility issues of SDN-based mobile networks and legacy systems. In legacy systems, co-existence with 4G equipment is the most important.

The recently approved 3GPP non-standalone 5G standard²⁰ talks about the scenarios where LTE's eNB and 5G NR's (New Radio) gNB (next generation NB) co-exist with EPC and/or NGC (Next Generation Core). The coexistence is possible through evolved base-station eLTE eNB which can connect to EPC as well as NGC and new interfaces, such as, Xx (between LTE eNB, connected to EPC, and gNB) or Xn (between eLTE eNB and gNB with either one or both are connected to NGC) [50]. NGC is a step towards a fully softwarized core by complete separation of control and user plane and modularizing the management functions. The non-standalone 5G allows an intermediate phase with backward compatibility to LTE.

Hybrid SDN network where traditional or legacy nodes co-exist with SDN-enabled nodes have been closely investigated in case of wired and fixed networks and the state of the art is way ahead of respective mobile network research. Firstly, an important point to note here is that OpenFlow protocol, which enables the switches to act as dump data pipes by following a flow table provided by the controller, can also be implemented on commercial Ethernet switches and routers [51] making them SDN-enabled. The simplest approach could be the dual stack approach where all network nodes have SDN interface along with an interface for normal processing [51]. This approach, however, requires substantial effort to

implement the software on each and every switch of the network.

Moreover, there are a number of proposals which has shown different mechanisms to successfully operate hybrid SDN networks with some SDN/OpenFlow switches and many non-SDN or legacy switches [52]–[55]. The performance of hybrid SDN networks show real promise and the architecture can actually provide a long-term solution rather than a quick fix for the transitional period. In [52] a small number of SDN switches, strategically placed, are shown to be sufficient for the network to reap the benefits of SDN and behave like a full SDN network by enforcing SDN policies, e.g., access control. In [54] and [55], the legacy switches are manipulated by special messages to forward the packet over the desired route. 10% penetration of SDN switches are shown to be sufficient to implement failure detection and route recovery mechanism through SDN controller in [53]. These approaches for managing hybrid networks, however, incur significant management complexity, as they control legacy and SDN switches via different mechanisms [54]. Moreover, the existing techniques on hybrid SDN networks successfully show implementation of SDN policies of routing, access control, flow isolation, failure recovery, etc., but it is not possible to efficiently virtualize the legacy resources via SDN controller without using additional virtualization mechanisms such as hypervisors etc.

All the above discussed schemes are designed for fixed networks and not for radio access network where the SDN-based RRUs will need to co-exist with eNodeBs, 2/3G base-stations, and RRUs of CloudRAN. RRU of CloudRAN may not exactly be the same as RRU of SDN-based mobile network. In the context of mobile networks, the research is way behind than the fixed wired networks and there are only very few attempts to propose solutions for the backward compatibility issue.

Mobileflow [56] is one of the earliest solutions to resolve the compatibility issue between SDN-based networks and legacy systems using virtualization techniques. It defines a mobile flow controller which can be used to implement the LTE control plane via network applications. The network

²⁰http://www.3gpp.org/news-events/3gpp-news/1929-nsa_nr_5g

applications provide all LTE control plane functionalities, e.g., control for S-GW (Serving Gateway) and P-GW (PDN or Packet data network Gateway) etc. Forwarding plane is defined through mobileflow forwarding engine which work similar to OpenFlow switches by following the routing rules given by the mobileflow controller. They also have radio interfaces to connect to UEs. The authors claim that operators can use this architecture in one part of the network to be used with the legacy equipment and then move to a more flat control model, without LTE control plane entities, in future by changing the virtual machines. They have validated the architecture through a prototype. However, the most important part of mobile network, i.e., eNodeB is not decomposed into the control and data plane and it is not shown how they will be mapped into the mobileflow forwarding engine and associated radio interface. In the prototype, an eNodeB was remained as it is and the mobileflow forwarding engine was used as the first node after eNodeB to direct the traffic. Although, not explicitly mentioned, mobileflow is in fact a SDN-based core network alternative to be used in place of 3G/4G core networks with option of moving towards a full SDN network in future through software update on the controller. The details about how to implement EPC through OpenFlow switches are also mentioned in [57].

In optical fiber domain, [58] showed SDN-controlled optical topology-reconfigurable mobile fronthaul achieving 10Gb/s peak rates with $<7\mu\text{s}$ back-to-back transmission latency and 29.6dB total power budget while maintaining backward compatibility with legacy fiber deployment.

A proxy based solution for backward compatibility of SDN-based core network is proposed in Softcell [59]. Each base-station has a proxy serving as GTP (Gateway Tunnel Protocol) end-point. Softcell implements LTE core network using commodity switches which then carry normal IP traffic between base-stations and Internet. Softcell, however, does not talk about softwarization of the base-station.

Heterogeneous CRAN [27] also claims backward compatibility with legacy cellular systems and other CloudRANs as along with RRUs and centralized BBU pool, it also introduces some high power nodes or base-stations providing control signaling. These high power base-stations are similar to macro base-stations. According to the authors of [27], these nodes are critical to guarantee backward compatibility where multiple heterogeneous radio networks converge and this architecture can take advantages of both CloudRAN and cellular networks.

Similar ideas were presented in [60] under the title of Fusion Net, where macro cells comprise of host layer providing anchoring points to the users. However, the traffic is dynamically transmitted and aggregated from a boosting layer consisting of Wireless LAN (WLAN) access points, small cells, RRUs, etc. The paper also discussed SDN and NFV as emerging technologies but did not bring them in the context of Fusion Net.

In [61], the discussion about backward compatibility leads to the mechanisms where SDN-based network (RAN and core) can be configured to be used as LTE 4G network. This is beneficial if 4G UE (User's Equipment) is being served by the network. However, the UE should have the appropriate appli-

cations to instantiate control and data transmissions. Moreover, the system still does not show the co-existence capability with 4G system where seamless handover and resource management could be possible between the two systems. According to [62], the main property for a SDN enabled backward compatible network is to implement standardized interfaces for interworking with legacy networks.

Another effort of addressing backward compatibility issue is presented in [63] where a 4G mobile communication system is shown to be used as a SDN-based mobile network via the separation of Edge Control Plane (ECP) and edge switches from the 4G core network. ECP contains all the intelligence and LTE control plane, i.e., MME, HSS, PCRF, etc., where as the 4G network only performs packet transmission. The access nodes, i.e., eNodeB, remain unchanged. [63] also talks about serving 3G and WiFi nodes through similar SDN-fixed 4G network. This scheme is a good proposal for first step towards the full SDN implementation but it requires substantial management rework in 4G networks. It is yet to see if it will make a good business case to revamp the control plane of already erected 4G network or will it be easier to install the new 5G systems, already arriving in the market, with similar SDN fixes as in [63]. The proposed SDN-based mobile network architecture in [64] maintains the 3GPP interfaces so the architecture can also be used with legacy network. The architecture moves EPC to cloud but leaves eNodeB unchanged.

1) Lessons Learned: In order to summarize the issue and state of the art of the relevant solutions, we believe that interoperability or backward compatibility may take different definitions but coexistence with 4G or legacy networks should include the following scenarios:

- 1) A 4G or legacy UE in the coverage of SDN-based mobile network. [61] tried to look at this type of scenarios where a SDN-based network can be programmed to serve a legacy/4G UE, although, the UE is required to have appropriate application. This scenario is, however, not really interesting from practical point of view since the upgrades in mobile handsets are much more rapid than the enhancements in the access network.
- 2) A legacy base-station or eNodeB is attached to a SDN-based core network. As shown in [61], the SDN-based network can be programmed to serve as a 4G network. In Softcell [59], the base-stations are assigned proxies which can act as GTP end-points to communicate with SDN-based core network. The major property identified in the literature for interworking of new and 4G systems is to maintain the standard 4G interfaces and GTP tunneling [62]. In 3GPP's 5G non-standalone architecture, this scenario is similar to option 4/4a/7/7a [50]. The solution is to enhance eNB and define additional interfaces. The standard also defines the situation when a 5G gNB (next generation NodeB) could be connected to an EPC through assistance from LTE eNB using a new interface Xx.
- 3) eNodeB (eNB) or heterogeneous legacy base-stations or phantom cells and RRUs of SDN-based mobile network share a geographical area. The SDN network should take

the spectrum resources used by the eNodeB or legacy cells into account while distributing resources among its RRU's. According to our search, no such scheme exists in the literature.

- 4) Handover between 4G and SDN-based mobile network. Mobility management mechanisms should be aware of the possible handover between the systems. For SDN-based core network, Softcell [59] suggests implementation of S1-MME and S10 interfaces so it can handover to and from legacy LTE EPC. 3GPP's recently approved standard for non-standalone 5G NR (New Radio) defines the handover between multiple RAT (Radio Access Technologies). The solution assumed a new interface between EPC and NGC (Next Generation Core) [50].

D. Disruptive Deployment

The issues of disruptive deployment and backward compatibility have a lot in common and the schemes discussed in the above section are in fact making it possible to deploy SDN networks in an evolutionary fashion. A good example is the recently approved 5G NR non-standalone standard which is an intermediate step towards full 5G NR deployment and it allows the operators to start experimenting with 5G equipment right away. However, the work on fixed networks are much more mature and sophisticated than wireless access networks and there are a number of investigations which showed that a small number of SDN switches are sufficient for the network to reap benefits of SDN in the transition phase [52], [54], [55], and [53].

For the radio access networks, schemes discussed in the previous section and introduced in [61], [63] provide the initial approaches towards evolutionary deployment of SDN-based mobile network building operators' confidence before major upgrade of the system. SDCN [9] talks about a phased approach towards fully programmable centralized RAN. In the first phase, the control or signaling plane of legacy systems, i.e., radio resource scheduling, hand-off, paging, etc., is moved to a logically centralized location or cloud. In the second phase the data planes of the base-stations can be implemented by SDR (Software Defined Radio) on commodity of the shelf (COTS) hardware. In the third phase CloudRAN is implemented with BBU pool using general purpose platforms and RRU's on the remote site.

Similar step-wise idea is presented in [64]. The proposed SDN-based mobile network architecture in [64] suggests a 3 step migration process from legacy network to fully complaint SDN network. It moves EPC to cloud while maintaining 3GPP interfaces with legacy nodes. In the second step, OpenFlow switch is introduced and GTP tunneling is replaced with MPLS tagging. In the third step, legacy nodes (S/P-GW) are removed from the network. The eNodeB stays the same in this architecture [57].

Softcell [59] proposed a SDN based LTE core network using commodity switches and routers. The major contribution lies in compressing thousands of rules and policies into manageable routing rules for off-the-shelf switches. They proposed that the carrier put a proxy on each base-station which can act

as GTP end-point. The Softcell switches then carry normal IP traffic between base-stations and core network. The network can have Softcells in one part while legacy core in the rest in the initial phase and can slowly move to full Softcell implementation. The proposal, however, does not cover the radio side of the network and base-stations are not made programmable.

Mobileflow [56] also allows part of the core network to be SDN-enabled but still work with legacy networks and base-stations. The architecture supports GTP tunneling and maintains the EPC standardized interfaces for compatibility. Independent implementation studies are required to confirm the claims of Softcell [59] and Mobileflow [56] that they can co-exist and work with legacy systems. Moreover, it is important to check the performance of such hybrid systems and how easy or difficult it is to implement SDN rules and policies on the flows.

An interesting classification of proposed SDN-based mobile network architectures based on revolutionary or evolutionary migration is done in [65]. Most of the evolutionary architectures are shown to have preserved GTP tunneling. OpenFlow, in general, does not support GTP tunneling and that is quoted as a major reason to use MPLS labeling [64] in SDN networks. Although, there are attempts to extend OpenFlow to support GTP tunneling [57]. [62] generalizes the evolutionary approach by identifying the importance of maintaining standardized interfaces by SDN network to have interworking with the legacy networks. Similar considerations were part of the evolutionary approach in [66] to move towards a flexible architecture for RANaaS (RAN as a Service). RANaaS chooses level of centralization according to the requirements and network state.

Three different alternative are being discussed for SDN deployment: overlay, white-box, or custom hardware²¹. Overlay technologies on existing infrastructure, such as, hypervisor for virtualization will offer minimal disruption but will not provide flexibility to use the hardware efficiently. Some optimization opportunities may be missed. Vendor-specific customized solutions using ASIC can also be deployed with minimal interruptions but will not provide a truly real-time programmable and economical platform. On the other hand, white-box solution requires a lot of standardization across the hardware and software domains to simplify the deployment and systematize the upgrade procedures.

1) *Lessons Learned:* Most of the ideas of evolutionary deployment of SDN-based mobile networks remains at a high-level and prototype implementation are necessary to ascertain their workability. From business perspective, the point solutions which can work with minimal disruption but can result in quick monetization are the most favorable. Now, the question is: what these point solutions will be in the SDN context filling in the puzzle pieces to complete the bigger picture.

The idea of non-standalone 5G NR (New Radio) in 3GPP's recently approved standard [50] is an intermediate step and a good trend to follow. The transitional phase is achieved

²¹<http://www.enterprisenetworkingplanet.com/netsysm/making-the-business-case-for-sdn-1.html>

through additional interface designs and evolving current LTE eNB into eLTE eNB which can connect to 5G core. 5G NR is a step towards full softwarization of the network through separation of control and user plane and modularizing of management functions.

E. SDN Specific Security Issues

The openness promised by SDN also introduces additional security risks and vulnerabilities [1]. Currently, carriers do not allow third party to meddle with the network protocols, routers, and switches, which results in an ossified network but its management and security is comparatively easier. Users' access to the networking fabric will increase occurrences of intentional and unintentional incidents affecting network operations. Moreover, the logically centralized controller is very critical component of the network. If the controller is compromised, the attacker will have tremendous access over the network [67], [68]. On the other hand, data is not co-located with control plane as in traditional network and the attacker needs to re-direct the data to get hold of it [67]. Moreover, the SDN virtualization abstraction layer integrates various ISP platforms while hiding the specific protocol details which makes security management comprehensive and manageable [69]. Furthermore, the capability of redirection or filtering of the flows based on packet content comes naturally with SDN and can enhance the network security [70]. According to [71], global network view, self-healing mechanisms, and the additional control capabilities give SDN great security advantage over traditional networks, though, the controller, forwarding plane and the links between them are exposed to new threats and attacks.

Another important issue with SDN networks is the configuration errors [72]. The network allows users and application to configure and program the network which may result in inconsistent policies and flow rules which could be both intentional or unintentional. It is very important to constantly check the network for inconsistent policies without adding a lot of monitoring overhead and overloading the controller with messages. Scalability and availability of the controller is also an issue with logically centralized SDN control plane. This issue is also discussed and looked at in the security related literature. A simple solution is to have distributed control where each controller is responsible for a cluster of switches with some overlap with other controllers to have robustness and resilience against failures [72]. Moreover, SDN are developed over GPP platforms which may have their own vulnerabilities. It is important to harden the platform or it will become potential surface for attackers [73].

Similar to other aspects of SDN, the work on security of SDN-based wired networks are more mature than in wireless SDN networks [67]. Although, the work is equally relevant for SDN-based mobile networks as even in the wireless settings, the SDN controller and links connecting it to the switches are still considered to be the most vulnerable points for attack [67], [74]. The wireless part of the system, however, has some unique issues, such as, user mobility, the co-existence of various generations of cellular systems and WiFi owned

by different operators, security of wireless link over shared medium, etc., [70].

Surveys of SDN-based networks' vulnerabilities and related solutions are presented in [67], [68], [72], [75]. Survey in [67] identifies possibilities of security threats and attacks in SDN in relevance with the traditional STRIDE (Spoofing, Tampering, Repudiation, Information disclosure, DoS (Denial of Service), Elevation of privilege) threat model and discusses respective work and proposed solutions. According to [67], DoS (Denial of Service) attacks are expected to be frequent in SDN networks than in the traditional networks as significant amount of data will be exchanged between the central controller and the switches. Spoofing may have less chances of occurrence in dynamic SDN network as it depends on tricking a network services based on obsolete information [67]. We have looked at each SDN layer and interfaces and discussed only the important issues identified in the literature. A comprehensive survey of security related work is outside the scope of present paper and we refer the readers to [67] and [72] (more recent) for thorough survey of security related research in SDN networks. The survey in [75] is exclusively for mobile networks.

1) *Controller*: The logically centralized controller is very critical component of the network. If the controller is compromised, the attacker will have tremendous access over the network [67], [68], [70], [75]. The idea of standby controller, when one of the controller fails, is discussed in [69]. There are several proposals for distributed control design for scalability and reliability [72], although, the distributed control plane and its interface for interworking are not properly explored. A Byzantine mechanism is also proposed where each network element is managed by multiple controllers. Optimization is used for controller assignment to minimize the number of controllers for a given Byzantine fault tolerance [72].

A simple DoS or DDoS (Distributed Denial of Service) attack can exhaust resources in the controller and also in the forwarding switches. Authentication for legitimate users can mitigate this issue up to an extent but a compromised application with legal credentials can inject faulty flows causing DoS. A flow-based anomaly detection mechanism can resolve this issue [67], [75]. [76] develops a connection migration technique on the data plane to classify successful TCP connections from unsuccessful ones in order to protect control plane from saturation attacks. Data plane can send network information to the control plane asynchronously via actuating triggers. According to [72], with SDN characteristics of dynamic flow table management and distributed control, the threat of SDN-specific DoS attack can be reduced.

On the other hand, the SDN controller provides unprecedented opportunities to develop adaptive and dynamic countermeasures for security threats. Although, the dynamic security control which can adapt to the changing networking conditions may also pose new challenges unknown to conventional networks [67]. A lot of work has been done in SDN-based firewalls which are mostly implemented as northbound API to the controller. The controller stores and dynamically update the Access Control List (ACL) whereas the traditional firewalls are static and may contain obsolete rules as they are manually

entered by network administrator [67]. In the existing security model, policies are in place to enforce the traffic to physically go through security middle-boxes and firewalls. In SDN, however, the paths are logically chosen by controller and it is the responsibility of the controller to ensure that firewalls etc., are enforced [73].

[72] surveys the access control mechanisms for OpenFlow where permission systems are developed to apply minimum privilege to the application instead of the current state where full privileges are open to all applications. Similarly, methods around enforcing the permissions are developed to secure the interface between controller and application along with authentication methods to stop unauthenticated users to access the controller. A summary of possible security threats, causes and solutions is presented in Table III.

Another issue raised in [77] is about the resilience problem of the current controller implementations, such as, Floodlight, OpenDayLight, POX, and Beacon. It is claimed that common application bugs are enough to crash existing controllers. From the security point of view, a simple malicious action, such as, changing the value of a data structure in memory can directly affect the working and reliability of the controller [77]. [75] and [73] point out that the packet without any match to the flow table is sent to the controller and an attacker can exploit this vulnerability and overwhelm the controller just by sending data. A possible solution is to design switches with enough processing power to forward maximum expected load and not plenty of packets [73]. According to [77], SDN development still need to go a long way to achieve the required security and dependability.

2) *Southbound interface*: Dedicated special connection can be provided to improve security of the links between controller and its switches, also called the southbound interface, but it does not eliminate the possibility of compromising the communication [67]. The term southbound may suggest direction of flow from the controller to the switches, although, this interface is bidirectional and we are using the terminology to conform with the SDN research community.

Most of the solutions developed for SDN are designed specifically for OpenFlow [67], although, the threat vectors are independent of any technology and based on the conceptual layers of basic SDN architecture [77]. An important issue with OpenFlow is that Transport Layer Security/Secure Socket Layer (TLS/SSL) encryption is optional and not enforced [67]. Moreover, TLS/SSL based communication is not strong enough to protect the control channel from IP based attacks [74]. Furthermore, the digital SSL certificate used for the communication is self-signed and is not very secure [73]. If the private key of the certificate is stolen, the attacker can eavesdrop and join the network.

If SSL encryption is not used, it is possible to launch an ARP (Address Resolution Protocol) spoofing attack on southbound interface where an attacker's MAC address is linked to a legitimate IP address. Counter-measures to ARP spoofing attack include packet-level monitoring [67]. Similarly, lack of encryption can also cause Man-in-the-Middle (MiM) attack. MiM can misdirect and drop flows and can scan the network and compromise confidentiality. Alternative encryption

methods along with proper auditing and logging methods are suggested to counter MiM and verify non-repudiation [67].

In a distributed controller environment, where each controller is responsible for a cluster of network switches, the interface between controllers to share important information is also vulnerable similar to the southbound interface and can have similar types of attacks [67], [75]. As far as OpenFlow is concerned, this interface is not properly researched and standardized yet and it is still to see what shape the interface will take.

3) *Northbound interface and application layer*: The northbound interface is also exposed to significant threats as the user applications with power to access and define network resources and processes can also abuse the network intentionally or unintentionally [67]. There is a lack of standardization for application access management regarding rules for granting privileges, binding mechanisms, auditing, trust mechanism between network management applications and controller, etc., [68], [70], [75].

In order to prevent any malicious application to gain access to the controller and network, some advanced solutions are also proposed besides establishing trust and authenticating the user [72]. These solutions uses techniques, such as, resolving the conflicting rules based on the priority of application, securing the controller core by sandboxing each application, and creating an isolation layer between application and control plane for fault isolation [72].

4) *Data plane*: Forwarding switches are vulnerable to different types of attacks, such as, DoS, modification or injection of fraudulent flow rules, compromise, etc. A compromised switch can discard, slow down, or deviate network traffic [70]. Even worse, it can launch a DoS attack on the controller [70]. The survey in [75] also mentions the possibility of flow table overflow in the switches due to an intelligent attacker sending flow requests with slight difference in parameters.

In OpenFlow, VAVE (Virtual Address Validation Edge) module verifies the address of external packets with no record in the flow table. The module is extended to include binding validation and a data-set of black and white lists to counter IP spoofing attacks [67]. Random and continuous change of host identities are also proposed as a counter-measure where an attackers scans the network to learn flow patterns [67]. According to [77], some attacks, such as spoofing, is not specific to SDN but has a much larger impact in the SDN. The authors give an example of spoofing the address of the controller and comment that a smart attack which would last only for few seconds just enough to install malicious flow rules is very detrimental and very hard to detect.

In a virtualized environment, it is also very important to build tools which can verify logical separation between virtual networks. In practical networks, a VM (Virtual Machine) may access some resources of another VM sharing the same physical platform, tampering the network information [67]. Formal verification based methods are being proposed to verify the isolation of traffic between network slices [72]. So far, the research is focused on OpenFlow and should be extended for broader SDN deployments.

Configuration issues where intentional or unintentional programming of the network causes conflict in policies and flow rules are also termed as a major problem specific to programmable SDN networks. Not just applications, but multiple entities with write access can override or misconfigure the related flows and compromise the switch functionality [72]. As new applications and new devices are added into the network, security techniques should be in place to make sure network correctness and operability [72]. According to [72], the area of configuration issues is one of the widely researched theme in SDN and solutions are proposed for detection of network errors and data plane verification. These proposals can be classified as real-time or non-real-time. In non-real-time solutions, symbolic executions are used to test OpenFlow applications for correctness and not causing the network to reach an inconsistent state. Binary decision diagrams are used to check intra-switch misconfiguration [72]. Flow policies and flow requests are also verified not to by-pass firewalls. These verifications may take several minutes and are not designed for real-time anomaly detection [72]. There are real-time tools as well for verification of flow tables (e.g., identifying routing loops, unavailable paths, etc.), conflict resolution for applications in a firewall, and detection of firewall violations [72]. According to [72], the real-time tools show the evolution in right direction, although, experimental deployments will be needed to bridge the gap between theoretical models and characteristics of live network.

The survey [72] also lists the northbound APIs which are designed for conflict resolution in policies. They extract non-overlapping policies from the flow rules before instructing the controller to install the flow rules into the switches. The work, however, should be extended for distributed control environment. Policy management techniques are also proposed for conflict resolutions across multiple modules in different layers and also for comprehensive policy update across all the relevant packets and flows [72]. According to [72], a concern with these policy conflict resolution is their scalability for larger applications or larger networks as all of these techniques are highly computationally extensive. Another alternative is to maintain a strongly consistent data store, instead of conflict resolution in policies, to maintain network consistency.

A summary of identified attacks and their solutions with relevance to SDN-based networks and specially SDN-based mobile networks is presented in Table III. As we have discussed before, most of the work is mainly focused over OpenFlow and its mechanisms. Table III is definitely not complete and as the state of the art moves forward, new threats and vulnerabilities will be identified. [68] presented a similar table but we have added information from other publications into the table and talked about the possible solutions to have clear idea about the current state of the art. Some attacks can be classified differently than Table III considering the multiple layers and interfaces it may affect as also noted in [72].

5) *Issues exclusive to SDN-based mobile networks:* The security analysis and counter-measures discussed above are designed for wired networks only and mobile environment

features and cross-domain interoperability is not considered in any of them [69]. In this section, we highlight the security issues and solutions specifically for cellular and wireless networks. As can be seen in this section, security of mobile networks has not gained a lot of attention in the research community.

a) *Vulnerabilities:* [69], [70] identify four major issues in SDN-based mobile networks besides the threats and attacks identified for the wired part of the network. These issues are: 1) mobility between different access technologies using different security protocols, 2) mobility between networks owned by different operators, 3) highly constrained computational capacity and storage, and 4) backward compatibility to older generations of mobile systems. These issues create unique security problems in mobile network domain.

The survey in [75] looked exclusively at the security issues of SDN-based mobile networks. The paper classified all attacks in terms of SDN layers as also done in the present paper and extends the scope to mobile terminal and radio access medium. Mobile terminals are in ever increasing danger of Trojans and viruses. The terminals are resource constrained in terms of computability and storage and cannot execute heavy duty intrusion detection modules [75]. It is important to customize the traditional tools to suit mobile terminals' need of light weight safety procedures. Malicious or legitimate RF interference can cause saturation attack at the access points. Threat of DoS from base-stations to the core network is considered low in [78] as the traffic from the users to the network or the base-station control traffic is generally low. [75] also surveyed the security literature in SDR (Software Defined Radio) area and identified RF interference, mobile terminal malware, and MAC tempering as some of the detrimental attacks in wireless environment.

Mobile user's privacy and confidential information can be leaked in an open backhaul network architecture [78]. Currently, the traffic is tunneled using IPSec to the core network to secure the mobile terminal's identity from any eavesdropper. Similarly, important control traffic can also be compromised over open backhaul. Another issue raised in [78] is about a malicious access point or base-station gaining access to the network by spoofing its location using a legitimate base-station location. Counter-measures based on GPS tracking can be developed to mitigate the attack. Moreover, the small base-station/access point/remote radio unit is also vulnerable to physical tempering.

b) *Secure Architectures:* Some work has been done in the SDN-based mobile network area which is also focused on OpenFlow. [79] adds vertical forwarding extension to OpenFlow to deal with mobility and access control management while forwarding to legacy elements of network. [80] identifies the lack of secure mobility support which make OpenFlow inapplicable in wireless and mobile systems. OpenFlow cannot handle switch mobility, e.g., in a moving train, and secure change of IP addresses in a fast and dynamic mobile network to maintain secure communication between the controller and the switch. The solution in [80] is based on HIP (Host Identity Protocol) security method to identify a host either by a host identifier or a host identity tag. A flow control agent is

TABLE III
SUMMARY OF IDENTIFIED VULNERABILITIES AND SOLUTIONS.

SDN layer/interface	Threats	Causes	Solutions	References
Application layer	Fraudulent access	Lack of access control	Access control	[68], [72], [75]
	Accountability	Lack of binding mechanisms	Secure standard for APIs	[68]
	Malicious application	Users' credential/identity theft	Anomaly detection	[70], [72], [75]
Northbound interface	Malicious access	Limited secure API	Strong access control	[68]
	Policy manipulation	Lack of binding mechanisms for applications	Secure standards for interface	[68], [72]
	Fraudulent rule insertion	Malicious applications	Access control	[68], [72]
Control Plane	DoS, DDoS	Flooding to affect legitimate flows	Flow level anomaly detection	[67], [68], [70] [72], [75]
	Compromise	Visibility and openness of controller	Isolation of controller, fail-safe mode for switches	[68], [70], [72]
	Unauthorized access	Lack of access control	Strong and adaptive access control	[68], [72]
	Privilege escalation	Tampering of access control	Auditing	[67]
Southbound interface	Man in the middle attack, repudiation	Optional use of TLS/SSL encryption	Encryption	[67], [68], [70], [72]
	TCP level attack	TLS is susceptible	HIP based IPSec tunneling	[68], [74]
	ARP spoofing	If SSL (optional) is not used	Packet level monitoring	[67]
	Scanning attacks	Scanning worms, improper controller access,	Encryption, standards for controller access, anomaly detection for scanning worms, anonymization, time-out randomization	[67]
Data plane	IP spoofing	DNS tampering	Authentication, black/white lists for flow rules	[67]
	Tampering	Illegitimate modification	Encryption	[67], [72], [75]
		Improper isolation between virtual networks	Tools to verify slice isolations	[67]
	Fraudulent flow rules	Compromised switch	Anomaly detection	[70], [72], [75]
Controller-controller interface	Similar to southbound interface			

also used in [80] to update controller of the new location information for location based services.

[74] proposed an architecture for SDN-based mobile networks using HIP (Host Identity Protocol), IPSec tunneling, and SecGW to provide secured control channel. Although, the evaluation used OpenFlow but the proposed architecture is independent of SDN protocol. Resilience against various attacks, such as, DoS, IP spoofing, eavesdropping, etc., is analyzed. The architecture is shown to be vulnerable to volume based DoS attacks and an access control method or firewall is recommended to be used to prevent such attacks. In a follow-up work a software defined monitoring and data collection scheme is implemented on top of the secure control channel to prevent, detect, and react to the security attacks [68].

[69] proposed a SDN-enabled security architecture for mobile networks where a security layer is introduced above the control layer to monitor the application and stop the malicious traffic from entering the radio environment. They proposed local agents at the wireless routers and access points for fast responsiveness. Since SDN hides the complexities of multiple domains and their specific protocols, it is relatively easier to provide comprehensive security through SDN architecture.

[81] proposed a three-level control plane structure for SDN-based mobile networks to implement a unified security, connection, mobility, and routing mechanism. They have included device controller (for UE), an edge controller, implementing 5G network functions through control applications, and an orchestration controller to coordinate utilization of cloud resources. Two versions of edge controller are identified in the paper, one implemented in the cloud and the other

TABLE IV
SUMMARY OF SECURE SDN-BASED MOBILE NETWORK ARCHITECTURES.

Proposal	Major security feature(s)	General attributes
[79]	Access control for legacy traffic	OpenFlow extended with vertical forwarding
[80]	HIP (Host Identity Protocol) w/ location update	OpenFlow with secure mobility support
[74]	HIP, IPSec, & SecGW for control channel	Independent of OpenFlow, resilience against DoS, IP spoofing, eavesdropping, but not volume based DoS
[68]	Traffic monitoring	Extension of [74]
[69]	Security layer over control layer for access control	Local agents at wireless APs, high-level design
[81]	3-level unified control for security, mobility, routing, etc.	High-level idea only
[82]	Exploit channel characteristics	Focused on D2D communication
[83]	NFVI-TP (NFV Infrastructure- Trust Platform)	Framework for security and trust within NFV-MANO
[84]	Authenticated handover	Low-latency and simple authentication via centralized controller
[85]	Authentication, access control, periodic key updates, routing conflict analyzer	3 OpenFlow GWs for UE, RAN, & network

in the mobile device for mission-critical control and out-of-coverage services. Without going into the details of the network functions, the paper describes that security along with radio resource management, mobility, etc., is implemented through control applications by the edge controller.

[82] addressed the computational limitations of small wireless devices by proposing a light-weight security framework for D2D communication. They define spatial transmission region for D2D link that can guarantee a minimum secrecy rate through exploiting the physical characteristics of the wireless channels. Analysis shows that as number of users increases, the achievable transmission capacity also increases due to availability of more D2D pairs.

A framework for security and trust for 5G is proposed in [83] in the context of virtualization and SDN. The authors argue that security mechanism in cloud computing are not sufficient for 5G security challenges and required extensions to provide sufficient security services and functions based on NFVI (NFV infrastructure) in a trustworthy and economical way with cooperation of various trustworthy VNFs (Virtual Network Functions). A security and trust framework is proposed within NFV MANO (Management and Orchestration) architecture with NFVI-TP (NFVI Trust Platform) embedded in NFVI by an authorized party. It ensures NFVI platform layer security by providing a root trusted module, a hardware resource, to ensure every component build upon it is certified as trusted. The framework is not evaluated through implementation though. Moreover, the framework and associated discussion is generic enough for fixed networks and does not address any particular issue regarding mobile networking.

The idea of taking benefit of SDN features to provide security services is used in [84], where a low-latency and simple authentication handover scheme is proposed for 5G dense HetNets. The major benefit comes through the centralized control of SDN which is also assumed to host an authentication handover module. Both access points and user need authentication before accessing the network. The simulation results show low handover delays than the traditional method.

In a recent publication [85], an SDN based 5G architecture

is proposed with three different OpenFlow gateways and controllers, one each for the mobile user, for the RAN, and for the network. The network is comprised of heterogeneous technologies. The authors identified potential threats and proposed security mechanisms, i.e., authentication, access control, periodic key updates, routing conflict analyzer, for communications between all types of the architectural entities and showed the proof of their feasibility and strength.

A summary of the secure architectures discussed above is presented in Table IV, where the major security feature (or features) and important architectural attributes are also highlighted. As it be seen from the table that access control is one of the popular mechanisms to secure northbound interface. Also, more detailed solutions emerged as extensions to OpenFlow and the once claiming to be independent of the physical infrastructure have used OpenFlow for proof of concept. The current state of the art with respect to secure SDN-based mobile networks still needs solutions for broader spectrum of threat models including all types of DoS attacks along with comprehensive evaluation of robustness of the solutions.

c) Tools: [86] proposed a security assessment scheme to quantify security levels to networks. The scheme is proposed for mobile SDN networks although it is generic enough to be used in other networks. The security level of a network is defined in terms of efforts to reach a target error state or in launching an attack. More efforts required to launch an attack would represent resilience or high security level of the system. The authors want to explore alternative methods for quantification of effort in their future work.

Security issues in embedded mobile devices are discussed in [87] and a learning intrusion detection system is proposed for OpenFlow networks. Statistically different traffic than the user-defined normal traffic is termed as anomalous. This work is not directly related to cellular systems rather it is designed for embedded mobile devices with strict constraints on transmission power and computational capacity. The idea of intrusion detection system to detect DoS though has application in cellular scenarios. A formal verification method is developed for OpenFlow to make sure systems' correctness

and security [88]. Malware protection for mobile devices based on OpenFlow protocol is discussed in [89]. The protection system uses real traffic analysis of connection establishment packets inside an OpenFlow controller.

Considering the storage and computational limitations of mobile devices for very heavy security applications, [90] proposed outsourcing security to cloud using OpenFlow virtual switch in the device and OpenFlow controller residing in the cloud along with security manager. This method takes away major computational load from the device, employs variety of security services with fast processing. The communication delay is an important factor affecting performance and so far it is not clear if faster processing overcomes the transmission latency or not [90].

6) *Summary of outstanding security Issues:* A comprehensive survey of SDN security solutions have been done in [67] and [72]. However, the solutions are still not mature enough to be used for production deployment. Here we list some of the take-away points from our survey of security literature for SDN networks.

- 1) **Comprehensive evaluation** Although, a lot of work has been done in SDN security but independent experimental evaluations and comprehensive assessment of solutions are required before the technology is ready for deployment.
- 2) **Controller security** Controller compromise and malfunction can cause huge damage to the network. DoS or DDoS attacks on the controller are the simplest and most expected attacks in SDN environment. Distributed control plane is an important paradigm for reliable and scalable network control. This paradigm is not explored in detail and the inter-controller interface is not yet defined.
- 3) **Access control** There is still lack of standardization for applications to access the controller. The permission schemes [72] should be thoroughly evaluated for correctness and computational feasibility.
- 4) **Southbound interface** A dedicated link can connect controller to its switch but appropriate level of encryption is needed to provide secure communication. The monetary implications of dedicated link and feasibility should be studied.
- 5) **Isolation between virtual networks** Real-time tools are needed to verify complete logical isolation between the virtual networks sharing the same physical resources.
- 6) **Error/anomaly detection** Anomaly or error detection is an important element in the network where there is high risk of controller/switch compromise. The techniques should be tested for overhead, detection accuracy and time, and it is also necessary to secure the anomaly detector from security attacks.
- 7) **Policy conflict resolution** The SDN network allows users and application to configure and program the network which may result in inconsistent policies and flow rules. There are real-time tools for policy control resolution but they have scalability issues while dealing with large applications or large networks.

- 8) **Data leakage and modification** It is shown in [72], that there are no solution till date for data leakage and modification in SDN networks. However, encryption is sought to be a valid solution for SDN networks as in traditional networks but the physical separation between controller and switches requires very important control information to be pass over the communication links. Their security should be comprehensively tested for the available encryption standards.
- 9) **Solutions conforming to SDN principles** SDN networks are supposed to have multi-vendor interoperability, third party application support, integration of virtualization [72], and fast and dynamic reconfigurations. Any security solution proposed for SDN should also conform to the norms of the network.

F. Clear and compelling business case

Despite the tremendous work in SDN and towards its deployment, an operator survey conducted in 2015 reveals that the majority thinks that lack of clear compelling business cases for SDN is the major obstacle in its widespread adoption and roll-out. The deployment cost came out as the second major obstacle²².

SDN's early adoption in enterprise networks is on its way in big companies, such as Google, Amazon, Facebook, Microsoft, AT&T, etc.,²³. A clear business case is present in all these circumstances, i.e., distributed organizations with many branches can benefit greatly by software and cloud based WAN (Wide Area Network) to provide low-cost Internet services easily and quickly.

Moreover, there is uncertainty around the benefits in terms of CAPEX and OPEX savings claimed by SDN technology. Virtualization, use of GPP, and resource pooling in cloud computing certainly result in cost savings but it is also discussed that cost reduction does not bring the true value of SDN in the picture, rather it is the faster roll-out capability with overall streamlining of network operations which will provide the economic benefits to operators and network owners. According to a recent survey²⁴, the flexibility and scalability of the network with respect to the demand are the main drivers behind investments in 5G technologies. SDN and NFV are the key enablers of a network that flex with the traffic. The fast-fail approach of SDN will result in quicker roll-out of new services but at this stage, it is unclear what services will be commercially viable and what revenues they will generate [1]. Revenue is a key factor in recovering deployment costs and the Return On Investment (ROI) for SDN is an important consideration. Although, the faster roll-out of services which can reduce the week-long processes to a few minutes may result in multi-million dollar revenue stream for the carrier for services where time to market is a significant success factor.

²²<http://www.lightreading.com/carrier-sdn/sdn-architectures/defining-use-cases-and-business-cases-for-sdn/a/d-id/716315>

²³<http://searchsdn.techtarget.com/answer/Whats-the-status-of-SDN-deployments-in-the-enterprise>

²⁴<https://www.sdxcentral.com/articles/news/carriers-5g-plans-rooted-sdnnfv-says-ixia-survey/2017/09/>

1) *Virtualization*: An NFV approach can reduce the number of physical servers, but NFV is not dependent on SDN and although SDN supports NFV by providing a flexible connectivity platform, virtualization can be achieved without it [1]. Service orchestration without any change in infrastructure could generate more benefits than SDN/NFV in the short run and a massive transformation of infrastructure can slow the ramp of savings²⁵. On the other hand, the SDN integrated NFV is able to support a sustainable business model for the long term future whereas the traditional custom stack based virtualization approaches will likely fail [91]. According to [91], TCO (Total Cost of Ownership) of traditional virtualization approaches tracks the growth in requirements whereas a common standardized platform based virtualization, e.g., integrated SDN and NFV, breaks the linkage between cost and requirement and can have 62% lower TCO than traditional approaches.

A ROI analysis by ACG Research [91] claims that the break-even point can be achieved in one year when traditional virtualization solutions are gradually phased out in favor of a common and standardized platforms, e.g., SDN, for NFV. This analysis considers network transformation cost including equipment and training and predicts over 350% ROI in 5 years for common platform based virtualization.

In [92], a techno-economic model is developed for SDN based mobile networks for CAPEX and OPEX savings and TCO (Total Cost of Ownership). The models predicts 68% CAPEX reduction, 63% OPEX reduction and 69% TCO reduction when SDN is adopted for mobile networks in comparison to the traditional networks. For base-station virtualization, [92] assumes CloudRAN architecture but did not consider the cost of fiber for fronthaul. Similar reductions are reported by ACG Research report [93] with 68% lower CAPEX and 67% lower OPEX by virtualizing EPC. Again, the prohibitive cost of fronthaul is not considered here as well.

2) *Use of GPP*: The use of GPP or commodity processors instead of special purpose hardware also reduces the cost significantly [1]. But this cost saving is based on the view that traditional router vendor equipment is over-priced, although, not everyone in the industry shares this view [1]. Moreover, considering the overall cost of re-organization of telco companies to suit the SDN model, the savings may not be as great as predicted. Similar concerns are also raised about the OPEX savings. The challenge of operating a network with white-box switches from one vendor controlled by a virtual machine from another vendor may be significant and OPEX may be higher than the current technologies [1].

3) *Resource pooling in cloud*: Recently, there are attempts to quantify the cost savings using cloud-based networking. An initial effort is made in [94] where the Cisco price list is used to determine the CAPEX of two use cases of mobile networks. The classical networks are compared against a hypothetical SDN-based network with similar configuration and another version of a SDN-based network with virtualization and sharing. It is shown that a SDN mobile network can save 13.81% CAPEX and with virtualization, the savings can go

up to 48.04%. The paper, however, made certain assumptions, such as, the Cisco switches can be upgraded to have OpenFlow which is not impossible. Although, similar assumptions for base-stations are not true. We still do not have a concept for a SDN-enabled base-station. The costs of fronthaul and backhaul are not considered in the study, and can be significant. Moreover, the study did not consider the associated cost of the disruptive nature of network transformation along with training of the staff [94]. In a follow-up presentation [95], the author discusses qualitative comparisons for OPEX savings considering the centralized and simpler control of SDN networks, though, concern is expressed regarding the initial rise in OPEX to get the new infrastructure to work.

[96] performs simulations using the OPNET simulator to compare baseband processing resources for CloudRAN with distributed RAN. The paper concludes that CloudRAN needs 4 times less resources than distributed RAN based on the daily load forecasts. The cost of fronthaul is not considered in this paper which could be prohibitive for various operational scenarios. In a follow-up paper, however, the authors provide detailed evaluation of split function fronthaul cost savings in terms of different multiplexing gains [33]. As expected, the fully centralized solution with all BBU functionalities in the cloud gives the best multiplexing gains but this puts a lot of capacity requirements on the fronthaul and can only be feasible for operators with cheap fronthaul access. For low traffic load, the BBU pool should have higher layer processing and requirements on the fronthaul can be relaxed [33]. The evaluation work in [97], shows a potential of a 30% reduced requirement for resources when base-stations opt for centralized processing versus the local processing. In the local processing option, each base-station is provided with the resources to handle peak time traffic.

The theoretical framework introduced in [98] shows 10-15% savings in capital expenditure per square kilometer when cloud-based RAN is compared against a traditional LTE network. However, the theoretical framework lacks some important information, such as, the cost of base-stations in cloud, and assumes that this cost will be lower than traditional systems. The traffic load, usage pattern, and revenue levels in 2015 from an average Finnish network is chosen as a reference case in [99] to compare the CAPEX and OPEX of SDN-based LTE versus regular LTE. It is shown that SDN reduces the network related annual CAPEX by 7.72% and OPEX by 0.31% compared to non-SDN LTE. The savings are very small comparing to the annual cost of a MNO (Mobile Network Operator) and do not present a compelling case for SDN. However, if cost savings can be translated through to profit, they would represent a significant rise in profit, which could act as a deployment incentive.

Another recent study [100] tries to find the central location and needed equipment for CloudRAN as a minimization problem from CAPEX and OPEX perspective. The results show that maximum or full centralization yields a minimum CAPEX solution for certain LTE-A configurations (40MHz). But interestingly, lower levels of centralization, i.e., more central locations with small IP/MPLS equipment, also yields up to 18% savings compared to fully centralization when

²⁵<http://blog.cimicorp.com/?p=2716>

higher capacity CPRI is available to cater for high rate LTE-A data (100MHz). The main reason behind this result is the price gap between big routers (6.72Tb/s and 48 slots) versus small routers (1.40Tb/s and 10 slots). Similarly, OPEX savings of up to 7% is shown compared to fully centralized solution mostly due to much higher power consumption of big routers in comparison to the small routers. The lower level centralization is shown to save 37% CAPEX and 82% OPEX when compared against fully distributed approach, i.e., the current cellular model. This analysis assumes the availability of fronthaul and the cost calculations only consider the BBU equipment at the central locations.

The fronthaul factor for the cost is considered in the cost model for CloudRAN developed in [101]. The paper considers the ratio of cost of fiber per km to the cost of one BBU. When this ratio is greater than unity, it is not beneficial to go for a centralized RAN. The paper also showed better results in case of partial centralized RAN solution when cost is minimized with respect to all the factors affecting it [101].

A summary of cost saving estimates are plotted in Fig. 8. Although, the estimates depend on specific method and assumptions used in the calculations. The numbers should not be compared in strict sense or interpreted as precise benefits. But, it is interesting to see that virtualization seems to have a better potential for cost savings than cloudification. It is not surprising, however, given the potential for end-to-end resource sharing with virtualization. Another factor affecting cloudification savings is the difference in prices of big and small routers [100] and it may not be cost-effective to replace multiple small routers with one big router. Moreover, an optimal design should also consider the cost of fronthaul.

4) *Products and business analysis:* As the research in SDN-based mobile networks is a step behind wired network developments, there are only a few wireless products emerging in the market whereas there are relatively more WAN-related SDN solutions and products. According to SDxCentral²⁶, the following mobile network related products are announced already:

- 1) **NEC's NFV CloudRAN²⁷** is a software application which runs on the central digital unit (DU), with COTS servers, but can delegate L1/L2 functionalities to the radio units (RUs) according to the fronthaul condition. It can work with both Ethernet and CPRI as well as one DU can have flexible functional split in different parts of the network.
- 2) **ASOCS's virtual Base-station²⁸** is a fully virtualized base-station with software of all layers running on GPP servers. It supports flexible baseband partitioning and provides interfaces for digital DAS (Distributed Antenna System). With ease of installation and integration with any carrier, it offers exceptional total cost of ownership.
- 3) **AltioStars vRAN solution²⁹** connects the intelligent RRU with the virtualized compute nodes over any trans-

port network. The intelligent RRU understands applications and schedule packets over the air to achieve required QoE (Quality of Experience). This may allow full CloudRAN experience even with constraint fronthaul. Initial trials with SK Telecom have been concluded already where key LTE features are validated. Major customers include SK Telecom, Dali Wireless, etc. Dali Wireless has recently announced the development of a patented virtual fronthaul interface as an intelligent aggregator-router. That interface makes it possible for multiple operators to connect to multiple BBUs, creating a multipoint-to-multipoint network³⁰.

- 4) **Advantech's Packetarium XLc carrier grade glade server³¹** is designed for NFV infrastructure for CloudRAN and mobile edge computing (MEC). It can host 9-12 Intel Xeon processor based blades and supports 720Gbps I/O and more than 1.2Tbps switching capacity.
- 5) **Netsias vRAN platform, called ProGRAN³²**, lets network operators slice the RAN and effectively allocate a portion of the network resources for specific application, such as, safety. Trials to integrate ProGRAN in Telefonica's virtualization project are already underway.
- 6) **AT&T's FlexWare**, previously called Network on-demand, is based on AT&T's integrated cloud platform for service orchestration. It uses SDN and NFV to provide software control of different functions for enterprise customers and can be used to offer MEC-capable services. FlexWare is part of AT&T's 5G trials which will also use mmWave spectrum along with WiFi to provide blanket coverage³³.
- 7) **BigSwitch's Big Mon (Big Monitoring Fabric)** can evaluate the traffic of millions of mobile subscribers allowing service providers to monitor their network performance and ensure ultra-high data performance³⁴.

There are a number of products and solutions in wired network domain. Some are relevant to mobile networking as well. It is impossible for us to mention all products because of space constraints but some major developments are:

- 1) **Nokia's Network Service Platform (NSP)** is purpose-built Carrier SDN software for service automation, network optimization and dynamic assurance for delivery of profitable, on-demand network services³⁵. Integrated with Nokia's Deepfield, NSP can perform real-time analytics for resource optimization and network security.
- 2) **IBM** has range of products available in SDN family designed to create a unified network architecture which

³⁰<https://www.sdxcentral.com/articles/news/dali-wireless-uses-sdn-to-virtualize-the-fronthaul/2017/10/>

³¹<https://www.sdxcentral.com/products/packetarium-xlc-carrier-grade-blade-server-for-virtual-service-edge/>

³²<https://www.sdxcentral.com/articles/news/netsias-platform-lets-operators-slice-dice-ran/2017/12/>

³³<https://www.sdxcentral.com/articles/news/att-will-use-flexware-platform-waco-5g-trial/2017/12/>

³⁴<https://www.sdxcentral.com/articles/news/big-switches-big-mon-keeps-eye-mobile-subs/2017/12/>

³⁵<https://networks.nokia.com/products/network-services-platform>

²⁶<https://www.sdxcentral.com>

²⁷<https://www.sdxcentral.com/products/nfv-c-ran/>

²⁸<https://www.sdxcentral.com/products/asocs-virtual-base-station-vbs/>

²⁹<https://www.sdxcentral.com/products/altio-star-vran-solution/>

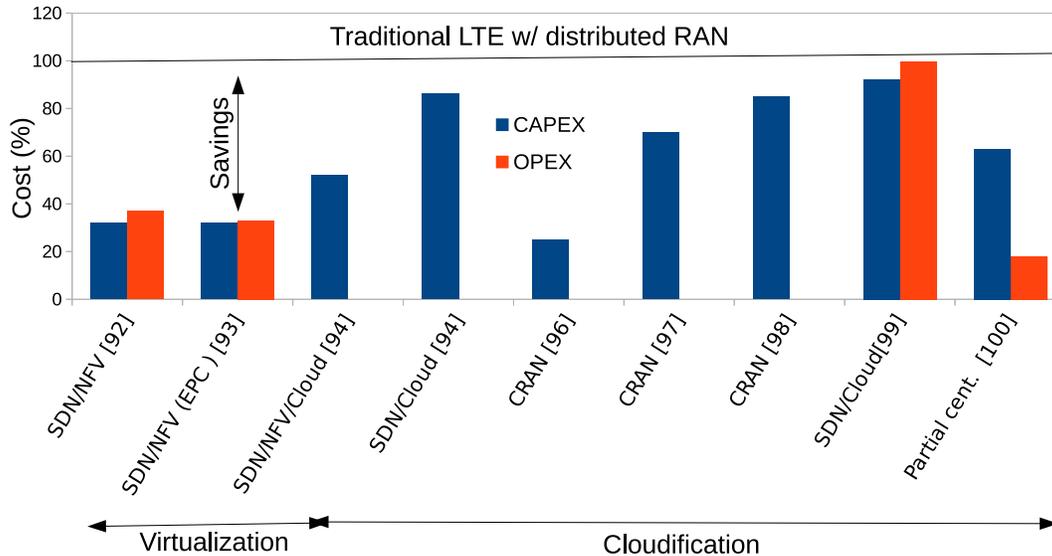


Fig. 8. CAPEX and OPEX saving estimates when SDN is used with/without virtualization and/or cloudification. CAPEX and OPEX saving estimates for CloudRAN (CRAN) are also plotted. The reference (100% cost) is traditional LTE network with distributed RAN. Note: details of every estimate method, assumptions, and limitations are discussed in the text.

enables cloud and big data analytics and optimizes the entire computing infrastructure, i.e., compute, storage and network resources³⁶.

- 3) An earlier example is **Juniper's High IQ**, a programmable networking approach, introduced in 2014. ACG Research analyzed three use cases for High IQ from simplifying CPE (Customer Premise Equipment), and pushing the VPN and firewall functions to the cloud, to real-time network self-optimization and elastic traffic. The report indicated compelling business cases through significant CAPEX and TCO savings in all the use cases³⁷.
- 4) In 2016, ACG Research's white paper compared **HPE pre-integrated NFV solution** to a bottom-up DIY solution and show the extra OPEX savings of 30% can be achieved with HPE pre-integrated solution only³⁸.

5) *Lessons Learned:* From this review there is generally consensus that a positive business case for SDN is evident. A significant issue of network transformation costs is identified, based on the disruption to operational approaches and networking infrastructure. The radical change in the communication network architecture, with full softwarization and centralization, focusing only on the broader CAPEX and OPEX savings in future does not account for transformation costs. A rapid transformation model is very unlikely to be followed, due to the high risk of service disruption.

A more practical alternative would be an incremental or evolutionary approach to the introduction of SDN into the operational networks. Although, there still need to be indi-

vidual business drivers or short-term monetization for each incremental implementation to make it appealing for carriers and network owners [1]. While the evolutionary path to full SDN deployment is also not clear, the individual business cases for the increments are among the big unknowns. Wherever, an enterprise has found some compelling cases to upgrade, the SDN point solutions happen [1]. Carriers will establish a roadmap of SDN/NFV solutions which will minimise disruption and show short term business benefits [1]. Thus, to determine the duration of transformation to a full production deployment of SDN is challenging.

IV. CURRENT STATE OF THE ART OF SDN-BASED MOBILE NETWORK ARCHITECTURE

As we have seen in the previous section that there exists a gap between the state where the SDN-based mobile network technology will be ready to be deployed and the current state of the solutions. But not all aspects of the SDN-based mobile network architectural development are covered in the previous section concerning major challenges. In order to really appreciate the current state of the art, we should also look at the major body of relevant literature. A detailed discussion about all the work under SDN-based mobile network architectural development is outside the scope of the current paper. In this section, we have discussed the major lessons from our study in the form of common and popular themes and approaches which may become part of the standards.

The earlier papers on SDN-based architectures talk mostly about the benefits as they have to make a case for SDN technology. The architectures they presented are mostly very generic and high-level. As time goes by, the proposed architecture still remains high-level but there are publications which are more focused on the issues and emphasize on details.

³⁶<https://www.sdxcentral.com/listings/ibm/>

³⁷<http://www.juniper.net/us/en/local/pdf/whitepapers/2000581-en.pdf>

³⁸<https://www.hpe.com/h20195/V2/Getdocument.aspx?docname=4AA6-4924ENW>

Recently, there has been a lot of work on specific network services, such as mobility management [102], resource management and sharing [103]–[105], CoMP (Coordinated Multi Point) [106], service chaining [107], network slicing [108] etc. In this paper, we stayed away from network services as a thorough coverage would make this paper too long.

There has been preliminary implementations with some basic results regarding overheads, cost savings, etc., but a big gap still exists between the current state of the art and the tech-ready state for production deployment. The testbeds available for implementation are also limited by GPP constraints as discussed in Section III-B and until we have the platforms suitable for signal processing, it is very difficult to see a real and comprehensive SDN-based mobile network. We have also listed the outstanding issues and problems related to SDN-based mobile network architecture.

The rest of the section is mainly organized in terms of SDN planes and what has been proposed for each of them and what ideas are repeatedly discussed among the research community. Table V shows the summarized landscape and an overview of the repeated themes, such as, the realization of EPC nodes as applications in SDN-based mobile network, distributed control plane, and data plane comprising of CloudRAN and heterogeneous radio technologies- e.g., LTE, 3G, and WiFi, etc. for backward compatibility- and OpenFlow (OF) switches. The themes are grouped under the associated SDN layer in Table V.

A. Application Plane

While transforming LTE architecture into SDN's 3 plane model, EPC control nodes, e.g., MME, HSS, and PCRF are usually suggested to be implemented as application plane modules as shown in Table V [4], [56], [81], [109], [110], [112], [113], [115]. Similar ideas are evident into the evolving 5G system's architecture from 3GPP [2], where mobility management and policy control functions are connected to other new functions, such as, authentication, application management, etc. via a common message bus. The architecture is a step closer to the softwarization or cloudification of core network.

However, the case with S/P-GWs is not very clear. In the 5G system architecture from 3GPP, the User Plane Function (UPF) to handle the user plane path of PDU sessions remains in data plane [2]. UPF in 5G architecture is supposed to be the equivalent of PGW-U (User plane part of PGW) in EPC. The equivalent of PGW-C (Control plane part of PGW) is SMF (Session Management Function) in 5G NGC (New Generation Core) [2]. The separation of control and user plane in 5G NGC is more comprehensive than in EPC. Although, this separation is not exactly the same as the separation of control logic and forwarding hardware defined in SDN.

A detailed study about which functions of LTE EPC should be moved to cloud in an SDN based system is done in [116]. They pointed out that as functional blocks move to the cloud, the cost of S/P-GWs goes down as the gateway hardware becomes simpler but data overhead and end-to-end latency

increase. The straight forward solution of moving all control functionalities to the cloud may result in unacceptable latency for some applications. Also, cloud infrastructure performance will be a critical factor for high frequency of control plane operations taking place at S/P-GWs.

On the other hand, moving signaling control and resource management logic to the cloud would allow S/P-GWs to be deployed on distributed elements but would also create bottleneck at the centralized control when forwarding rules are exchanged. Moving resource management back to the data plane would make it more independent and resilient but it would not be possible to take the advantage of centralization through optimal resource allocation. The authors also suggested a hybrid approach where the scenario and traffic requirements would dictate the design but this approach would require state synchronization and orchestration between cloud and data plane to avoid redundant assignments.

A comparative cost analysis of SDN-based EPC and software only EPC, where LTE control and user (C/U) plane are both implemented in cloud, is presented in [114]. The key finding of this study is the realization that a pure software solution where C/U planes are in the cloud is much more expensive and it is economical to leave the user plane (S/P-GWs) outside the cloud and perhaps on the dedicated hardware. An optimal solution for placing SGW-C- i.e.- the controller part of S-GW, in the network is presented in [111] using game theoretic approaches. The algorithm finds a trade-off between load reduction on SGW-C and reducing relocation of S-GW which is costly for the mobile operators.

In [109], SGW-C, i.e., the control part of S-GW is used as an application along with MME module and PGW-C (control part of P-GW) with an OpenFlow controller. The data plane, or SGW-D, is comprised of advanced OpenFlow based switches with GTP support. The numerical evaluation show reduction in signaling load with OpenFlow implementation of EPC than the original 3GPP EPC. This work was extended in [110], where all functional blocks of EPC including S/P-GW control part are implemented as application over mobile controller. The signaling load in full implementation of EPC is shown to be lower than that of [109].

B. Control Plane

One of the important issues in SDN-based networks is controller scalability. Clustering or distributed control plane is explored as a potential solution, although, many important questions, such as, inter-controller coordination, etc., remain unanswered. Hierarchical control planes are also proposed to cater for the time-sensitive control processes on one hand and the need for centralized view for optimal operations on the other.

1) *Hierarchical/Distributed Control Plane:* Different proposals are put forward which define domains or clusters for each controller to resolve the scalability issue [81], [113], [117]–[119], [122]–[124]. The controllers can be connected to each other in a distributed manner or there could be a higher level controller to coordinate between the lower level controllers. The size of the domain or clusters can be chosen to

TABLE V
SUMMARY OF RESEARCH IN SDN-BASED MOBILE NETWORK ARCHITECTURES.

SDN Layer	Popular Ideas	References	Key Issues/Solutions/Ideas
Application layer	1- EPC control nodes as applications/functions	[4], [56] [81], [109] [112], [113] [110], [115]	1- S/P-GW split in C/U planes, [2], [109]–[111] 2- S/P-GW CP part is used as application 3- CP/UP split is economically optimal [114]. 4- Centralized control may create bottlenecks [116]
Control layer	1- Distributed control	[113], [117] [118], [119]	1- Delay limit determines cluster size 2- Handover between clusters [118] 3- Optimal controller placement [111], [120] 4- Local controllers for time-critical tasks [3], [121]
	2- Hierarchical control	[81], [122] [123], [124]	1- Different controller for core, RAN, orchestration, etc. 2- Fine-grained control difficult for OpenFlow [119]
	3- Global Network View (GNV)	[125], [126] [115]	1- Store/update network state information 2- Local storage for delay-sensitive data
	4- HetNets/multiRAT	[62], [122] [127], [129]	1- Handover via virtual switches [127] and through virtualization [124], [128] 2- Joint resource allocation [9], [103], [129] 3- Traffic offloading [130] 4- End-to-end QoS [129] 5- Energy efficiency [128]
Data plane	1- CloudRAN	[131], [132] [97], [128] [8], [66] [135], [136]	1- Fronthaul capacity and latency 2- Partial/adaptive centralization [66], [132]–[134] 3- Heterogeneous CRAN [129] 4- xhaul:backhaul & fronthaul [137]
	2- Integration of SDR	[138], [139]	
	3- Exclusively OpenFlow	[4], [8], [57], [104], [107] [64], [109], [113], [140] [119], [142], [143]	1- Compression of policies [59] 2- Support for (GTP) tunneling [57], [140]–[142] 3- Performance requirements/optimization [4], [140]
	4- Control/data split in RAN	[11], [40], [136]	1- Management and virtualization as control modules [11] 2- Switching-off strategies [136]
Integrated virtualization	1- 5 options to integrate SDN in NFVI	[144]–[146] [15], [147]	1- Controller be part of OSS/BSS; 2- exists as an entity in NFVI; 3- exists as physical network function (PNF); 4- be instantiated as VNF [148]; 5- be in VIM (Virtualised Infrastructure Manager) [144].
	2- Integration in SDN	[4], [11] [8], [64] [124], [139]	1- Role of hypervisors [4], [8], [149], [150] 2- Virtualization layer above infrastructure [4], [124], [151] 3- Virtualization services as control plane modules [11]
SDN applications	1- Resource sharing	[39], [104], [105]	1- Co-primary spectrum sharing [105] 2- Via virtualization [104] 3- Via Smart gateways [39]
	2- Mobility management	[102], [152], [153]	1- DMM implementation via X2 [102], [152] 2- Via global network view [153]
	3- Multi RAT scenarios	[130], [154], [155] [156]	1- Traffic offloading [130] 2- End2end routing [154] 3- SDN-based backhaul [155] 4- On-demand small cells [156]

satisfy the delay constraints. Hierarchical control plane proposals suggest 2 or 3-tier architectures for scalability [113], ease in management specially in heterogeneous networks [122], and for fulfilling delay requirements [117], [122].

A distributed controller architecture is presented in [118], where controller-domains are defined as the eNodeBs served by a particular controller. Handover processes are also defined between eNodeBs in a controller domain and within different controller domains. The handover processes and packet sizes are also revised from 3GPP standards to reduce the signaling overhead. The proof-of-concept, i.e., testbed implementation or simulation experiments are missing from this study to ascertain that the design is complete and adequate for 5G capacity and speeds. The controller domain idea makes the SDN control plane scalable but important details about controller placement and optimization are also not available in this paper. Optimal controller placement is also an open issue. Solutions are presented in [111], [120] using optimization function and game theoretic approaches.

SoftRAN [3] also observes that it is cost-effective to leave data plane functionality to the base-stations along with some part of control plane for delay-sensitive decisions, but stressed on the coordination of closely-deployed BS in a dense network though a centralized controller such that the dense base-stations in an area can be viewed as virtual big base-station (big BS).

Similar ideas of a regional controller which possibly resides in cloud and local controller, located in a macro cell, have been used in [121]. The SDN-based mobile network model for dense deployments [121] is an outcome of FP7 project CROWD with the focus on energy optimization. The paper also proposed a mobility management scheme for their model.

The architectures presented in [122], [124] defines a local SDN controller (LSC) for scalability which are responsible for heterogeneous wireless networks in an area and are connected to core SDN controllers (CSC). The paper [124] goes into the details of LSC and what functions it is supposed to perform including content caching, resource and mobility management functions. The authors also showed the backhaul traffic reduction with content caching via simulation experiments. The virtualization layer lies on top of physical infrastructure and each VM (Virtual Machine) communicate to LSC via a local SDN agent. The implementation results are still preliminary and does not show the interworking of different LSCs and does not include offloading mechanisms over heterogeneous technologies.

[81] proposed a three-level control plane structure for SDN-based mobile network. They have included device controller (for UE), an edge controller, implementing 5G complaint network functions through control applications, and an orchestration controller to coordinate utilization of cloud resources. Two versions of edge controller are identified in the paper, one implemented in the cloud and the other in the mobile device for mission-critical control and out-of-coverage services. The paper does not go in the detail of physical structure or fronthaul or backhaul issues rather points out the requirements for the control plane architecture. Following up from [81], the authors in [123] define the initial attachment and service

requests procedure for the 3-level control plane structure. They show that reconfigurability of the network to use appropriate cloud resources can improve the latency up to 75% compared to 3GPP Release 12 complaint 4G systems.

[119] describes a high-level architecture for mobile networks with one controller each for core network and RAN. They use a coordinator to provide end-to-end services through both controllers. Some services, such as, firewall, encryption, etc., are implemented through middle-boxes and the controllers decide the flows traversing through these middle-boxes. The paper identifies major challenges of this approach, such as, providing fine-grained control which is not yet possible with OpenFlow. Moreover, the paper provides preliminary performance results through simulations.

2) *Global network view (GNV)*: The paper [125] introduces the concept of global network view (GNV) which stores information about network state and can be accessed by network services. The storage can be local for delay-sensitive information, such as, channel state information. The rest of the information regarding flows and load can be stored in a common place. The protocol stack in the architecture of [125] is in the modular form which can be orchestrated by the control plane for implementation of specific function, e.g., eNodeB or S-GW.

The use of GNV with locally and globally stored information is also discussed in [126] but only for the context of access network. A similar idea is introduced in [115], where an additional plane called 'knowledge plane' is added on top of the application plane in SDN-based mobile network architecture. This knowledge plane is responsible for providing a global view of network usage which can be used to device new applications to optimize resource utilization. The useful information in the knowledge plane consists of network usage, traffic load, congestion information, etc.

3) *HetNets*: The current mobile network environment, in general, consists of heterogeneous radio technologies, such as, 2G/3G/4G and WLAN. Most likely, the future operating scenarios will also have similar characteristics. SDN has the potential of exploiting the different radio resources in a comprehensive manner as the details of the technology are hidden under the abstraction layer and it is much easier to optimize the utilization of available resources in a unified way [122]. Although, the mobility management, joint resource allocation, end-to-end QoS (Quality of Service) assurance, and security issues pose additional challenges specially when UE moves between networks of different ISPs using heterogeneous technologies [129].

Special network selection switches are proposed to achieve handover between different technologies [127] and virtualization is also used as a mean to connect different networks to a controller [124], [128]. Traffic offloading schemes are also coming up recently [130]. In [103], a resource allocation mechanism with holistic view is introduced for SDN-based heterogeneous networks. Most of the work in this respect remains at conceptual level and more exploration is needed [9], [62], [157]. This is one area where optimal resource utilization solution can provide immediate monetization opportunity.

In [62], a high-level heterogeneous network architecture is

presented with a SDN controller providing network control functionalities to RAN, transport network, service providers, virtual operators, etc. The paper did not go into the details of implementations or low-level architecture of the controller but discusses high-level requirements from different interfaces. The discussion in [157] also remains at high-level in describing the scenarios and benefits of using SDN to control and operate a HetNet.

The approach discussed in [9] proposes a SDN controller cloud providing control to data plane including heterogeneous wireless access systems and BBU cloud for CloudRAN. The SDN controller will provide control to the EPC cloud implemented in GPP as well as physical switches for the core network. The paper provides details about different components of data plane and how they can be specifically controlled with associated issues and benefits. However, the paper still remains a high-level conceptual contribution.

HetSDN [127] also proposed seamless transfer between heterogeneous RANs via virtual switch. They introduced two new components, i.e., SDN-NS (SDN Network Selection) and SDN-HA (SDN Home Agent). SDN-NS uses a virtual switch to seamlessly direct packets to the most suitable wireless interface and SDN-HA is used as the anchor for all traffic for the UE. Another idea to provide seamless movement between LTE and WiFi is through SDA (Software Defined Access) [158]. In this architecture, a separate controller, i.e., SDA controller is added on the control plane which has a client side SDA-u on the UE. The also have another element SDA-g (SDA gateway) on the data path which is used as an anchor. Prototype implementation shows proof of design and concept.

VCell [128] describes an architecture with heterogeneous cells, i.e., macro, micro, femto, etc., where all resources are virtualized in a logically centralized pool. The resources are allocated to the user from a 3D grid of time, space, and frequency resource blocks. The SDN controller assigns resources and manages interference. The main contribution of this paper is to consider heterogeneous cell types while building up the resource grid. The issues with centralization, such as, fronthaul latency and capacity, processing latency of commodity hardware, etc., are not discussed in [128].

Heterogeneous CloudRANs are also studied in literature where small RRUs and macro RRUs are distinguished by their service area [136]. These networks may also contain small or macro base-stations based on the availability of resources and chosen functional split. The macro base-station or macro RRU is shown to provide signaling for the whole service area and small base-stations and RRUs can only be used on-demand optimizing energy efficiency while satisfying QoS constraints.

C. Data Plane

CloudRAN seems to be preferred choice for SDN-based mobile network layout, although, it is not always beneficial to centralize all the processing in the cloud [97], [128]. The case for partial centralization of processing units is discussed under various themes, such as, in RANaaS (RAN as a Service) the level of centralization is chosen according to the need and network state [66], [132]. A plastic or adaptive architecture

which provides various levels of centralization according to the operational requirements [133] could be very desirable from business point of view as it provide further optimization of cost but the technical issues involved in realizing such architecture are not clear at this point. [134] discusses the partial centralization option in their architecture to cater for the fronthaul and backhaul constraints.

[129] pointed out that, in a CloudRAN integrated HetNet, if multiple standards are being used in the same spectrum, the RRUs can support them only partially. Whenever the standard changes, the BBU is forced to restart instead of sharing multi-standards resource directly.

[137] suggests an architecture for integrating existing and new fronthaul and backhaul networks into a flexible unified 5G transport solution using SDN/NFV-based management and orchestration (MANO) technology. The paper explores the resource management functional blocks and discusses some use cases.

Issues and advantages related to the integration of SDR (Software Defined Radio) with SDN are also explored in the literature [138], [139]. Preliminary results have shown that the integrated network with the potential of using spectrum holes and white spaces performs significantly well than the network without SDR [138].

1) *Exclusively OpenFlow based:* OpenFlow seems to be the choice of the researchers at least for now [4], [8], [57], [64], [104], [107], [109], [113], [119], [140], [142], [143]. OpenFlow and SDN may even appear to be synonyms in some places. During our study, we only came across one paper [150] which advocated the use of ForCES (IETF Forwarding and Control Element Separation framework) instead of OpenFlow. ForCES includes abstraction model for both SDN and NFV as described in RFC 5812 [150]. The major issues in using OpenFlow for the data plane in a cellular network are the lack of GTP tunneling support and the transformation of complex billing and management policies into simple and scalable routing rules for OpenFlow switches.

Among the earliest work on SDN-based mobile network, OpenRoads or OpenFlow Wireless presents a prototype where wireless APs (WiFi and WiMax) are augmented with OpenFlow [159]. A centralized controller, Network OS (NOX), is used to host the application plug-ins and translate the policies into routing rules for OpenFlow. Applications of mobility management were executed over the prototype. Authors suggested similar OpenFlow additions for LTE, although, LTE base-station and core network is way more complex than a WiFi or WiMax access point and it is not straightforward to translate management and billing policies into simple routing rules [59]. Softcell [59] remarked that the policies of a cellular network translates into thousands of rules and it is important to come up with some logical aggregation techniques so the rules can be fit into today's routers and switches.

The design in [143] put all the EPC control functions in controller which translates the flow rules for OpenFlow switches used in place of S-GW. P-GW is eliminated from the design. The author claims efficient mobility support with the proposed EPC architecture than the legacy one. The paper overlooked the real problem of scalability and translation of

huge lists of policies into simple flow rules for the switch [59]. Until both problem remain unsolved, it would be difficult to realize the architecture.

[57] added relevant extensions in the OpenFlow protocol for GTP tunneling and encapsulation. In [141], a new dynamic GPRS tunneling protocol is presented and evaluated for SDN based core network. The protocol allows elastic use of data plane resources and enable cloud to provide on-demand packet processing. Terminating GTP tunnel at the cloud would incur additional delays than terminating it on the fast data path. The paper is preliminary work and more evaluation is needed to understand the scalability and performance with real networks.

The integration of OpenFlow in EPC is discussed in [142], where the S-GW is replaced with OpenFlow switches and detailed steps are discussed to keep GTP tunneling and vertical protocol stack for backward compatibility. The paper later introduced different processes like creating, modifying, and deleting a bearer or a session without GTP tunneling for lightweight architecture.

In [140], OpenFlow 1.4.0 is extended and optimized for 3GPP EPC including the design of controller, switch, and protocol. The paper first highlights the requirements including the clear definition of northbound APIs, GTP support, latency, etc. The switch is evaluated within OpenEPC implementation for proof-of-concept and scalability.

CellSDN [4] identifies additional requirements from OpenFlow switches to provide network services, such as, deep packet inspection, header compression, etc. They emphasized on reducing the controller's load by means of local control agents at the base-station and switches. The interfaces between controller, local agents, and data plane are also defined in terms of requirements and available protocols.

2) *Control signaling and data split*: The approach called CONCERT [11] separates control signaling from data transmission as defined in BCG2 architecture and Phantom cell concept (cf. Section II-C). It uses SDN switches with all radio interfaces, computational resources, and servers. These switches are then controlled by the control plane entity called a conductor. All management and virtualization services are defined as control plane modules. Application scenarios are highlighted to focus on the benefits of SDN-based architectures. Technical issues are also briefly explored.

The authors in [40] also touched upon the control-data split in their deployment scenarios but did not go into the details of implementation. [136] also discussed the idea in relation to heterogeneous CloudRANs, where macro base-station or RRU can provide signaling to the area and small base-stations or RRUs can only be used on-demand. The switching-off strategies are also discussed in [136].

D. Integrated Virtualization

1) *SDN from NFV Perspective*: As mentioned earlier SDN and NFV are complimentary technology enablers for 5G and [146] provides an overview of both these technologies in the context of 5G network requirements and architecture. From NFV perspective there are different options for deploying SDN within the NFV infrastructure and the various deployment

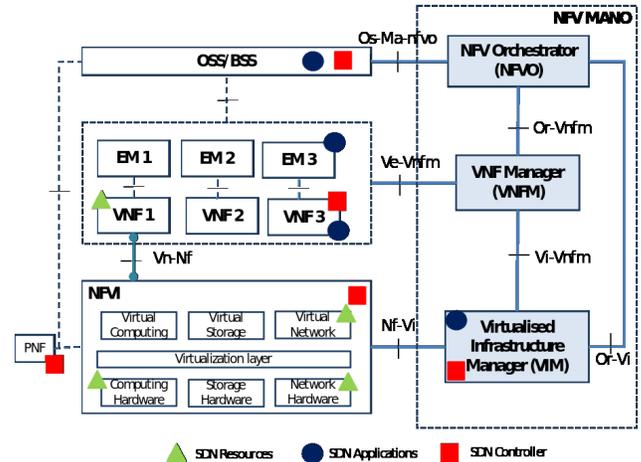


Fig. 9. Possible options of positioning SDN Resources, SDN Controller and SDN Applications in NFV Architectural Framework [146].

options have been analyzed in detail by ETSI NFV in [145]. Figure 9 [146] summarizes the possible options of integrating SDN application, SDN resources and SDN controller with different entities within the NFV MANO and NFV architecture. There are five integration options for SDN controller to either (i) be part of OSS/BSS (Operations/Business Support Systems), (ii) exist as an entity within NFVI (NFV Infrastructure), (iii) exist as a PNF (Physical Network Function), (iv) be instantiated as a VNF (Virtual Network Function), or (v) be integrated within the VIM (Virtualised Infrastructure Manager). The latter approach, for example, is supported by the ONF SDN architecture [144] and is also an adopted approach by open source OPNFV platform³⁹. OPNFV prescribes the integration of SDN controllers like ODL (OpenDayLight) and ONOS (Open Network Operating System) with OpenStack VMM (Virtual Machine Manager). It should be noted that from ETSI NFV perspective, OpenStack platform for all practical purposes is considered as a VIM platform.

There are also some prominent research projects like 5G NORMA⁴⁰ that leverages on the SDN and NFV concepts in order to develop a novel mobile network architecture that shall provide the necessary adaptability in a resource efficient way able to handle fluctuations in traffic demand resulting from heterogeneous and dynamically changing service portfolios and to changing local context.

2) *NFV in SDN Domain*: As the SDN architectural proposals are concerned, NFV integrated SDN architectures make more sense from business as well as technical points of view [4], [8], [11], [64], [124], [139]. Designs may include the role of hypervisors [4], [8], [149], [150] or virtualization layer above the infrastructure or forwarding plane [4], [124], [151]. In CONCERT [11], the virtualization and management services are provided as control plane modules. These archi-

³⁹<https://www.opnfv.org>

⁴⁰<https://5gnorma.5g-ppp.eu/>

tures are high-level designs and the relevant issues are not fully explored.

One of the initial attempts to combine NFV and SDN technologies together for mobile EPC is given in [148]. It shows the complexities and issues involved in achieving the virtualization capacity in a softwarized core network. Firstly, the authors show the use of OpenFlow switch network to replace S/P-GWs of EPC. Important questions left unanswered here, such as, how the huge lists of policies in PCRF can be translated into manageable flow rules for the OpenFlow switches. Secondly, the combined NFV and SDN architecture is proposed which shows two different instantiations of the control plane, i.e. one for NFV and the another for end-to-end connectivity. The two instantiations are connected to each other and also to the NFV management and orchestration unit, at supposedly higher control level. The EPC control plane is also mapped as VNF (Virtual Network Functions) in the NFV infrastructure. Many issues are unanswered regarding the design. Most importantly, how these VNFs will interact with SDN controller to maintain connectivity. Scalability is another issue which needs to be tackled in a centralized EPC architecture. A layered approach providing clear demarcation between actions needed for connectivity and providing services is probably a better way to understand the problem instead of modeling a broad set of NFV and SDN aspect into one flat design as also mentioned in [148].

SoftAir [8] proposed an SDN-based mobile network architecture with network hypervisor to orchestrate virtualization. They have proposed wireless hypervisor and switch hypervisor for SDN enabled base-station and SDN enabled switches. They have designed SDN enabled base-station such that all processing stays at the cloud except modulation/demodulation which are located with antennas at RRUs. The split would reduce the load on CPRI supported fronthaul.

The paper [150] advocates the use of ForCES (IETF Forwarding and Control Element Separation framework) instead of OpenFlow. The authors build a proof-of-concept focusing on GTP tunneling. The proof-of-concept involved a hypervisor which instantiate VMs and deploy necessary LFB (Logical Forwarding Block), an element of the forwarding plane in ForCES. They build an application to provision and destroy GTP tunnels and collect statistics. The authors claim that the framework can scale-in and out as required but there are not analytical results presented in the paper to support the claim. Proof-of-concept was shown to the IETF meetings but no results are presented in the paper.

A high-level architecture is also presented in [151], where a network virtualization layer is laid over the physical infrastructure and under the service enablement layer which seems to be very close to control plane in concept. Architecture is using cognitive and cloud optimization domain and is named as CONE architecture. As it is also mentioned in the paper, many details are left out of the paper and only preliminary design is discussed.

The proposed SDN-based mobile network architecture in [64] moves EPC to cloud to benefit from NFV while maintaining 3GPP interfaces with legacy nodes. It suggests moving some functionalities of eNodeB, if required, to the cloud

but the prototype uses the unchanged eNodeB. The proposed architecture is deployed over a testbed using two eNodeB from Nokia, OpenFlow enabled MPLS switch, open source S/P-GW and probing tools. The results show downtime of up to 2 seconds in live migration of a MME running on a VM instance.

Mobileflow [56] is one of the earliest solutions to resolve the compatibility issue between SDN-based networks and legacy systems using virtualization techniques. OpenRAN [7], [160] is also a high level idea of a SDN-based mobile network through virtualization. A SDN controller would first create virtual RRU and virtual BBU using the physical resources and then dynamically optimize them according to the application.

[161] develops an interesting concept of Hybrid-eNodeB (HeNB) which hosts a virtualized EPC (vEPC) along with all eNodeB functionalities. vEPC is the most fundamental version of EPC without the complexities of detailed control. vEPC needs the connections to physical EPC, although not all the time, to get the necessary information about billing, roaming, etc. HeNB greatly simplifies the deployment of 4G LTE and even allow standalone and self-powered UAVs (Unmanned Aerial Vehicles) and LAPs (Low Altitude Platforms) to host HeNB. [149] presents the idea of virtualized EPC using a general purpose node (GPN). GPN is a core-class server with a hypervisor to provide virtual instances of EPC. The main purpose is to reduce cost through improvement in resource utilization. EPC nodes and interfaces remain unchanged in this work along with GTP tunneling.

In a recent publication [139], the authors presented an architecture for WNV (Wireless Network Virtualization) which is similar to NFV concept but only for access networks. Their architecture places the virtualization layer or hypervisor as southbound interface between the physical network components or the data plane and the virtual resource management or the control plane. The virtual network functions are placed in the application plane. The paper also listed the outstanding issues and advantages of such an architecture.

The proprietary virtualized access (vAccess) development platform from Freescale [48] is created to facilitate the development of SDN/NFV products. It is built over OpenStack cloud computing platform with hardware accelerators for L1 real-time processing and Linux patch to support bounded guarantees for application start time. The vAccess platform provides the basic structure to build LTE base-station software.

In 2013, Huawei [131] unveiled SoftCom strategy. SoftCom is a holistic approach to network architecture based on cloud computing, SDN, and virtualization. The four key elements of SoftCom are cloudification of equipment through separation of hardware and software, cloudification of network through separation of forwarding and control, transforming traditional infrastructure to a cloud infrastructure, and transforming Telecom-oriented systems to Internet-oriented systems. For RAN, SoftCom is focusing on moving the control plane for small cells to macro base-station and centralization of BBU processing as in CloudRAN. SoftCom promises to develop an eco-system for simplification of network operations and innovation of open business models.

E. SDN Applications

In this section, the discussion about application of SDN in improving and optimizing network management is included only to admire how the SDN research space is expanding and focus is turning towards details. There is huge amount of work in this respect specially since 2014, and it is outside the scope of the present paper to provide a comprehensive survey on applications.

1) *Resource Sharing*: A framework for co-primary spectrum sharing among MNOs (Mobile Network Operators) is developed in [105]. The framework is based on an SDN architecture integrating Mobileflow [56] and SoftRAN [3]. Co-primary spectrum sharing enables optimal resource utilization among different operators. Mobileflow [56] is the core network part of this scheme while SoftRAN [3] is evolved to enable access network sharing. The big base-station concept of SoftRAN is used in the paper for all base-stations of all operators. The logically centralized controller then assigns the resources based on the resource management rules. The scheme does not go into the details of SLA (Service Level Agreements) between MNOs and does not identify the business incentives and benefits. Security issues are also left out of the paper.

Radio resource sharing is also discussed in [104], where virtualization of LTE eNodeB is used to facilitate lease on-demand the physical infrastructure and resources of one operator to the flows of another operator. The framework is named OpeNB (Open eNodeB) and it is based on SDN and OpenFlow technologies. Besides OpenFlow controller, another controller called OpeNB controller provides signaling and management control for resource sharing. A main controller provides hand shaking between networks of different operators. System-level simulations show improved performance in terms of packet drop rate and throughput.

Another proposal for SDN-based smart gateways to connect S/P-GW of multiple operators is presented in [39]. Here a SDN orchestrator is responsible to manage multiple smart-gateways in order to provide demand-based uplink capacity to the connected small cells. The smart-gateways allow resource sharing among multiple operators through SDN orchestrator.

2) *Mobility Management*: In [102], the authors looked at the requirements of DMM (Distributed Mobility Management) and how SDN architecture can satisfy them. They also showed a simple example of how DMM can be implemented in a virtualized LTE environment. Their follow-up paper also provides simulation results of SDN-based DMM implementation [152]. Their performance evaluation results show that seamless mobility management is possible when X2 path is used between eNodeBs.

The benefits of SDN in mobility management with global network view and flow-based control are highlighted in [153]. The authors discussed mobility management in SDN architectures with different centralization levels. They identified issues regarding session continuity and scalability of handover and showed how mobility management can be easily integrated with autonomic management mechanisms to optimize resource utilization.

3) *Multi RAT scenarios*: A SDN enabled traffic offloading scheme is proposed in [130]. A SDN controller runs software

based RRM (Radio Resource Management) and PCRF (Policy Charging and Rules Function) modules to determine the offloading from LTE to WiFi depending on the network state and flow requirements. If both networks are unable to satisfy traffic requirements, the packets are dropped. Simulation results show 15% less packet drop with the SDN-enabled offloading scheme. RRM is responsible for updating the network state information and PCRF is responsible for policy enforcement. The paper shows a very important use case where SDN can provide performance improvement in a straight forward manner.

A detailed sub-optimal online solution for end to end routing in mobile networks based on SDN and cloud computing is presented in [154]. The objective is to maximize the amount of traffic accepted over time. The problem is formulated and an online algorithm is developed and tested via simulations for multi RAT (Radio Access Technology) scenario.

A wireless/wired backhaul solution based on SDN is presented in [155] for small cells. The prototype performance is compared against 802.11s and significant improvement is observed when channel conditions are used to decide which packets to transmit over the backhaul. Similar idea is also presented in [156] where some details of the on-going work on implementation are also discussed. An SDN controller can adaptively power on/off small cells based on the demand and traffic metrics.

F. Outstanding Issues

Recent years have seen massive growth in SDN related research but there are still many open questions. The main focus of the present paper is on the major challenges. In this section, we limit ourselves to the open issues and problems in architectural domain only. Summarizing the literature in the current section, we identified the following important issues which should be looked at by the research community.

- 1) **Policy compression** PCRF policies and rules need a consistent aggregation mechanism to be able to translate into routing rule for commercially available and OpenFlow-enabled routers and switches.
- 2) **Access network layout** CloudRAN seems to be the choice for SDN-based access network but the fronthaul capacity and latency requirements create big hurdles in the production deployment. 3GPP's LTE Release 12 also defines Phantom cells to realize the densification of base-stations. Thorough investigation is needed to find the optimal alternative for different operating scenarios.
- 3) **Signaling/data split** Separation of control signaling and data transmission is proposed to achieve huge energy efficiency in BCG2 architecture and Phantom cell concept (cf. Section II-C). We found only three papers which proposed to integrate the idea into a SDN-based mobile network architecture [11]. The 85-90% saving potential, as reported in [19], is significant enough to encourage more exploration in this area. While designing the architecture for a new system, it may not be difficult to adopt the signaling/data split paradigm.

- 4) **Level of centralization** Moving all functionalities to cloud may not be a desirable choice for most of the operating scenarios. One reason is the limitations posed by GPP in processing real-time signals. Another is the inhibitive cost of fiber deployment if it is not already available. This area is being explored, mostly in terms of cost, but comprehensive investigation is still missing which can also look at the partial centralization scenarios and quantify the negatives, such as, the inability to carry out CoMP etc.
- 5) **Scalability** Inter-cluster communication between the controllers responsible for clusters of base-stations has not been explored at all.
- 6) **GPP implementation** EPC control nodes are suggested to be implemented as applications but carrier grade implementations are needed to ascertain the realizability of the idea.
- 7) **Edge computing** An opposite concept to centralization, edge computing also brings promise of performance enhancement and cost savings. It is vital to bring the two ideas together to see how they interplay and how practical it is to develop a plastic architecture that could be optimized for a given operational scenario.
- 8) **Integrated SDN and NFV** This area needs more investigation. Specially with practical limitations, we may not see fully centralized and fully software based system at least in the beginning. How these limitations affect the realization of NFV would be important and in fact may guide the SDN designs and development.

V. RELATED WORK

Although, the development in the area of SDN-based mobile networks is trailing behind the work in wired and fixed networks but there are still a number of survey papers available in this domain. Now the question is, why do we need one more? In this section, we present a brief summary of the available surveys and while they all target to explore some aspect of the complex problem and moves forward the collective understanding of the research community, none has focused on the issues delaying the realization of SDN-based mobile systems. As per our knowledge, this is the first attempt to provide a holistic picture of the state of the art with emphasis on the weakest links.

The survey in [162], although is relatively recent but, presents a case for SDN by highlighting the issues with the current networks, such as, heterogeneous access networks, network ossification, increase in cost for operator amid dwindling revenues, etc., and how SDN can provide solutions. The paper, then, summarizes work in SDN and NFV for mobile and wireless networks and identifies open technical issues, such as, heterogeneous system's support, decomposition of control functions from protocol layers, scalability, customized and open API development, virtual machine migration, etc. The paper also advocates the joint design of SDN and NFV. SDN provides unprecedented visibility into the network allowing virtualization functions to have a clean abstraction to slice the resources.

The survey in [163] discusses all SDN related research in wireless domain, including wireless LAN, Cellular, mesh networks, sensor networks, etc. Three publications- SoftCell [59], CellSDN [4], and SoftRAN [3]- related to cellular networks are mentioned in the survey, not surprising for such an early survey paper. The paper mostly lists the publications for each class of wireless networks and identifies the respective advantages and technical issues, such as, latency and channel variability for power allocation.

An almost similar list of publications in cellular domain appeared in [164], i.e., SoftCell [59], SoftRAN [3], Mobileflow [56], and OpenRadio [6]. This paper discusses the benefits, SDN brings to the access networks, such as, virtualization, interference management, and mobile traffic management. The challenges and issues in realization of SDN-based mobile networks are also briefly discussed, such as, actual cost saving margins, scalability of controller, security, isolation between network slices, estimation of channel load for proper resource allocation, and handover between different service providers. Similar survey of SoftCell [59], SoftRAN [3], Mobileflow [56], and CROWD project [121] can be found in [165]. The paper also discussed open research problems such as SDN-enabled cross-layer MIMO and heterogeneous radio technologies.

The survey in [166] has a brief section on architectural designs of SDN-based mobile networks and they have classified them according to their closeness to CloudRAN or mobile edge computing. This paper is rather an overview of different issues including a brief history, high-level discussion about business case and challenges and technical problems, and standardization activities in this space.

In a recent and most comprehensive survey of SDN and virtualization research for LTE mobile networks [65], the authors have provided a general overview of SDN and virtualization technologies and their respective benefits. They have developed a taxonomy to survey the research space based on the elements of modern cellular systems, e.g., access network, core network, and backhaul. Within each class, the author further classified the material in terms of relevant topics, such as, resource virtualization, resource abstraction, mobility management, etc. They have also looked at the use cases in each class. It is the most comprehensive survey one could find in the radio access network research relevant to SDN. The thrust of the survey is complementary to the present paper. If the readers want better understanding about the material covered in Section IV, they are recommended to read [65]. On the other hand, the open challenges briefly discussed at the end of [65] and the relevant work under each challenge are discussed in detail in the present paper.

The survey in [167] covers the holistic wireless space including cellular, WLAN, mesh, sensor, and home networks. In cellular networks, the major contribution of this paper is to cover only the significant papers in RAN and core network architectures with the focus on the inclusion of virtualization in the design and use of OpenFlow. The present paper provides much deeper discussion about the work in cellular SDN space and a much comprehensive coverage of the research area.

A much recent survey [168] covers the proposals for

SDN/NFV architectures for EPC, i.e., core network only. The authors found three different approaches to re-architect EPC, i.e., virtualizing EPC (vEPC) using NFV, decoupling control and user planes in vEPC with SDN technology, and fully SDN realized core network. They then classified the proposals according to the four attributes, namely, 1- architectural approach (evolutionary or revolutionary), 2-technology adoption (full or partial adoption of SDN or NFV or integrated NFV/SDN technology), 3- functional implementation (migration to VMs, decomposition of functions, merging of multiple functions into single unit), and 4- deployment strategy (distributed or centralized).

Another recent survey [169] briefly surveys all technologies and applications associated with 5G. The survey also touches upon SDN and only superficially covers some research work under the theme. A more in-depth analysis of some of SDN-based mobile network architectures, i.e., [3], [4], [6], [8], [56], [59], [121], [138], [159], are presented in [170] in terms of ideas presented in the proposals and their limitations. The survey in [171] looks at the proposals for softwarization and cloudification of cellular networks in terms of optimization and provisions for energy harvesting for sustainable future. The gaps in the technologies are also identified. All of the above mentioned surveys, however, have a broader scope than just SDN-based mobile network architecture and they have only looked at some SDN papers appropriate for the major theme of their survey papers.

VI. SUMMARY

In this paper, we argue that transforming the current mobile network infrastructure to a SDN-enabled architecture may take more time than expected. With 5G around the corner (5G NR (New Radio) non-standalone standard approved in December 2017 and full standard is expected in 2018), it is likely that we may just see some point solutions based on SDN rather than a full scale roll-out of the technology. We identified six major roadblocks in realization of SDN-based mobile networks, i.e., fronthaul, latency of general purpose platforms, backward compatibility, disruptive deployment, SDN specific security vulnerabilities, and clear and compelling business case. We also looked at the state of the solutions under each issue. The major lessons learned from our study are summarized in the following subsections.

A. Fronthaul

- 1) Fronthaul's capacity and latency requirements need optical fiber links between the remote radio units (RRU) and the centralized processing unit. This is perhaps the biggest hurdle facing SDN deployments in places where fiber is not available. Countries with rich fiber coverage are in an advantageous position.
- 2) Microwave can be used in place of optical fiber but only for low rate CPRI (upto 2.5Gbits/s, future advances may support up to 10Gbps⁴¹)

⁴¹<http://www.cablefree.net/wirelesstechnology/4glte/cpri-front-haul-technology/>

- 3) CPRI, an internal base-station protocol to transport digitized waveform from cabinet to antenna heads, is the most widely used transport protocol for RRU-cloud connection but it has the worst data transport efficiency e.g., a 20MHz LTE channel can carry up to 150Mbps in downlink but requires CPRI rate of 2.5Gbps [23].
- 4) Other than capacity, delay and jitter are important issues for fronthaul. According to [25], delay should remain under 100 microsecond and jitter should be under 65 nanoseconds for CPRI.
- 5) Functional split is a promising solution to cater for fronthaul capacity constraints. It is highly probable that layer 1 and part of layer 2 would remain at the remote antenna sites, i.e., at RRUs and the rest of the layered protocol would move to the cloud. There is no clear winner in this case. You loose multiplexing gain when move away from full centralization option. Operating conditions and fronthaul constraints are the main drivers to choose the split option. Chinamobile's NGFI [29] is also a promising option in practical deployments.
- 6) Use of Ethernet in fronthaul with or without CPRI is also very likely in future as it is a mature transport technology with much better efficiency than CPRI.
- 7) The most important development in this regard is IEEE 1914 working group (NGFI). The working group has two active projects: IEEE 1914.1 is studying NGFI of Chinamobile and IEEE 1914.3 is looking at encapsulation of digitized radio signal (I/Q samples) into Ethernet frames for fronthaul transmission.

B. Latency of GPP

GPPs are not designed for real-time signal processing and the delays could be of the order of tens of microseconds [42]. There has been significant efforts in implementing the base-station components in GPP. So far, the data rates up to 43.8Mbps are supported on a 20MHz channel by an implementation of LTE uplink Rx PHY in GPP by Microsoft [42]. It is still a long way to go as far as 5G goals of 1000x capacity and 100x data rate are concerned.

Hybrid designs with some hardware accelerators for highly-computational blocks, e.g., turbo coding may provide a practical approach to meet the targets. GPP latency is shown to have decreased with processor speed and number of processor cores.

C. Backward Compatibility

A practical realization of SDN should be able to work with the legacy systems specially 4G systems. The major property to hold is to keep the LTE interfaces and GTP tunneling in the new system as well. From innovation point of view, it is not a good idea as with time better alternatives are also surfacing e.g., MPLS labeling provides a faster alternative than GTP tunneling. Moreover, there has not been any work on UE handover between the two systems and regarding their co-existence in the same geographical area. The recently approved 5G NR non-standalone standard defines additional interfaces

and an evolved LTE eNodeB, i.e., eLTE eNB, to provide backward compatibility to 4G systems [2].

D. Disruptive Deployment

An evolutionary approach for deployment is considered better and acceptable to the network owners and operators as any disruption would cause huge revenue loss. The ideal deployment trajectory will consist of point solutions leading to the full SDN-based systems. As seen before, the LTE interfaces and GTP tunneling are necessary for interworking with legacy systems and are kept in all proposals for evolutionary deployments studied in this paper. Most of the ideas, however, still remains at a conceptual level and prototype implementation are necessary to ascertain their workability.

E. SDN Specific Security Issues

Security is perhaps the most researched area in SDN domain. Although, the development in wired networks are more advanced than in the wireless networks. SDN brings additional vulnerabilities to the network but at the same time provide enormous ease in controlling the network as well as visibility throughout the system. Exploiting these properties, better security systems can be introduced and implemented. The salient points from our study are given below.

- 1) Controller is the most critical component of the SDN network and its compromise would give tremendous access to the attacker. DoS or DDoS attack on the controller would also be extremely devastating. Distributed control plane is an important paradigm for reliable and scalable network control which should be explored in detail and the inter-controller interfaces should be defined.
- 2) There is still lack of standardization for secure northbound and southbound interfaces. A dedicated link can connect controller to its switches but appropriate level of encryption is needed to provide secure communication. The monetary implications of dedicated link and feasibility should be studied.
- 3) Programmability of the system may lead to policy inconsistencies and configuration errors. There are real-time tools for policy control resolution but they have scalability issues while dealing with large applications or large networks.
- 4) Independent experimental evaluations and comprehensive assessment of solutions are required to check them for robustness, interoperability with third-party applications, integration of virtualization, and fast and dynamic reconfiguration.
- 5) GPPs should also be hardened so the weaknesses in the processors cannot be exploited to gain control over the network or damage it. Moreover the current controller implementations, i.e., Floodlight, OpenDayLight, POX, and Beacon, also have resilience issues and common application bugs are enough to crash them.
- 6) Real-time tools are needed to verify complete logical isolation between the virtual network sharing the same physical resources.

- 7) It is shown in [72], that there are no solution till date for data leakage and modification in SDN networks. However, encryption is sought to be a valid solution for SDN networks as in traditional networks but the physical separation between controller and switches requires very important control information to be pass over the communication links. Their security should be comprehensively tested for the available encryption standards.
- 8) Most of the development is focused on OpenFlow and should be extended for broader SDN deployments. Moreover, OpenFlow cannot handle switch mobility yet.
- 9) Security research related to SDN-based mobile networks is lagging behind. Wireless shared medium, mobility between different access technologies and different operators' network, constrained computational capacity, and backward compatibility to legacy networks are termed as important security issues relevant only to mobile networks.

F. Clear and Compelling Business Case

Virtualization, use of GPP instead of specialized hardware, and resource pooling in the cloud are the key ideas behind CAPEX and OPEX savings in SDN networks. Actual savings depend on the operating conditions, such as, availability or non-availability of optical fiber fronthaul, traffic load, equipment cost etc. In some business studies, 68% lower CAPEX and 67% lower OPEX are reported when LTE EPC is virtualized. Although, virtualization is not dependent on SDN and immediate gains are possible when traditional methods are used for virtualization. SDN, however, provides a sustainable business model for long-term future.

Some analyst also believe that the cost reduction due to GPP is based on the assumption that specialized hardware is over-priced which is not always true. Considering the cost of transforming telco companies to suit the new SDN paradigm, the cost reduction may not be as great as predicted in some studies.

However, this has also been pointed out clearly that the real benefit of SDN is not in the cost savings but it is in providing the fail-fast capability to the networks. Carriers with such capability, where the deployment cycle for new technologies can be cut down to hours from days, will have a unique edge over others and eventually nobody could survive without it.

The disruptive model of transformation of communication network with full softwarization and centralization is very unlikely to be followed as none of the saving promises and monetary benefits can justify them. An incremental approach where SDN is introduced through point solutions in an already erected network seems to be more plausible.

G. Current State of the Art

In our summary of architectural developments in SDN-based mobile networks, we list the repeated themes in the literature, such as, EPC nodes are mostly featured as application plane modules, SDN-based mobile networks may be close to CloudRAN in design with partial centralization where full

centralization is costly, distributed control plane is scalable, OpenFlow is preferred choice of researchers at least for now, and an integrated SDN/NFV design is more desirable from business point of view. Outstanding issues are discussed in Section IV-F.

There is still a gap between the state of the art and the point where SDN would be tech-ready for production deployment. SDN provides a novel and challenging playing field for research community but it is important to get the priorities right and emphasis should be on solving the problems which matter the most.

ACKNOWLEDGMENT

Partially funded by the EC H2020-ICT-2014-2 project 5G NORMA (www.5gnorma.5g-ppp.eu)

REFERENCES

- [1] C. Gallon, "Carrier software defined networking," Ofcom, Tech. Rep., March 2014. [Online]. Available: https://www.ofcom.org.uk/_data/assets/pdf_file/0013/32143/sdn_report.pdf
- [2] 3GPP, "3gpp ts 23.501 v15.0.0 (2017.12)-system architecture for the 5g system, release 15," Tech. Rep., December 2017. [Online]. Available: <http://www.3gpp.org/DynaReport/23501.htm>
- [3] A. Gudipati, D. Perry, L. E. Li, and S. Katti, "Softran: Software defined radio access network," in *Proceedings of the Second ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking*, ser. HotSDN '13. New York, NY, USA: ACM, 2013, pp. 25–30. [Online]. Available: <http://doi.acm.org/10.1145/2491185.2491207>
- [4] L. Li, Z. Mao, and J. Rexford, "Toward software-defined cellular networks," in *Software Defined Networking (EWSDN), 2012 European Workshop on*, Oct 2012, pp. 7–12.
- [5] S. Bhaumik, S. P. Chandrasekhar, M. K. Jataprolu, G. Kumar, A. Muralidhar, P. Polakos, V. Srinivasan, and T. Woo, "Cloudiq: A framework for processing base stations in a data center," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, ser. Mobicom '12. New York, NY, USA: ACM, 2012, pp. 125–136. [Online]. Available: <http://doi.acm.org/10.1145/2348543.2348561>
- [6] M. Bansal, J. Mehlman, S. Katti, and P. Levis, "Openradio: A programmable wireless dataplane," in *Proceedings of the First Workshop on Hot Topics in Software Defined Networks*, ser. HotSDN '12. New York, NY, USA: ACM, 2012, pp. 109–114. [Online]. Available: <http://doi.acm.org/10.1145/2342441.2342464>
- [7] M. Yang, Y. Li, D. Jin, L. Su, S. Ma, and L. Zeng, "Openran: A software-defined ran architecture via virtualization," *SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 549–550, Aug. 2013. [Online]. Available: <http://doi.acm.org/10.1145/2534169.2491732>
- [8] I. F. Akyildiz, P. Wang, and S.-C. Lin, "Softair: A software defined networking architecture for 5g wireless systems," *Computer Networks*, vol. 85, pp. 1 – 18, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1389128615001632>
- [9] J. Lai, L. Jiang, M. Lei, A. Abdollahpouri, and W. Fang, "Software-defined cellular networking: A practical path towards 5g," *Int. J. Commun. Netw. Distrib. Syst.*, vol. 14, no. 1, pp. 89–105, Nov. 2015. [Online]. Available: <http://dx.doi.org/10.1504/IJCNSD.2015.066019>
- [10] S. M. Raza, D. S. Kim, and H. Choo, "The proposal for sdn supported future 5g networks," in *Proceedings of the 2014 Conference on Research in Adaptive and Convergent Systems*, ser. RACS '14. New York, NY, USA: ACM, 2014, pp. 180–185. [Online]. Available: <http://doi.acm.org/10.1145/2663761.2664237>
- [11] J. Liu, T. Zhao, S. Zhou, Y. Cheng, and Z. Niu, "concert: a cloud-based architecture for next-generation cellular systems," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 14–22, December 2014.
- [12] J. Wu, S. Guo, J. Li, and D. Zeng, "Big data meet green challenges: Greening big data," *IEEE Systems Journal*, vol. 10, no. 3, pp. 873–887, Sept 2016.
- [13] Z. Zaidi, V. Friderikos, and M. A. Imran, "Future ran architecture: Sd-ran through a general-purpose processing platform," *IEEE Vehicular Technology Magazine*, vol. 10, no. 1, pp. 52–60, March 2015.
- [14] S. Abdelwahab, B. Hamdaoui, M. Guizani, and T. Znati, "Network function virtualization in 5g," *IEEE Communications Magazine*, vol. 54, no. 4, pp. 84–91, April 2016.
- [15] ETSI NFV ISG, "GS NFV-MAN 001 V1.1.1 Network Function Virtualisation (NFV); Management and Orchestration," Dec. 2014.
- [16] Alcatel-Lucent, "The lte network architecture: A comprehensive tutorial," in *White Paper ID CPG0599090904, published as a chapter in LTE The UMTS Long Term Evolution: From Theory to Practice (Wiley 2009)*, 2009. [Online]. Available: www.alcatel-lucent.com/4g-consumer-communications
- [17] G. Rittenhouse, "Green wireless networks," in *Keynote presentation in IEEE WCNC 2012*, April 2010. [Online]. Available: <http://wcnc2012.ieee-wcnc.org/Keynote.html>
- [18] H. Ishii, Y. Kishiyama, and H. Takahashi, "A novel architecture for lte-b :c-plane/u-plane split and phantom cell concept," in *Globecom Workshops (GC Wkshps), 2012 IEEE*, Dec 2012, pp. 624–630.
- [19] I. Godor, "D3.3: Final report on green network technologies," INFOS-ICT-247733 EARTH, Tech. Rep., June 2012. [Online]. Available: <https://www.ictearth.eu/publications/deliverables/deliverables.html>
- [20] T. N. L. (Aricent), "Cloud ran: Benefits of centralization and virtualization," in *White Paper, Aricent*, 2016, White Paper. [Online]. Available: <https://www.aricent.com/whitepapers/cloud-ran-benefits-centralization-and-virtualization>
- [21] J. Wu, "Green wireless communications: from concept to reality [industry perspectives]," *IEEE Wireless Communications*, vol. 19, no. 4, pp. 4–5, August 2012.
- [22] Q. Luo, W. Fang, J. Wu, and Q. Chen, "Reliable broadband wireless communication for high speed trains using baseband cloud," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, p. 285, Sep 2012. [Online]. Available: <https://doi.org/10.1186/1687-1499-2012-285>
- [23] K. Murphy, "Centralized ran and fronthaul," in *White Paper, Ericsson Inc.*, May 2015, White Paper. [Online]. Available: http://www.ospmag.com/files/pdf/whitepaper/C-RAN_and_Fronthaul_White_Paper.pdf
- [24] M. Peng, C. Wang, V. Lau, and H. V. Poor, "Fronthaul-constrained cloud radio access networks: insights and challenges," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 152–160, April 2015.
- [25] D. Chitimalla, K. Kondepudi, L. Valcarenghi, M. Tornatore, and B. Mukherjee, "5g fronthaul-latency and jitter studies of cpri over ethernet," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 2, pp. 172–182, Feb 2017.
- [26] N. J. Gomes, P. Chanclou, P. Turnbull, A. Magee, and V. Jungnickel, "Fronthaul evolution: From {CPRI} to ethernet," *Optical Fiber Technology*, vol. 26, Part A, pp. 50 – 58, 2015, next Generation Access Networks. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1068520015000942>
- [27] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: a new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 126–135, December 2014.
- [28] O. C. A. W. Group, "Opfwp001.0216- openfog architecture overview," in *White Paper, OpenFog Consortium*, February 2016, White Paper. [Online]. Available: <https://www.openfogconsortium.org/wp-content/uploads/OpenFog-Architecture-Overview-WP-2-2016.pdf>
- [29] C. I. I, Y. Yuan, J. Huang, S. Ma, C. Cui, and R. Duan, "Rethink fronthaul for soft ran," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 82–88, September 2015.
- [30] N. Nikaein, "Processing radio access network functions in the cloud: Critical issues and modeling," in *Proceedings of the 6th International Workshop on Mobile Cloud Computing and Services*, ser. MCS '15. New York, NY, USA: ACM, 2015, pp. 36–43. [Online]. Available: <http://doi.acm.org/10.1145/2802130.2802136>
- [31] A. S. Thyagaturu, Z. Alharbi, and M. Reisslein, "R-FFT: function split at IFFT/FFT in unified LTE CRAN and cable access network," *CoRR*, vol. abs/1708.08902, 2017. [Online]. Available: <http://arxiv.org/abs/1708.08902>

- [32] G. Mountaser, M. L. Rosas, T. Mahmoodi, and M. Dohler, "On the feasibility of mac and phy split in cloud ran," in *(to appear) IEEE Wireless Communications and Networking Conference 2017*. IEEE, March 2017.
- [33] A. Checko, A. P. Avramova, M. S. Berger, and H. L. Christiansen, "Evaluating c-ran fronthaul functional splits in terms of network level energy and cost savings," *Journal of Communications and Networks*, vol. 18, no. 2, pp. 162–172, April 2016.
- [34] C. Y. Chang, R. Schiavi, N. Nikaiein, T. Spyropoulos, and C. Bonnet, "Impact of packetization and functional split on c-ran fronthaul performance," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–7.
- [35] C. Y. Chang, N. Nikaiein, and T. Spyropoulos, "Impact of packetization and scheduling on c-ran fronthaul performance," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–7.
- [36] A. D. L. Oliva, X. C. Perez, A. Azcorra, A. D. Giglio, F. Cavaliere, D. Tiegelbekkers, J. Lessmann, T. Haustein, A. Mourad, and P. Iovanna, "Xhaul: toward an integrated fronthaul/backhaul architecture in 5g networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 32–40, October 2015.
- [37] G. E. T. G. N. N. M. M. R. A.-O. C. Lanzani, P. Tumber, "Fronthaul evolution toward 5g: Standards and proof of concepts," in *Design & Reuse*, 2016, Industry Article. [Online]. Available: <https://www.design-reuse.com/articles/40008/fronthaul-evolution-toward-5g-standards-and-proof-of-concepts.html>
- [38] M. Z. Shakir, M. A. Imran, X. Wang, J. Wu, A. Ghosh, H. Lundqvist, and L. Liu, "Smart backhauling and fronthauling for 5g networks: from precoding to network architecture [guest editorial]," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 10–12, October 2015.
- [39] A. S. Thyagaturu, Y. Dashti, and M. Reisslein, "Sdn-based smart gateways (sm-gws) for multi-operator small cell network management," *IEEE Transactions on Network and Service Management*, vol. 13, no. 4, pp. 740–753, Dec 2016.
- [40] M. Artuso, A. Marcano, and H. Christiansen, "Cloudification of mmwave-based and packet-based fronthaul for future heterogeneous mobile networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 76–82, October 2015.
- [41] A. S. Thyagaturu, A. Mercian, M. P. McGarry, M. Reisslein, and W. Kellerer, "Software defined optical networks (sdons): A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 4, pp. 2738–2786, Fourthquarter 2016.
- [42] K. Tan, H. Liu, J. Zhang, Y. Zhang, J. Fang, and G. M. Voelker, "Sora: High-performance software radio using general-purpose multi-core processors," *Commun. ACM*, vol. 54, no. 1, pp. 99–107, Jan. 2011. [Online]. Available: <http://doi.acm.org/10.1145/1866739.1866760>
- [43] N. Kai, S. Jianxing, C. Kulilin, and K. K. Chai, "Td-lte enodeb prototype using general purpose processor," in *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*, Aug 2012, pp. 822–827.
- [44] Z. Zhu, P. Gupta, Q. Wang, S. Kalyanaraman, Y. Lin, H. Franke, and S. Sarangi, "Virtual base station pool: Towards a wireless network cloud for radio access networks," in *Proceedings of the 8th ACM International Conference on Computing Frontiers*, ser. CF '11. New York, NY, USA: ACM, 2011, pp. 34:1–34:10. [Online]. Available: <http://doi.acm.org/10.1145/2016604.2016646>
- [45] C. Y. Yeoh, M. H. Mokhtar, A. A. A. Rahman, and A. K. Samingan, "Performance study of lte experimental testbed using openairinterface," in *2016 18th International Conference on Advanced Communication Technology (ICACT)*, Jan 2016, pp. 617–622.
- [46] J. Wang, J. Xu, Y. Yang, and H. Xu, "Gpp based open cellular network towards 5g," *China Communications*, vol. 14, no. 6, pp. 189–198, 2017.
- [47] I. Gomez-Miguel, A. Garcia-Saavedra, P. D. Sutton, P. Serrano, C. Cano, and D. J. Leith, "srslte: An open-source platform for LTE evolution and experimentation," *CoRR*, vol. abs/1602.04629, 2016. [Online]. Available: <http://arxiv.org/abs/1602.04629>
- [48] W. Rouwet, "Next generation access network development platform and vaccess," March 2015, White Paper. [Online]. Available: https://cache.freescale.com/files/32bit/doc/white/_paper/VBTSWP.pdf
- [49] T. Kazaz, C. Van Praet, M. Kulin, P. Willemen, and I. Moerman, "Hardware accelerated sdr platform for adaptive air interfaces," in *ETSI Workshop on Future Radio Technologies : Air Interfaces*, 2016, pp. 1–10.
- [50] 3GPP, "3gpp tr 38.801 v14.0.0 (2017.03)-study on new radio access technology: Radio access architecture and interfaces, release 14," Tech. Rep., April 2017. [Online]. Available: <http://www.3gpp.org/DynaReport/38801.htm>
- [51] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "Openflow: Enabling innovation in campus networks," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, pp. 69–74, Mar. 2008. [Online]. Available: <http://doi.acm.org/10.1145/1355734.1355746>
- [52] D. Levin, M. Canini, S. Schmid, F. Schaffert, and A. Feldmann, "Panopticon: Reaping the benefits of incremental sdn deployment in enterprise networks," in *2014 USENIX Annual Technical Conference (USENIX ATC 14)*. Philadelphia, PA: USENIX Association, Jun. 2014, pp. 333–345. [Online]. Available: <https://www.usenix.org/conference/atc14/technical-sessions/presentation/levin>
- [53] M. Markovitch and S. Schmid, "Shear: A highly available and flexible network architecture marrying distributed and logically centralized control planes," in *2015 IEEE 23rd International Conference on Network Protocols (ICNP)*, Nov 2015, pp. 78–89.
- [54] C. Jin, C. Lumezanu, Q. Xu, H. Mekky, Z.-L. Zhang, and G. Jiang, "Exerting fine-grained path control over legacy switches in hybrid networks," in *TR 16-035: Technical Report, University of Minnesota*, September 2016, Technical Report. [Online]. Available: http://www.cs.umn.edu/sites/cs.umn.edu/files/tech/_reports/16-035.pdf
- [55] C. Jin, C. Lumezanu, Q. Xu, Z.-L. Zhang, and G. Jiang, "Telekinesis: Controlling legacy switch routing with openflow in hybrid networks," in *Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research*, ser. SOSR '15. New York, NY, USA: ACM, 2015, pp. 20:1–20:7. [Online]. Available: <http://doi.acm.org/10.1145/2774993.2775013>
- [56] K. Pentikousis, Y. Wang, and W. Hu, "Mobileflow: Toward software-defined mobile networks," *IEEE Communications Magazine*, vol. 51, no. 7, pp. 44–53, July 2013.
- [57] J. Kempf, B. Johansson, S. Pettersson, H. Lning, and T. Nilsson, "Moving the mobile evolved packet core to the cloud," in *2012 IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct 2012, pp. 784–791.
- [58] N. Cvijetic, A. Tanaka, K. Kanonakis, and T. Wang, "Sdn-controlled topology-reconfigurable optical mobile fronthaul architecture for bidirectional comp and low latency inter-cell d2d in the 5g mobile era," *Opt. Express*, vol. 22, no. 17, pp. 20 809–20 815, Aug 2014. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-22-17-20809>
- [59] X. Jin, L. E. Li, L. Vanbever, and J. Rexford, "Softcell: Scalable and flexible cellular core network architecture," in *Proceedings of the Ninth ACM Conference on Emerging Networking Experiments and Technologies*, ser. CoNEXT '13. New York, NY, USA: ACM, 2013, pp. 163–174. [Online]. Available: <http://doi.acm.org/10.1145/2535372.2535377>
- [60] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, "5g on the horizon: Key challenges for the radio-access network," *IEEE Vehicular Technology Magazine*, vol. 8, no. 3, pp. 47–53, Sept 2013.
- [61] R. Trivisonno, R. Guerzoni, I. Vaishnavi, and D. Soldani, "Sdn-based 5g mobile networks: Architecture, functions, procedures and backward compatibility," *Trans. Emerg. Telecommun. Technol.*, vol. 26, no. 1, pp. 82–92, Jan. 2015. [Online]. Available: <http://dx.doi.org/10.1002/ett.2915>
- [62] C. J. Bernardos, A. de la Oliva, P. Serrano, A. Banchs, L. M. Contreras, H. Jin, and J. C. Zuniga, "An architecture for software defined wireless networking," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 52–61, June 2014.
- [63] Y. Kyung, T. M. Nguyen, K. Hong, J. Park, and J. Park, "Software defined service migration through legacy service integration into 4g networks and future evolutions," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 108–114, September 2015.
- [64] J. Costa-Requena, J. L. Santos, V. F. Guasch, K. Ahokas, G. Prem-sankar, S. Luukkainen, O. L. Prez, M. U. Itzazelaia, I. Ahmad,

- M. Liyanage, M. Ylianttila, and E. M. de Oca, "Sdn and nfv integration in generalized mobile network architecture," in *2015 European Conference on Networks and Communications (EuCNC)*, June 2015, pp. 154–158.
- [65] V.-G. Nguyen, T.-X. Do, and Y. Kim, "Sdn and virtualization-based lte mobile network architectures: A comprehensive survey," *Wireless Personal Communications*, vol. 86, no. 3, pp. 1401–1438, 2016. [Online]. Available: <http://dx.doi.org/10.1007/s11277-015-2997-7>
- [66] P. Rost, C. J. Bernardos, A. D. Domenico, M. D. Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wbbs, "Cloud technologies for flexible 5g radio access networks," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 68–76, May 2014.
- [67] I. Ahmad, S. Namal, M. Ylianttila, and A. Gurtov, "Security in software defined networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2317–2346, Fourthquarter 2015.
- [68] M. Liyanage, I. Ahmed, M. Ylianttila, J. L. Santos, R. Kantola, O. L. Perez, M. U. Itzazelaia, E. M. d. Oca, A. Valtierra, and C. Jimenez, "Security for future software defined mobile networks," in *2015 9th International Conference on Next Generation Mobile Applications, Services and Technologies*, Sept 2015, pp. 256–264.
- [69] A. Y. Ding, J. Crowcroft, S. Tarkoma, and H. Flinck, "Software defined networking for security enhancement in wireless mobile networks," *Computer Networks*, vol. 66, pp. 94 – 101, 2014, Leonard Kleinrock Tribute Issue: A Collection of Papers by his Students. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1389128614001133>
- [70] D. He, S. Chan, and M. Guizani, "Securing software defined wireless networks," *IEEE Communications Magazine*, vol. 54, no. 1, pp. 20–25, January 2016.
- [71] M. Dabbagh, B. Hamdaoui, M. Guizani, and A. Rayes, "Software-defined networking security: pros and cons," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 73–79, June 2015.
- [72] S. Scott-Hayward, S. Natarajan, and S. Sezer, "A survey of security in software defined networks," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 623–654, Firstquarter 2016.
- [73] C. A. Ezefibe and Y. R. Shayan, "Towards virtualisation and secured software defined networking for wireless and cellular networks," in *2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, May 2016, pp. 1–5.
- [74] M. Liyanage, M. Ylianttila, and A. Gurtov, "Securing the control channel of software-defined mobile networks," in *Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2014*, June 2014, pp. 1–6.
- [75] M. Chen, Y. Qian, S. Mao, W. Tang, and X. Yang, "Software-defined mobile networks security," *Mobile Networks and Applications*, vol. 21, no. 5, pp. 729–743, 2016. [Online]. Available: <http://dx.doi.org/10.1007/s11036-015-0665-5>
- [76] S. Shin, V. Yegneswaran, P. Porras, and G. Gu, "Avant-guard: Scalable and vigilant switch flow management in software-defined networks," in *Proceedings of the 2013 ACM SIGSAC Conference on Computer & #38; Communications Security*, ser. CCS '13. New York, NY, USA: ACM, 2013, pp. 413–424. [Online]. Available: <http://doi.acm.org/10.1145/2508859.2516684>
- [77] D. Kreutz, F. M. V. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, Jan 2015.
- [78] V. Vassilakis, I. Moscholios, B. Alzahrani, and M. Logothetis, "On the security of software-defined next-generation cellular networks," in *IEICE Information and Communication Technology Forum (ICTF)*, July 2016. [Online]. Available: <http://repository.uwl.ac.uk/2827/>
- [79] G. Hampel, M. Steiner, and T. Bu, "Applying software-defined networking to the telecom domain," in *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2013, pp. 133–138.
- [80] S. Namal, I. Ahmad, A. Gurtov, and M. Ylianttila, "Enabling secure mobility with openflow," in *2013 IEEE SDN for Future Networks and Services (SDN4FNS)*, Nov 2013, pp. 1–5.
- [81] R. Guerzoni, R. Trivisonno, and D. Soldani, "Sdn-based architecture and procedures for 5g networks," in *1st International Conference on 5G for Ubiquitous Connectivity*, Nov 2014, pp. 209–214.
- [82] R. Atat, L. Liu, H. Chen, J. Wu, H. Li, and Y. Yi, "Enabling cyber-physical communication in 5g cellular networks: challenges, spatial spectrum sensing, and cyber-security," *IET Cyber-Physical Systems: Theory Applications*, vol. 2, no. 1, pp. 49–54, 2017.
- [83] Z. Yan, P. Zhang, and A. V. Vasilakos, "A security and trust framework for virtualized networks and software-defined networking," *Security and Communication Networks*, vol. 9, no. 16, pp. 3059–3069, 2016, sCN-14-0760.R1. [Online]. Available: <http://dx.doi.org/10.1002/sec.1243>
- [84] X. Duan and X. Wang, "Authentication handover and privacy protection in 5g hetnets using software-defined networking," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 28–35, April 2015.
- [85] Y. Fu, Z. Yan, H. Li, X. L. Xin, and J. Cao, "A secure sdn based multi-rans architecture for future 5g networks," *Computers & Security*, vol. 70, no. Supplement C, pp. 648 – 662, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167404817301785>
- [86] S. Luo, M. Dong, K. Ota, J. Wu, and J. Li, "A security assessment mechanism for software-defined networking-based mobile networks," *Sensors*, vol. 15, no. 12, pp. 31 843–31 858, 2015. [Online]. Available: <http://www.mdpi.com/1424-8220/15/12/29887>
- [87] R. Skowrya, S. Bahargam, and A. Bestavros, "Software-defined ids for securing embedded mobile devices," in *IEEE High Performance Extreme Computing Conference, HPEC*, September 2013.
- [88] R. W. Skowrya, A. Lapets, A. Bestavros, and A. Kfoury, "Verifiably-safe software-defined networks for cps," in *Proceedings of the 2Nd ACM International Conference on High Confidence Networked Systems*, ser. HiCoNS '13. New York, NY, USA: ACM, 2013, pp. 101–110. [Online]. Available: <http://doi.acm.org/10.1145/2461446.2461461>
- [89] R. Jin and B. Wang, "Malware detection for mobile devices using software-defined networking," in *Proceedings of the 2013 Second GENI Research and Educational Experiment Workshop*, ser. GREE '13. Washington, DC, USA: IEEE Computer Society, 2013, pp. 81–88. [Online]. Available: <http://dx.doi.org/10.1109/GREE.2013.24>
- [90] G. Hurel, R. Badonnel, A. Lahmadi, and O. Festor, *Outsourcing Mobile Security in the Cloud*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 69–73. [Online]. Available: http://dx.doi.org/10.1007/978-3-662-43862-6_9
- [91] K. Grenier, "Business case for a common nfv platform," ACG Research, Tech. Rep., August 2015. [Online]. Available: <http://acgcc.com/business-case-for-a-common-nfv-platform/>
- [92] C. Bouras, P. Ntaranos, and A. Papazois, "Cost modeling for sdn/nfv based mobile 5g networks," in *2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Oct 2016, pp. 56–61.
- [93] R. Haim, "Total cost of ownership study, virtualizing the mobile core," ACG Research, Tech. Rep., July 2015. [Online]. Available: http://www.affirmednetworks.com/wp-content/uploads/2015/07/TCO-Report_7.13.15_ACG-Template.pdf
- [94] B. Naudts, M. Kind, F. J. Westphal, S. Verbrugge, D. Colle, and M. Pickavet, "Techno-economic analysis of software defined networking as architecture for the virtualization of a mobile network," in *2012 European Workshop on Software Defined Networking*, Oct 2012, pp. 67–72.
- [95] B. Naudts, "Techno-economic analysis of software defined networking as architecture for the virtualization of a mobile network," in *IEEE Software Defined Networks for Future Networks and Services (SDN4FNS)*, Nov 2013, Panel Presentation. [Online]. Available: http://sites.ieee.org/sdn4fns/files/2013/11/SDN4FNS_panel_presentation_Bram_Naudts.pdf
- [96] A. Checko, H. L. Christiansen, and M. S. Berger, "Evaluation of energy and cost savings in mobile cloud ran," in *Proceedings of OPNETWORK 2013*, ser. OPNETWORK '13. OPNET, 2013. [Online]. Available: http://orbit.dtu.dk/fedora/objects/orbit:124961/datastreams/file_f0352259-e81d-440c-ad42-b417a02e5999/content
- [97] W. Wu, L. E. Li, A. Panda, and S. Shenker, "Pran: Programmable radio access networks," in *Proceedings of the 13th ACM Workshop on Hot Topics in Networks*, ser. HotNets-XIII. New York,

- NY, USA: ACM, 2014, pp. 6:1–6:7. [Online]. Available: <http://doi.acm.org/10.1145/2670518.2673865>
- [98] V. Suryaprakash, P. Rost, and G. Fettweis, “Are heterogeneous cloud-based radio access networks cost effective?” *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2239–2251, Oct 2015.
- [99] N. Zhang and H. Hmminen, “Cost efficiency of sdn in lte-based mobile networks: Case finland,” in *2015 International Conference and Workshops on Networked Systems (NetSys)*, March 2015, pp. 1–5.
- [100] A. Asensio, P. Saengudomlert, M. Ruiz, and L. Velasco, “Study of the centralization level of optical network-supported cloud ran,” in *2016 International Conference on Optical Network Design and Modeling (ONDM)*, May 2016, pp. 1–6.
- [101] H. Yeganeh and E. Vaezpour, “Fronthaul network design for radio access network virtualization from a capex/opex perspective,” *Annals of Telecommunications*, vol. 71, no. 11, pp. 665–676, 2016. [Online]. Available: <http://dx.doi.org/10.1007/s12243-016-0538-3>
- [102] M. Karimzadeh, L. Valtulina, and G. Karagiannis, “Applying sdn/openflow in virtualized lte to support distributed mobility management (dmm),” in *Proceedings of the 4th International Conference on Cloud Computing and Services Science*, ser. CLOSER 2014. Portugal: SCITEPRESS - Science and Technology Publications, Lda, 2014, pp. 639–644. [Online]. Available: <http://dx.doi.org/10.5220/0004946106390644>
- [103] S. Kang and W. Yoon, “Sdn-based resource allocation for heterogeneous lte and wlan multi-radio networks,” *J. Supercomput.*, vol. 72, no. 4, pp. 1342–1362, Apr. 2016. [Online]. Available: <http://dx.doi.org/10.1007/s11227-016-1662-6>
- [104] S. Costanzo, D. Xenakis, N. Passas, and L. Merakos, “Openb: A framework for virtualizing base stations in lte networks,” in *2014 IEEE International Conference on Communications (ICC)*, June 2014, pp. 3148–3153.
- [105] P. Spapis, K. Chatzikokolakis, N. Alonistioti, and A. Kaloxylas, “Using sdn as a key enabler for co-primary spectrum sharing,” in *IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications*, July 2014, pp. 366–371.
- [106] D. Boviz, N. Abbas, G. Aravinthan, C. S. Chen, and M. A. Dridi, “Multi-cell coordination in cloud ran: Architecture and optimization,” in *2016 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Oct 2016, pp. 271–277.
- [107] Y. Zhang, N. Beheshti, L. Beliveau, G. Lefebvre, R. Manghirmalani, R. Mishra, R. Patney, M. Shirazipour, R. Subrahmaniam, C. Truchan, and M. Tatipamula, “Steering: A software-defined networking for inline service chaining,” in *2013 21st IEEE International Conference on Network Protocols (ICNP)*, Oct 2013, pp. 1–10.
- [108] K. Katsalis, N. Nikaein, E. Schiller, R. Favraud, and T. I. Braun, “5g architectural design patterns,” in *2016 IEEE International Conference on Communications Workshops (ICC)*, May 2016, pp. 32–37.
- [109] M. R. Sama, S. B. H. Said, K. Guillooard, and L. Suci, “Enabling network programmability in lte/epc architecture using openflow,” in *2014 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2014, pp. 389–396.
- [110] V. Nguyen and Y. Kim, “Proposal and evaluation of sdn-based mobile packet core networks,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, p. 172, 2015. [Online]. Available: <http://dx.doi.org/10.1186/s13638-015-0395-1>
- [111] A. Ksentini, M. Baggaa, and T. Taleb, “On using sdn in 5g: The controller placement problem,” in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.
- [112] J. Costa-Requena, R. Kantola, J. Llorente, V. Ferrer, J. Manner, A. Y. Ding, Y. Liu, and S. Tarkoma, “Software defined 5g mobile backhaul,” in *1st International Conference on 5G for Ubiquitous Connectivity*, Nov 2014, pp. 258–263.
- [113] M. H. Kabir, “A novel architecture for sdn-based cellular network,” *International Journal of Wireless & Mobile Networks (IJWMN)*, vol. 6, no. 6, pp. 71–85, December 2014.
- [114] X. An, W. Kiess, J. Varga, J. Prade, H. J. Morper, and K. Hoffmann, “Sdn-based vs. software-only epc gateways: A cost analysis,” in *2016 IEEE NetSoft Conference and Workshops (NetSoft)*, June 2016, pp. 146–150.
- [115] A. Bradai, K. Singh, T. Ahmed, and T. Rasheed, “Cellular software defined networking: a framework,” *IEEE Communications Magazine*, vol. 53, no. 6, pp. 36–43, June 2015.
- [116] A. Basta, W. Kellerer, M. Hoffmann, K. Hoffmann, and E. D. Schmidt, “A virtual sdn-enabled lte epc architecture: A case study for s-/p-gateways functions,” in *2013 IEEE SDN for Future Networks and Services (SDN4FNS)*, Nov 2013, pp. 1–7.
- [117] I. Elgendi, K. S. Munasinghe, and A. Jamalipour, “A three-tier sdn architecture for densenets,” in *2015 9th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Dec 2015, pp. 1–7.
- [118] E. B. Hamza and S. Kimura, “A scalable sdn-epc architecture based on openflow-enabled switches to support inter-domain handover,” in *2016 10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, July 2016, pp. 272–277.
- [119] M. Yang, Y. Li, B. Li, D. Jin, and S. Chen, “Service-oriented 5g network architecture: an end-to-end software defining approach,” *International Journal of Communication Systems*, vol. 29, no. 10, pp. 1645–1657, 2016. [Online]. Available: <http://www.ingentaconnect.com/content/bpl/dac/2016/00000029/00000010/art00004>
- [120] S. Aurox, M. Drxler, A. Morelli, and V. Mancuso, “Dynamic network reconfiguration in wireless densenets with the crowd sdn architecture,” in *2015 European Conference on Networks and Communications (EuCNC)*, June 2015, pp. 144–148.
- [121] H. Ali-Ahmad, C. Cicconetti, A. de la Oliva, V. Mancuso, M. R. Sama, P. Seite, and S. Shanmugalingam, “An sdn-based network architecture for extremely dense wireless networks,” in *2013 IEEE SDN for Future Networks and Services (SDN4FNS)*, Nov 2013, pp. 1–7.
- [122] V. Yazici, U. C. Kozat, and M. O. Sunay, “A new control plane for 5g network architecture with a case study on unified handoff, mobility, and routing management,” *IEEE Communications Magazine*, vol. 52, no. 11, pp. 76–85, Nov 2014.
- [123] R. Trivisonno, R. Guerzoni, I. Vaishnavi, and D. Soldani, “Towards zero latency software defined 5g networks,” in *2015 IEEE International Conference on Communication Workshop (ICCW)*, June 2015, pp. 2566–2571.
- [124] V. G. Vassilakis, I. D. Moscholios, B. A. Alzaharani, and M. D. Logothetis, “A software-defined architecture for next-generation cellular networks,” in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.
- [125] X. Mi, Z. Tian, X. Xu, M. Zhao, and J. Wang, “No stack: A sdn-based framework for future cellular networks,” in *2014 International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Sept 2014, pp. 497–502.
- [126] Y. Lu, X. Su, J. Zeng, L. Rong, C. Yang, and X. Xu, “Software defined radio access network in 5g mobile network,” in *2015 10th International Conference on Communications and Networking in China (ChinaCom)*, Aug 2015, pp. 132–136.
- [127] K. Chen, R. Izard, H. Hu, K. C. Wang, J. Martin, and J. Deng, “Hetsdn: Exploiting sdn for intelligent network usage in heterogeneous wireless networks,” in *2016 IEEE/ACM 24th International Symposium on Quality of Service (IWQoS)*, June 2016, pp. 1–6.
- [128] R. Riggio, K. Gomez, L. Goratti, R. Fedrizzi, and T. Rasheed, “V-cell: Going beyond the cell abstraction in 5g mobile networks,” in *2014 IEEE Network Operations and Management Symposium (NOMS)*, May 2014, pp. 1–5.
- [129] S. Sun, M. Kadoch, L. Gong, and B. Rong, “Integrating network function virtualization with sdr and sdn for 4g/5g networks,” *IEEE Network*, vol. 29, no. 3, pp. 54–59, May 2015.
- [130] M. Amani, T. Mahmoodi, M. Tatipamula, and H. Aghvami, “Programmable policies for data offloading in lte network,” in *2014 IEEE International Conference on Communications (ICC)*, June 2014, pp. 3154–3159.
- [131] Huawei, “Huawei softcom, reshaping the future of network architecture,” in *Report, Huawei*, 2013, Report. [Online]. Available: www.huawei.com/ilink/en/download/HW_204467
- [132] D. Sabella, P. Rost, Y. Sheng, E. Pateromichelakis, U. Salim, P. Guitton-Ouhamous, M. D. Girolamo, and G. Giuliani, “Ran as a service: Challenges of designing a flexible ran architecture in a cloud-based

- heterogeneous mobile network,” in *2013 Future Network Mobile Summit*, July 2013, pp. 1–8.
- [133] V. Jungnickel, K. Habel, M. Parker, S. Walker, C. Bock, J. F. Riera, V. Marques, and D. Levi, “Software-defined open architecture for front- and backhaul in 5g mobile networks,” in *2014 16th International Conference on Transparent Optical Networks (ICTON)*, July 2014, pp. 1–4.
- [134] P. Arnold, N. Bayer, J. Belschner, and G. Zimmermann, “5g radio access network architecture based on flexible functional control / user plane splits,” in *2017 European Conference on Networks and Communications (EuCNC)*, June 2017, pp. 1–5.
- [135] C. Yang, Z. Chen, B. Xia, and J. Wang, “When icn meets c-ran for hetnets: an sdn approach,” *IEEE Communications Magazine*, vol. 53, no. 11, pp. 118–125, November 2015.
- [136] F. Han, S. Zhao, L. Zhang, and J. Wu, “Survey of strategies for switching off base stations in heterogeneous networks for greener 5g systems,” *IEEE Access*, vol. 4, pp. 4959–4973, 2016.
- [137] X. Costa-Perez, A. Garcia-Saavedra, X. Li, T. Deiss, A. de la Oliva, A. di Giglio, P. Iovanna, and A. Moore, “5g-crosshaul: An sdn/nfv integrated fronthaul/backhaul transport network architecture,” *IEEE Wireless Communications*, vol. 24, no. 1, pp. 38–45, February 2017.
- [138] H. H. Cho, C. F. Lai, T. K. Shih, and H. C. Chao, “Integration of sdr and sdn for 5g,” *IEEE Access*, vol. 2, pp. 1196–1204, 2014.
- [139] E. J. Kitindi, S. Fu, Y. Jia, A. Kabir, and Y. Wang, “Wireless network virtualization with sdn and c-ran for 5g networks: Requirements, opportunities, and challenges,” *IEEE Access*, vol. 5, pp. 19 099–19 115, 2017.
- [140] J. Mueller, Y. Chen, B. Reichel, V. Vlad, and T. Magedanz, “Design and implementation of a carrier grade software defined telecommunication switch and controller,” in *2014 IEEE Network Operations and Management Symposium (NOMS)*, May 2014, pp. 1–7.
- [141] J. Heinonen, T. Partti, M. Kallio, K. Lappalainen, H. Flinck, and J. Hillo, “Dynamic tunnel switching for sdn-based cellular core networks,” in *Proceedings of the 4th Workshop on All Things Cellular: Operations, Applications, & Challenges*, ser. AllThingsCellular ’14. New York, NY, USA: ACM, 2014, pp. 27–32. [Online]. Available: <http://doi.acm.org/10.1145/2627585.2627587>
- [142] J. Page and J. M. Dricot, “Software-defined networking for low-latency 5g core network,” in *2016 International Conference on Military Communications and Information Systems (ICMCIS)*, May 2016, pp. 1–7.
- [143] S. Chourasia and K. M. Sivalingam, “Sdn based evolved packet core architecture for efficient user mobility support,” in *Proceedings of the 2015 1st IEEE Conference on Network Softwarization (NetSoft)*, April 2015, pp. 1–5.
- [144] ONF, “Relationship of sdn and nfv, issue 1, onf tr-518,” ONF, Tech. Rep., October 2015. [Online]. Available: https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/onf2015.310_Architectural_comparison.08-2.pdf
- [145] E. N. ISG, “Gs nfv-eve 005 v1.1.1 network function virtualisation (nfv); ecosystem; report on sdn usage in nfv architectural framework,” ETSI, Tech. Rep., December 2015. [Online]. Available: http://www.etsi.org/deliver/etsi_gs/NFV-EVE/001_099/005/01.01.01_60/gs_NFV-EVE005v010101p.pdf
- [146] F. Z. Yousaf, M. Bredel, S. Schaller, and F. Schneider, “Nfv and sdn - key technology enablers for 5g networks,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 11, pp. 2468–2478, Nov 2017.
- [147] G. NORMA, “Deliverable d3.2, 5g norma network architecture intermediate report,” ERC, Tech. Rep., January 2017. [Online]. Available: https://5gnorma.5g-ppp.eu/wp-content/uploads/2017/03/5g_norma_d3-2.pdf
- [148] M. R. Sama, L. M. Contreras, J. Kaippallimalil, I. Akiyoshi, H. Qian, and H. Ni, “Software-defined control of the virtualized mobile packet core,” *IEEE Communications Magazine*, vol. 53, no. 2, pp. 107–115, Feb 2015.
- [149] F. Z. Yousaf, J. Lessmann, P. Loureiro, and S. Schmid, “Softpec; dynamic instantiation of mobile core network entities for efficient resource utilization,” in *2013 IEEE International Conference on Communications (ICC)*, June 2013, pp. 3602–3606.
- [150] E. Haleplidis, D. Joachimpillai, J. H. Salim, K. Pentikousis, S. Denazis, and O. Koufopavlou, “Building softwarized mobile infrastructures with forces,” in *2016 23rd International Conference on Telecommunications (ICT)*, May 2016, pp. 1–5.
- [151] V. Ziegler, T. Theimer, C. Sartori, J. Prade, N. Sprecher, K. Albal, and A. Bedekar, “Architecture vision for the 5g era,” in *2016 IEEE International Conference on Communications Workshops (ICC)*, May 2016, pp. 51–56.
- [152] L. Valtulina, M. Karimzadeh, G. Karagiannis, G. Heijenk, and A. Pras, “Performance evaluation of a sdn/openflow-based distributed mobility management (dmm) approach in virtualized lte systems,” in *2014 IEEE Globecom Workshops (GC Wkshps)*, Dec 2014, pp. 18–23.
- [153] S. Kuklinski, Y. Li, and K. T. Dinh, “Handover management in sdn-based mobile networks,” in *2014 IEEE Globecom Workshops (GC Wkshps)*, Dec 2014, pp. 194–200.
- [154] Z. Cao, S. S. Panwar, M. Kodialam, and T. V. Lakshman, “Enhancing mobile networks with software defined networking and cloud computing,” *IEEE/ACM Transactions on Networking*, vol. PP, no. 99, pp. 1–14, 2017.
- [155] A. Hurtado-Borras, J. Pal-Sol, D. Camps-Mur, and S. Sallent-Ribes, “Sdn wireless backhauling for small cells,” in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 3897–3902.
- [156] R. Santos and A. Kessler, “A sdn controller architecture for small cell wireless backhaul using a lte control channel,” in *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, June 2016, pp. 1–3.
- [157] S. Sun, L. Gong, B. Rong, and K. Lu, “An intelligent sdn framework for 5g heterogeneous networks,” *IEEE Communications Magazine*, vol. 53, no. 11, pp. 142–147, November 2015.
- [158] V. Sagar, R. Chandramouli, and K. P. Subbalakshmi, “Software defined access for hetnets,” *IEEE Communications Magazine*, vol. 54, no. 1, pp. 84–89, January 2016.
- [159] K.-K. Yap, R. Sherwood, M. Kobayashi, T.-Y. Huang, M. Chan, N. Handigol, N. McKeown, and G. Parulkar, “Blueprint for introducing innovation into wireless mobile networks,” in *Proceedings of the Second ACM SIGCOMM Workshop on Virtualized Infrastructure Systems and Architectures*, ser. VISA ’10. New York, NY, USA: ACM, 2010, pp. 25–32. [Online]. Available: <http://doi.acm.org/10.1145/1851399.1851404>
- [160] M. Yang, Y. Li, L. Hu, B. Li, D. Jin, S. Chen, and Z. Yan, “Cross-layer software-defined 5g network,” *Mob. Netw. Appl.*, vol. 20, no. 3, pp. 400–409, Jun. 2015. [Online]. Available: <http://dx.doi.org/10.1007/s11036-014-0554-3>
- [161] K. Gomez, T. Rasheed, L. Reynaud, and L. Goratti, “Fme: A flexible management entity for virtualizing lte evolved packet core,” in *2014 IEEE Network Operations and Management Symposium (NOMS)*, May 2014, pp. 1–4.
- [162] M. Yang, Y. Li, D. Jin, L. Zeng, X. Wu, and A. V. Vasilakos, “Software-defined and virtualized future mobile and wireless networks: A survey,” *Mob. Netw. Appl.*, vol. 20, no. 1, pp. 4–18, Feb. 2015. [Online]. Available: <http://dx.doi.org/10.1007/s11036-014-0533-8>
- [163] N. A. Jagadeesan and B. Krishnamachari, “Software-defined networking paradigms in wireless networks: A survey,” *ACM Comput. Surv.*, vol. 47, no. 2, pp. 27:1–27:11, Nov. 2014. [Online]. Available: <http://doi.acm.org/10.1145/2655690>
- [164] S. Tomovic, M. Pejanovic-Djurisic, and I. Radusinovic, “Sdn based mobile networks: Concepts and benefits,” *Wireless Personal Communications*, vol. 78, no. 3, pp. 1629–1644, 2014. [Online]. Available: <http://dx.doi.org/10.1007/s11277-014-1909-6>
- [165] J. Li, P. Liu, and H. Li, “Software-defined cellular mobile network solutions,” *ZTE Communications*, vol. 12, no. 2, pp. 28–33, June 2014.
- [166] T. Chen, M. Matinmikko, X. Chen, X. Zhou, and P. Ahokangas, “Software defined mobile networks: concept, survey, and research directions,” *IEEE Communications Magazine*, vol. 53, no. 11, pp. 126–133, November 2015.
- [167] I. T. Haque and N. Abu-Ghazaleh, “Wireless software defined networking: A survey and taxonomy,” *IEEE Communications Surveys Tutorials*, vol. 18, no. 4, pp. 2713–2737, Fourthquarter 2016.

- [168] V. G. Nguyen, A. Brunstrom, K. J. Grinnemo, and J. Taheri, "Sdn/nfv-based mobile packet core network architectures: A survey," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–1, 2017.
- [169] M. A. Al-Namari, A. M. Mansoor, and M. Y. I. Idris, "A brief survey on 5g wireless mobile network," *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 8, no. 11, 2017. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2017.081107>
- [170] S. K. Tayyaba and M. A. Shah, "5g cellular network integration with sdn: Challenges, issues and beyond," in *2017 International Conference on Communication, Computing and Digital Systems (C-CODE)*, March 2017, pp. 48–53.
- [171] D. A. Temesgene, J. Nez-Martnez, and P. Dini, "Softwarization and optimization for sustainable future mobile networks: A survey," *IEEE Access*, vol. 5, pp. 25 421–25 436, 2017.