New Concepts for Traffic, Resource and Mobility Management in Software-Defined Mobile Networks

Ijaz Ahmad*, Madhusanka Liyanage, Suneth Namal
Mika Ylianttila, Andrei Gurtov
University of Oulu, Finland

Marcus Eckert, Thomas Bauschert
Chair for Communication Networks
Chemnitz University of Technology, Germany

Zoltán Faigl, László Bokor
BME Mobile Innovation Centre
Budapest, Hungary

Ece Saygun, Hasan Anıl Akyıldız
Ericsson R&D Center
İstanbul, Turkey

Oscar Lopez Perez, Mikel Uriarte Itzazelaia
Nextel S.A.
Zamudio, Spain

Berna Özbek, Aydın Ulaş
Izmir Institute of Technology, ARGELA A.Ş.
İzmir/Turkey, İstanbul/Turkey

***Abstract*—The evolution of mobile telecommunication networks is accompanied by new demands for the performance, portability, elasticity, and energy efficiency of network functions. Network Function Virtualization (NFV), Software Defined Networking (SDN), and cloud service technologies are claimed to be able to provide most of the capabilities. However, great leap forward will only be achieved if resource, traffic, and mobility management methods of mobile network services can efficiently utilize these technologies. This paper conceptualizes the future requirements of mobile networks and proposes new concepts and solutions in the form of Software-Defined Mobile Networks (SDMN) leveraging SDN, NFV and cloud technologies. We evaluate the proposed solutions through testbed implementations and simulations. The results reveal that our proposed SDMN enhancements supports heterogeneity in wireless networks with performance improvements through programmable interfaces and centralized control.**

***Index Terms*—Traffic Management; Resource Management; Mobility Management; Software-Defined Networking; Network Function Virtualization; 5G**

I. INTRODUCTION

Software-Defined Mobile Networks (SDMN) extend the idea of Software-Defined Networking (SDN) and Network Function Virtualization (NFV) into mobile networks. The aim is improving the scalability and adaptability of the mobile network architectures to varying and diverse traffic demands by leveraging existing host and network virtualization technologies besides SDN. NFV decouples software implementations of network functions from the hardware resources. SDN separates the network control functions from the data forwarding elements and logically centralizes the network control functions to enable programmability of the networking equipment through programmable interfaces [1].

Current mobile networks face a number of challenges such as the unprecedented growth in user traffic, high Operational Expenses (OpEx) and Capital Expenses (CapEx), and the complexity in network management. To meet these challenges, mobile service providers and operators seek new solutions such as opportunities in cloud networking technologies, SDN and NFV. SDN is a natural choice for network architectures since it simplifies the network control and enables network management from a central vantage point, e.g. physical or virtual servers running in a data center. NFV technologies not only minimize OpEx and CapEx, but also enhance the scalability and availability of the network resources.

The combined benefits of NFV and SDN are the main motives behind the integrated architecture presented in this paper. We present new concepts for the SDMN architecture leveraging NFV and SDN to solve the challenges of resource, traffic and mobility management faced by the legacy mobile networks. Furthermore, new opportunities are sought out and analyzed theoretically and experimentally to meet the user requirements and address the limitations of SDN and NFV technologies for future mobile networks.

This paper is organized as follows: In Section II, we provide an overview of the existing work and present a high-level SDMN architecture. We also provide a brief overview of the European Telecommunications Standards Institute (ETSI) NFV reference architecture. Section III gives the details of our proposed concepts with their implementation frameworks and the expected benefits. Section IV presents the evaluation of the proposed concepts through testbed implementations or simulations. Section V concludes the paper and presents our future plans.

II. BACKGROUND

A. Related Work

There exist proposals that try to solve the challenges faced by current mobile network architectures with the help of cloudification, NFV, and SDN. Software-defined Radio Access Network (SoftRAN) [2] proposes a software-defined centralized control plane for radio access network. Conventionally,

ISBN 978-3-901882-80-7

radio access networks have been considered as a collection of base stations making independent control decisions by having loose distributed coordination. SoftRAN introduces a centralized software-defined control plane for radio access networks that abstracts all the base stations in a defined geographical area as a single, virtual, and big base station. This framework can effectively perform load balancing, interference management, throughput maximization, and global network utility enhancement.

SoftCell [3] is a scalable architecture proposed to support fine-grained policy implementation for mobile devices in cellular networks. SoftCell enables operators to realize high-level service policies to efficiently utilize the resources based on the subscriber attributes and applications. SoftCell leverages from aggregating traffic along the dimensions of policy, base station, and User Equipment (UE) by using an identifier associated to them. Implemented on top of the Floodlight SDN controller [4], the SoftCell platform is evaluated using real traces from a large LTE (Long Term Evolution) network.

CellSDN [5] is another solution which attempts to simplify the design and management of cellular data networks, while enabling new services and prototyping using open source LTE implementation. Proposal for SDN and NFV integration in mobile network architectures is presented in [6]. In [6] the authors listed the mandatory requirements for the adoption of new technologies in mobile networks and demonstrated the feasibility of SDN and NFV for 5G mobile networks.

B. Architecture

LTE uses all-IP based core network architecture called the Evolved Packet Core (EPC). EPC contains network control elements such as MME (Mobility Management Entity), HSS (Home Subscriber Server), PCRF (Policy and Charging Rule Function) and the control planes of S/P-GW(Serving/Packet Gateway). We propose the integration of these elements and other network services including network monitoring, load balancing, resource optimization, security and charging policy services into the cloud to benefit from NFV and SDN.

SDMN uses the concepts of SDN and NFV to solve the challenges faced by current mobile network architectures in integrating diverse services and network functions. Leveraging SDN, SDMN concentrates network control functions into a logically centralized control plane. The network control system using a Network Operating System (NOS) provides an abstract view of the underlying network infrastructure to network services and functions implemented on top of the control plane [1], [6], [7]. SDMN deploys these network control functions and other network services on top of the control plane in the core network cloud as shown in the high-level architecture presented in Fig. 1.

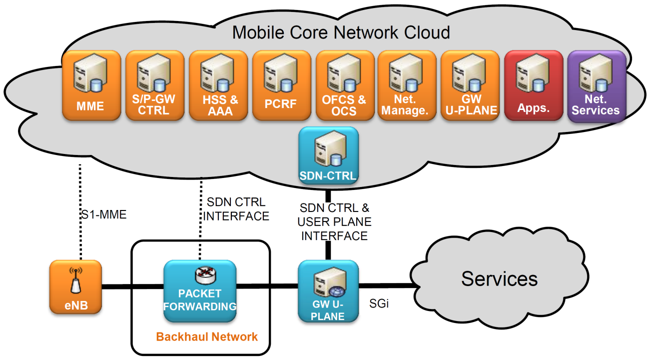


Fig. 1. High-Level architecture of SDMN.

The ETSI NFV Industry Specification Group is on track to define an NFV ecosystem using the works of other standards organizations and identifying new gaps for them. The ETSI NFV reference architecture [8] provides a good reference for the SDMN architecture. NFV architecture has three main parts. (1) Network Function Virtualization Infrastructure (NFVI) to realize host and network virtualization above physical resources; (2) Virtual network Functions (VNF) which are attached to the container interfaces of virtual machines and are connected by virtual networks; and (3) Management and Orchestration of NFVI and VNF. The orchestrator is the high-level hook for management and is aware of NFVI resources and VNF instances.

In addition to three main functional blocks, the NFV architecture must be compatible with service providers' existing management systems, Operational and Business Support Systems (OSS/BSS) and their legacy network equipment in order to support management functions such as network inventory management, service provisioning, network configuration and fault management. Further details of the NFV reference architecture can be found in ETSI NFV Open Area [8]. Fig. 2 is a simplified version of the ETSI NFV architecture that presents our architectural components of SDMN which are further described in detail in the following Section III.

III. PROPOSED ARCHITECTURE COMPONENTS

In this section the proposed components for the SDMN architecture are described. Fig. 2 maps our architectural components of SDMN to the SDN-based virtual and cloud network network architecture. It is mainly a functional diagram, but also illustrates that different software-based functions of SDMN can run as VNFs. Resource, traffic and mobility management require changes in and between the network control functions, VNF manager and Orchestrator. SDN controllers provide abstractions and northbound Application Programming Interfaces (APIs) for the NFVI network control of the functions required for management and control of network resources.

As shown in Fig. 2, the components constitute the orchestrator and can interact with the core network cloud. Virtualized instances of the components connect to the virtual infrastructure domain with the help of the VNF manager using the NVFI API. Besides the virtual data forwarding plane, e.g. OpenFlow switches, the virtual infrastructure comprises virtual instances of some of the proposed components. The functionality of each of the component is described below in detail.

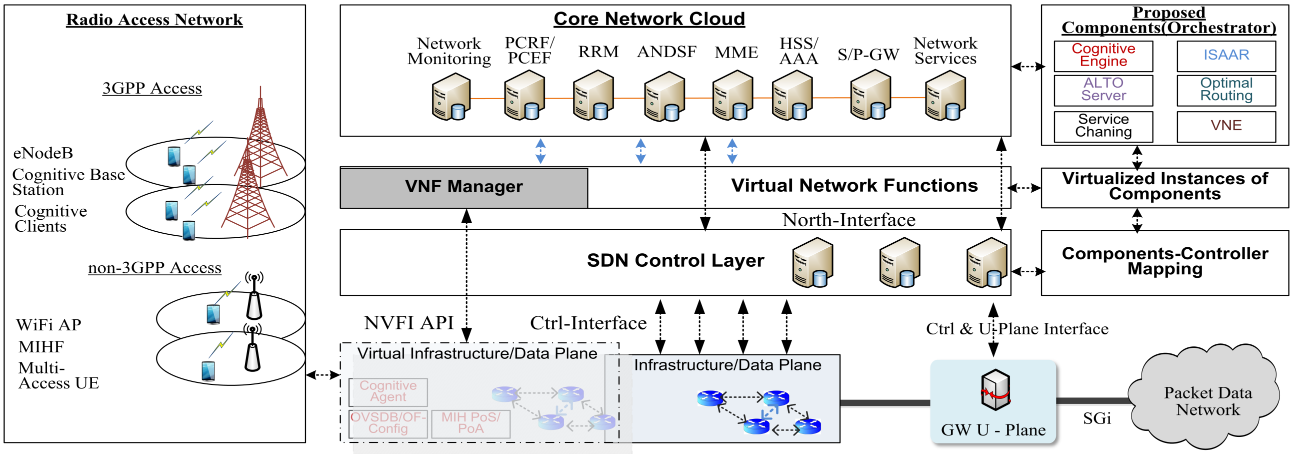


Fig. 2. Proposed architecture and its components.

A. Optimized Resource and Traffic management

One goal of traffic management is to decide how to route traffic in a network in order to balance several objectives such as maximizing throughput, balancing the link utilization across the network, minimizing the power consumption, controlling the bandwidth allocated to competing flows, minimizing latency and ensuring reliable operation when traffic patterns change or parts of the network fail [9], [10], [11]. We examine the joint routing and traffic management optimization problem to minimize the cost based on the number of active switches while satisfying rate requirements of the flows under the assumption that the traffic load in the network is low (night time traffic case). We implement a heuristic framework with the proposed routing algorithm to obtain the solution that satisfies the constraints for a given bandwidth and network topology. In order to verify the algorithm, the rate constraints depending on the applications are assigned and the cost based on power consumption on switches is calculated in the given network topology.

The optimal routing and resource management module (referred to as Optimal Routing in the architecture) which is connected to a northbound agent implementing the API, receives the initial topology and the capacity of the links from the SDN controller. When a request for a path between two nodes is requested, the controller relays this request to the application and the heuristic algorithm finds the optimum path while reducing the power consumption and returns back to the controller to set up the path via southbound interfaces. The optimization application also communicates with the switch management module (referred to as OF-Config/OVSDB in Fig. 2, representing the most commonly used protocols for switch management) to shut down and turn on switches to reduce power consumption.

B. Service Chaining

An important part of the calculation of end-to-end route for a given flow in todays network control is including an array of inline services, such as DPI, firewalls, NAT, content caching, video transcoding, etc. Inline services can be hosted on physical hardware, on virtual machines or in the cloud. In fact, almost all traffic in mobile networks visit a pre-defined sequence of inline services en-route to its destination today. Such a sequence is commonly referred to as a service chain. Even though traffic flows are needed to en-route only a subset of the services, all traffic in the network traverses each of the services that cause massively over-provisioning of services. There is a need to orchestrate a service chain for each flow in order to serve applications in an efficient manner.

With recent advances in SDN, operators and service providers are highly interested in deploying SDN enabled service chaining solution to provide optimally utilizing compute and network resources, deploy personalized end-user services, and reduce OpEx and CapEx. Many vendors have already been working on developing SDN enabled service chaining solutions [12] [13]. Performing flexible and dynamic traffic steering for service chaining, optimizing the cloud realization of functions and services, and optimizing a service chaining route for a given network flow jointly and dynamically are the some of the main research areas in SDN enabled service chaining. Following issues will be considered jointly for efficient service chaining: (1) Potential route congestion towards servers which host the in-line services, (2) Load of the servers which host these services and creation of new hosts per service if needed, (3) Service chaining and route optimization for the case of some services in the chain being hosted over the cloud, for example within Network as a Service (NaaS) model.

Next, we summarize the state-of-the-art in relation to these three issues. An experimental SDN-based test-bed is introduced in [14], where three data centers form a 125 km ring of single-mode optical fiber. The test-bed shows an automatic triggering of live VM migration when the server utilization exceeds 75%. Martini et al. [15] propose an SDN-based architecture for the orchestration of dynamic resource chaining in cloud data centers.

Efficient handling of dynamic updates to virtualized applications, orchestration of chains for large services, and integrating

SDN with multiple cloud controllers and data centers can be listed as some of the most important challenges. Dynamicity is also necessary for virtual server migration of an existing service in the chain during the course of a flow. This migration might be due to load balancing or network congestion purposes. Creation of high performance service-chaining applications is essential to meet traffic demand and higher quality expectations from customers while reducing capital and operational expenses associated with their networks.

Our approach will be to develop a controller module with a mechanism taking jointly into account server loads for inline services and routes towards these servers, and cloud interaction for services provided over the cloud. Getting periodic updates and integrating with OSS/BSS in order to provide input for the mechanism is part of the research. Another important part will be demonstrating results by chaining selected services.

C. Combined Virtual Mobile Core Network Function Placement and Topology Optimization

The problem of finding an optimum mapping of virtual nodes and links onto a given physical substrate network is considered as the virtual network embedding (VNE) problem. For VNE, some solutions have been already proposed in the scientific literature. These approaches consider the optimum embedding of a virtual network with predefined topology onto a given physical substrate network. Thus, the virtual network topology (graph) itself is fixed and not subject to optimization.

We focus on the virtual mobile core network embedding problem and develop a novel approach for combining the optimization of the virtual network topology with the VNE optimization [16]. It relies on the joint embedding of individual core network service chains where a core network service chain denotes the sequence of mobile core VNFs a user or control plane traffic flow has to traverse. For the proposed model, the mobile core VNFs are assumed to comprise the same functionality and interfaces than the core network elements of the 3GPP LTE Evolved Packet Core architecture. A core network service chain is further decomposed into a user plane sub-chain (RAN Traffic Aggregation Point - SGW - PGW - Internet Exchange Point) and several control plane sub-chains (RAN Traffic Aggregation Point - SGW), (MME - HSS), (MME - SGW).

The traffic demand is specified as follows: For each core network service chain the (user and control plane related) bandwidth requirements on the virtual links between the VNFs as well as the processing, storage and switching (throughput) resources the VNFs pose on physical nodes are given.

The optimization target is to find a feasible embedding of the core network service chains according to given physical network link/node capacities and capabilities, so as to minimize the total cost for the mobile core network embedding while accommodating the traffic demand. We formulate the objective function as a linear combination (with weight factors) of three cost terms: basic cost that occurs if any VNF is placed on a physical substrate node, cost per occupied storage, processing and switching unit on a physical substrate node and cost per occupied capacity unit on a physical link.

The new VNE optimization model was evaluated by means of two different physical network topology examples taken from SNDlib [17] with 50 and 12 nodes, respectively. It was assumed that all nodes have server capability, i.e. are able of hosting any mobile core VNF whereas only a predefined subset of them can provide connectivity to the Internet. The evaluation showed that the new approach is capable of finding good solutions for realistic problem instances within few minutes of computation time.

As a next step we intend to include the option of a decomposed SGW/PGW user/control plane in the optimization model and also to consider resilience constraints.

D. Application-Layer Traffic Optimization in SDMNs

Endpoint selection function is an integral part of distributed services, should it be content delivery using peer-to-peer or content distribution networks or a value-added service of an internet service provider. Its significance will increase as the elasticity and portability of network functions come true in the telecommunication sector and new scenarios, e.g., in-network caching or service chaining emerge. Existing application-level or cross-layer techniques aiding the discovery and ranking of endpoints often have limitations regarding accuracy, convergence and other properties [18]. It is worthwhile to establish a network information service, which provides explicit, fresh and aggregated view of all actors (network provider, content distributor etc.) on network costs and is able to provide well-informed guidance for clients.

Application-Layer Traffic Optimization (ALTO) is a protocol published by IETF in September 2014 (RFC 7285). The two main information elements provided by an ALTO server to ALTO clients are network and cost maps. A network map consists of host-groups. Cost maps define one-way connections between host-groups and assign cost values to each one-way connection. Ranking of endpoints is based on multiple cost types. ALTO service provides appropriate level of abstraction of network and cost maps, enforcing the policies of network operators and content distributors, but keeping the privacy of network topology information.

In Fig. 2, we present a new scenario where the ALTO client is integrated in the SDN controller (using the ALTO-Server). The details of this service can be found in [19]. Originally, depending on the use case, the ALTO client should be deployed either in the UE or in the tracker or resource directory of the content distributor. Integration into the SDN controller makes ALTO service transparent for the endpoints; furthermore any application type can be dynamically selected to get ALTO guidance based on operator policies. In SDN transport segments the network elements have built-in support for flow rewrite rules, which are needed by redirection mechanisms.

The ALTO server function could be part of the Orchestrator or the virtual infrastructure manager in the NFV framework and could run as a VNF or a network function. ALTO servers should collect abstracted network information from network controllers and content distributors. Network controllers should provide dynamic network information [19]. The Orchestrator can dynamically set ALTO service policies. An important challenge is the dynamic provisioning and merging of the network and cost maps from content distribution networks.

E. Quality of Service/Experience Enforcement

In order to achieve acceptable service quality, the broad spectrum of mobile Internet services requires differentiated handling and forwarding of the respective traffic flows. 3GPP allows for such service differentiation by means of dedicated GPRS Tunneling Protocol (GTP) tunnels, which are specifically set up and potentially updated as the mixture of client initiated service consumption changes. The ISAAR (Internet Service quality Assessment and Automatic Reaction) framework [20] augments existing quality of service functions in mobile networks by a flow based and network centric Quality of Experience (QoE) monitoring and enforcement. It is meant to be an all in one solution for service quality monitoring and QoE enforcement within an operator network. It consist of three functional blocks: Quality Monitoring (QMON), Quality Rules Entity (QRULE) and Quality Enforcement (QENF). The task of QMON is flow detection and assessment as well as QoE estimation. Within QRULE enforcement policy rules for each observed flow are created and QENF performs the flow manipulation according to these rules. The architecture and the operation of the original ISAAR framework without SDN and NFV support are described in detail in [20].

The current version of ISAAR makes use of SDN and NFV concepts to enhance its flexibility. SDN is applied for traffic identification as well as for steering the traffic of interest to a monitoring node. NFV is used to virtualize the functions of ISAAR like monitoring or enforcement and to dynamically place these functions at different network and cloud locations.

F. Media Independent Mobility Management in SDMNs

NFV/SDN technologies introduce enormous flexibility to support and deploy novel mobility management schemes. Only a few existing articles have analyzed the potential of such a dynamically manageable architecture in addressing the problems of futures heterogeneous wireless setups. Guimaraes et al. [21] enhance OpenFlow with IEEE 802.21 Media Independent Handover (MIH) [22] capabilities to optimize link connectivity establishment in SDNs, but they do not consider complete 3GPP or ETSI NFV architectures in their work. Although 3GPP has already standardized techniques for vertical mobility, these procedures are limited in using cross-layer information for mobility optimization and far from being media independent, proactive and seamless. In [23], Knaesel et al. propose a MIH-enabled Evolved Packet System (EPS) to provide seamless handovers between 3GPP and non-3GPP wireless technologies. However, their proposal does not consider programmable and virtualized mobile architectures.

Our concept maps IEEE 802.21 MIH functions into the proposed NFV architecture and integrates an OpenFlow-based handover execution scheme, in order to provide efficient, proactive, fine grained, QoS/QoE-aware and seamless mobility management. MIH is used to optimize handovers among different access networks, while OF configures network resources to proactively establish and manage communication paths for user level data flows. As Fig. 2 shows, the proposed SDN controller extensions gather mobility related cross-layer information by relying on dynamic information exchange using MIH event/command services, static data of MIIS (Media Independent Information Service) and the policy control of PCRF (all implemented as VNFs in the telco cloud). Note that MME should also became MIH-aware, as hand over (HO) signaling and decision procedures of 3GPP technologies are mainly handled by this node.

Based on MIH protocol messages and also using data from other SDN controller modules, the architecture manages resources during handover events, and configures flow paths in a proactive and highly dynamic manner according to the actual connectivity options. Using this concept, the impact of inter-technology handovers on the user flows can be minimized, the user and data planes of mobility management can be separated, optimal transmission routes can be continuously maintained and flow-level decisions can be made and executed to support efficient offloading situations. A challenge still to be addressed here is the orchestration of multiple runs of optimization, i.e., how to harmonize decisions of HO events triggered by link going down indications due to user mobility with HO events triggered by other control mechanisms such as network initiated smart offloading algorithms.

G. SDN-enabled Cognitive Radio Network

Cognitive networking brings automation in communication networks by sensing contextual changes in the network, adapting to those contextual changes and applying control loop systems to learn and update itself for future actions without an intervention. Thus, cognitive networks have the potential to provide high bandwidth, adaptive and robust communication system through their ability to observe the current state of the network, analyze it, and adapt to the available resources in an efficient manner. Besides being self-aware, cognitive networks need self-configuration to immediately and dynamically adapt to the network environment [24].

Cognitive networking has not been realized into practice due to the challenges lying in the current architectures besides its own intricacies. For example, cognitive networks require Software Adaptable Network (SAN) elements that consist of network elements which are tunable at run time. SAN elements constitute the building block of cognitive networking that provide physical control of networked systems and action space for the cognitive process. However, the currently deployed network architectures are tightly coupled with the underlying hardware making the cognitive notion of networking really difficult to implement and deploy [24]. Due to the tight coupling of the control and forwarding planes, the network elements are not tunable at run-time.

Therefore, we propose the integration of cognitive networking with SDN to achieve self-configuration and automation. SDN offers programmability of network elements and global visibility of the network state in the centralized control plane. A NOS in the SDN control plane maps the entire network to different services and functions that can be implemented as SDN applications. The programmability enhances the dynamism of communication networks by introducing programmable interface in networking equipment and the global visibility enables SDNs to react abruptly to the changing network environments.

From the implementation perspectives, the SDN controller implements cognition by leveraging cognitive algorithms and a role manager. These entities are deployed in the cognitive engine that interact with the core network as depicted in Fig. 2. The role manager is the function that synchronizes the user sessions among different network segments. The centralized network architecture facilitates the proposed solution that enhances the management efficiency and resource optimization. Henceforth, the benefits delivered to mobile users include network independency, seamless roaming, resource sharing, enhanced redundancy, improved bandwidth, and coverage. One question yet to be addressed is the QoS management where the same level of QoE is guaranteed across technologically different communication domains. The basic integrated architecture has been implemented in a testbed and the results are explained in the following section.

IV. IMPLEMENTATION AND RESULTS

In this section, the proposed architecture components are evaluated through testbed implementations, simulations and/or theoretical evaluations. Various experiments are carried out to analyze the performance of each of the proposed architecture components. The experimental setups and results present the preliminary performance of the proposed schemes in a standalone manner (i.e., without sophisticated integration of the particular schemes). The details of the experimental setups, results from the experiments and analysis are as follows.

A. Software Defined Cognitive Network

The software-defined cognitive networking architecture is composed of SDN enabled WiFi access points, OpenFlow SDN controller and the cognitive radio access network as shown in Fig. 3. The OpenFlow domain consists of OF-enabled wireless access points that forward traffic in-between wireless and LAN interfaces. The SDN controller is implemented with a common bridge between LAN, wireless LAN and Cognitive network. Three Linux-Enriched Wireless open-Access Research Platform (LE-WARP) boards [25] working on 2.4 GHz and 5 GHz ISM band were used in the testbed. One WARP board worked as a cognitive base station and two were used as cognitive clients. The Cognitive network is a conceptual realization of the Software-Defined Radio (SDR) enabled with a separate functional element namely Cognitive Engine. The cognitive engine resides in a laptop attached to the base station and the controller.

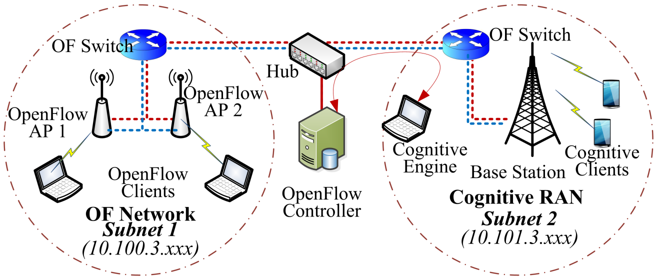


Fig. 3. SDN enabled WiFi and Cognitive RAN testbed.

We conducted two sets of experiments. In the first set of experiments we used a single frequency slot and in the second experiments we used eight frequency slots to see performance improvement due to the cognition process among eight frequency channels. Since the main aim of the experiments is enabling cognition from the centralized control plane of the SDMN architecture, we show the results achieved from the cognition process among the eight frequency slots in Fig. 4.

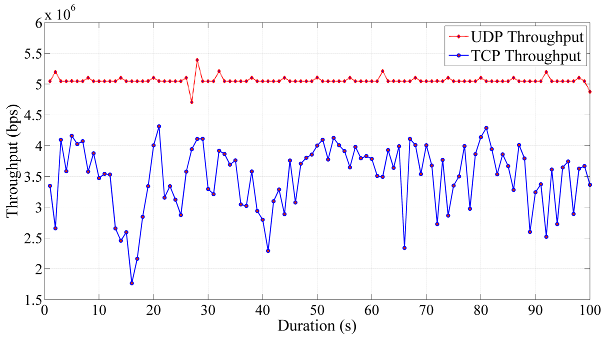


Fig. 4. Testbed throughput results.

We performed TCP and UDP throughput experiments by using IPerf [26] to collect statistics between the OpenFlow wireless and cognitive radio clients. The average UDP and TCP throughput using a single frequency slot remained approximately 1.05 Mbps, whereas the average UDP and TCP throughput for the eight frequency slots remained 5.15 Mbps and 3.5 Mbps respectively. We measured the throughput for the duration of 100 seconds. The results of the experiment reveal that the centralized control provided by SDN enables frequency cognition on one hand, and interworking of heterogeneous radio access networks on the other hand.

B. Optimized Resource and Traffic management

The optimized resource and traffic management algorithm has been tested on the following topology shown in Fig. 5. The test setup consists of 20 OpenFlow switches and three base stations. The link capacities are 0.1 Gbps, 0.2 Gbps, 0.5 Gbps and 1 Gbps.

Using this topology, we have tested several scenarios for flow requirements with different bandwidth requests. In Fig. 6,

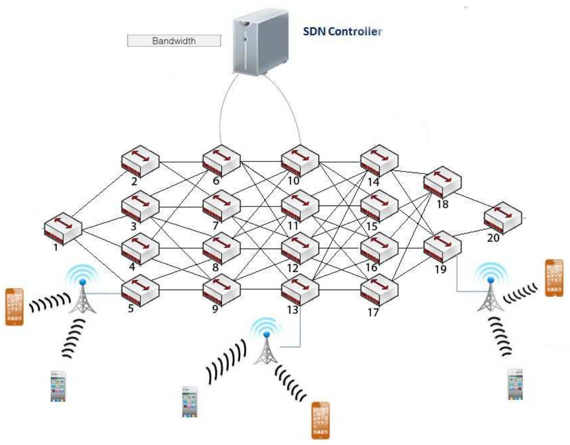


Fig. 5. Network Model for Optimized Resource and Traffic Management.

we have demonstrated that the power consumption can be reduced by about 75% in the network by applying the proposed algorithm which shuts down redundant switches in the network.

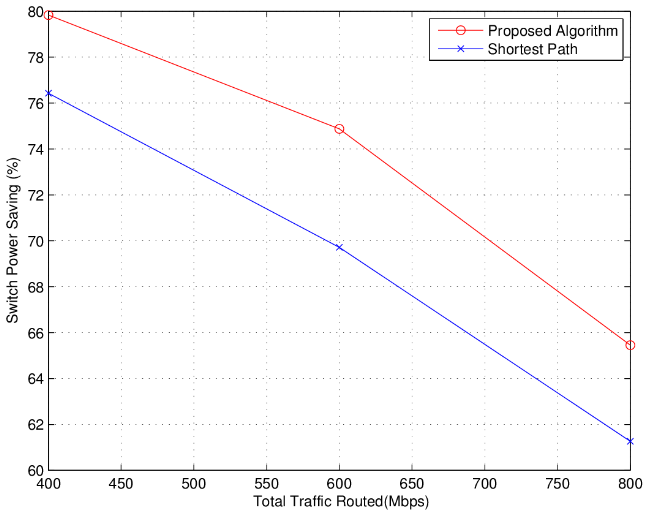


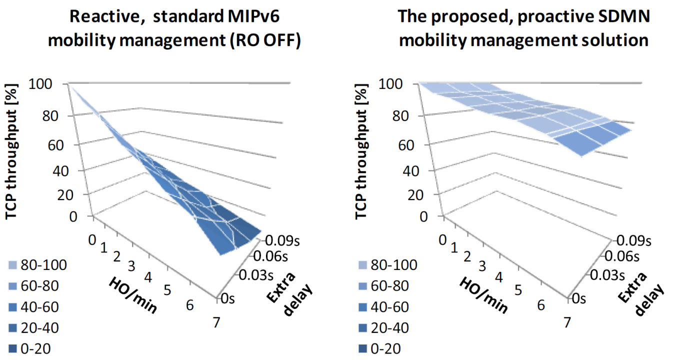
Fig. 6. Performance results of the proposed scenario with throughput requirements of all flows being 100 Mbps.

C. Proactive and Media Independent Mobility Management for SDMNs

The main scenario for the initial performance evaluation of the proposed OpenFlow-based SDMN mobility management focuses on highlighting the advances of the proactive behavior in a simulated network environment. The applied framework for the experiments is INET/OMNeT++ with the OFOMNET OpenFlow model extension [27]. In order to provide a highly configurable, extensible, and adequate model for other legacy and advanced mobility management schemes, we have integrated OFOMNET into our IPv6-based Host Identity Protocol simulation framework called HIPSim++ [28]. The model is built on the top of the 1.99.3 version of INET which is an extension and TCP/IP model collection of the component-based, modular OMNeT++ 4.2 discrete event simulation environment [29].

The mobile terminal moves between different wireless access points (APs) in the simulated mobile architecture. Therefore, it loses and builds up radio connections during the evaluation period according to the movement path and access coverages. The OpenFlow-based mobility management handles handover events based on the MIH supported cross-layer messages received by the controller from the mobiles, APs and the OF switches: the SDN forwarding plane configures flow paths in a proactive and highly dynamic manner according to the actual connectivity and other context information.

In our simulations legacy Mobile IPv6 is used as the comparison base and several key performance indicators are applied, like Handover Latency, UDP Packet Loss, different objective quality of experience metrics and TCP throughput. Fig. 7 depicts the results in regard of the latter: TCP throughput proportion of the two protocols under analysis (i.e., standard MIPv6 without Routing Optimization and our proposed solution) in a one minute communication session between the mobile node and its correspondent entity performed at different handover frequencies from 0 to 7 and different extra delay values between the access and core network entities (0 - 90 ms). The gain of our proactive, MIH-based solution is eye-catching especially when the circumstances are deteriorating (i.e., the number of handovers is increasing). In case of the highest handover frequency, the proactive OpenFlow-based SDMN solution shows more than 350% increase in TCP throughput. Moreover, the average gain of our advanced NFV-integrated scheme is above 100%.


Reactive, standard MIPv6 mobility management (RO OFF)

The proposed, proactive SDMN mobility management solution

Fig. 7. Standard MIPv6-like (left) vs. proactive OpenFlow-based (right) mobility management performance.

V. CONCLUSION

This paper has highlighted several important challenges of traffic, resource and mobility management in mobile networks and proposed new concepts for tackling them. There exist opportunities in SDN and NFV, however, a unified approach to leverage those technologies in the wireless domain in general and in cellular networks in particular is lacking. The elasticity of SDN/NFV technologies provides great flexibility for deploying new services, simplifying network management, and enabling more efficient resource utilization. This paper is a leap towards SDN and NFV based mobile network architecture to reap the benefits of both technologies in a unified approach. We have proposed SDMN architecture components to overcome the highlighted limitations in current wireless networks and provide a way-ahead for future mobile networks. The proposed architecture components are tested in real testbeds or simulation environments to show the feasibility of these components. In the future we will assess the dependencies and gains of joint operation of the components and define the necessary interfaces for integration.

ACKNOWLEDGMENT

This work has been performed in the framework of the CELTIC-Plus project C2012/2-5 SIGMONA. The organizations on the source list would like to acknowledge the contributions of their colleagues to the project.

REFERENCES

[1] M. Liyanage, A. Gurtov, and M. Ylianttila, *Software Defined Mobile Networks (SDMN): Beyond LTE Network Architecture*. John Wiley & Sons, 2015.

[2] 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. ACM, 2013, pp. 25–30.

[3] 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*. ACM, 2013, pp. 163–174.

[4] Developing floodlight modules. Floodlight OpenFlow controller. Big Switch. [Online]. Available: http://www.projectfloodlight.org/floodlight/

[5] L. E. Li, Z. M. Mao, and J. Rexford, "CellSDN: Software-defined cellular networks," *Computer Science, Princeton University, Princeton, NJ, USA, Tech. rep*, 2012.

[6] J. Costa-Requena, J. Llorente Santos, V. Ferrer Guasch, K. Ahokas, G. Premsankar, S. Luukkainen, I. Ahmad, M. Liyanage, M. Ylianttila, O. Lopez Perez, M. Uriarte Itzazelaia, and E. Montes de Oca, "SDN and NFV integration in generalized mobile network architecture," in *Networks and Communications (EuCNC), 2015 European Conference on*, June 2015, pp. 154–158.

[7] I. Ahmad, S. Namal, M. Ylianttila, and A. Gurtov, "Security in software defined networks: A survey," *Communications Surveys Tutorials, IEEE*, vol. PP, no. 99, pp. 1–1, 2015.

[8] ETSI NFV ISG, Published specifications. ETSI NFV Industry Specification Group. [Online]. Available: http://docbox.etsi.org/ISG/NFV/Open

[9] T. Feng, J. Bi, and K. Wang, "Joint allocation and scheduling of network resource for multiple control applications in SDN," in *Network Operations and Management Symposium (NOMS), 2014 IEEE*, May 2014, pp. 1–7.

[10] H. Egilmez and A. Tekalp, "Distributed QoS Architectures for Multimedia Streaming Over Software Defined Networks," *Multimedia, IEEE Transactions on*, vol. 16, no. 6, pp. 1597–1609, Oct 2014.

[11] J. Zhang, K. Xi, M. Luo, and H. Chao, "Load balancing for multiple traffic matrices using SDN hybrid routing," in *High Performance Switching and Routing (HPSR), 2014 IEEE 15th International Conference on*, July 2014, pp. 44–49.

[12] Architecture Enhancement for Flexible Mobile Service Steering. 3GPP Specification. [Online]. Available: http://www.3gpp.org/DynaReport/23718.htm

[13] J. Halpern and C. Pignataro, "Service Function Chaining (SFC) Architecture," *draft-ietf-sfc-architecture-07 (work in progress)*, 2015.

[14] R. Cannistra, B. Carle, M. Johnson, J. Kapadia, Z. Meath, M. Miller, D. Young, C. Decusatis, T. Bundy, G. Zussman, K. Bergman, A. Carranza, C. Sher-DeCusatis, A. Pletch, and R. Ransom, "Enabling autonomic provisioning in SDN cloud networks with NFV service chaining," in *Optical Fiber Communications Conference and Exhibition (OFC), 2014*, March 2014, pp. 1–3.

[15] B. Martini, D. Adami, A. Sgambelluri, M. Gharbaoui, L. Donatini, S. Giordano, and P. Castoldi, "An SDN orchestrator for resources chaining in cloud data centers," in *Networks and Communications (EuCNC), 2014 European Conference on*. IEEE, 2014, pp. 1–5.

[16] A. Baumgartner, V. Reddy, and T. Bauschert, "Mobile core network virtualization: A model for combined virtual core network function placement and topology optimization," in *Network Softwarization (NetSoft), 2015 1st IEEE Conference on*, April 2015, pp. 1–9.

[17] S. Orlowski, M. Pióro, A. Tomaszewski, and R. Wessäly, "Sndlib 1.0survivable network design library (2007) optimization online preprint," in *Proceedings of the 3rd International Network Optimization Conference (INOC 2007)*, 2007, pp. 163–174. [Online]. Available: http://sndlib.zib.de

[18] V. K. Gurbani, V. Hilt, I. Rimac, M. Tomsu, and E. Marocco, "A survey of research on the application-layer traffic optimization problem and the need for layer cooperation," *Communications Magazine, IEEE*, vol. 47, no. 8, pp. 107–112, 2009.

[19] Z. Faigl, Z. Szabo, and R. Schulcz, "Application-layer traffic optimization in software-defined mobile networks: A proof-of-concept implementation," in *Telecommunications Network Strategy and Planning Symposium (Networks), 2014 16th International*. IEEE, 2014, pp. 1–6.

[20] M. Eckert and T. M. Knoll, "ISAAR (Internet Service Quality Assessment and Automatic Reaction) a QoE monitoring and enforcement framework for internet services in mobile networks," in *Mobile Networks and Management*. Springer, 2013, pp. 57–70.

[21] C. Guimaraes, D. Corujo, R. Aguiar, F. Silva, and P. Frosi, "Empowering software defined wireless networks through media independent handover management," in *Global Communications Conference (GLOBECOM), 2013 IEEE*, Dec 2013, pp. 2204–2209.

[22] I. . W. Group *et al.*, "IEEE standard for local and metropolitan area networksPart 21: Media independent handover," *IEEE Std*, vol. 802, p. 2008, 2009.

[23] F. J. Knaesel, P. Neves, and S. Sargento, "Ieee 802.21 mih-enabled evolved packet system architecture," in *Mobile Networks and Management*. Springer, 2012, pp. 61–75.

[24] I. Ahmad, S. Namal, M. Ylianttila, and A. Gurtov, "Towards software defined cognitive networking," in *New Technologies, Mobility and Security (NTMS), 2015 7th International Conference on. IEEE*, 2015, pp. 1–5.

[25] M. Jokinen and H. Tuomivaara, "LE-WARP: Linux enriched design for wireless open-access research platform," in *Proceedings of the 4th International Conference on Cognitive Radio and Advanced Spectrum Management*. ACM, 2011, p. 16.

[26] A. Tirumala, F. Qin, J. Dugan, J. Ferguson, and K. Gibbs, "Iperf: The TCP/UDP bandwidth measurement tool," *http://dast.nlanr.net/Projects*, 2005.

[27] D. Klein and M. Jarschel, "An openflow extension for the omnet++ inet framework," in *Proceedings of the 6th International ICST Conference on Simulation Tools and Techniques*, ser. SimuTools '13. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013, pp. 322–329.

[28] L. Bokor, S. Nováczki, L. T. Zeke, and G. Jeney, "Design and Evaluation of Host Identity Protocol (HIP) Simulation Framework for INET/OMNeT++," in *Proceedings of the 12th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, ser. MSWiM '09. ACM, 2009, pp. 124–133.

[29] A. Varga and R. Hornig, "An Overview of the OMNeT++ Simulation Environment," in *Proceedings of the 1st International Conference on Simulation Tools and Techniques for Communications, Networks and Systems & Workshops*, ser. Simutools '08. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008, pp. 60:1–60:10.