Project Number:	**CELTIC / CP2012-2-5**
Project Title:	SDN Concept in Generalized Mobile Network Architectures– SIGMONA
Document Type:	I (Internal)

Document Identifier:	D 1.1
Document Title:	**Consolidated view of SIGMONA Software Defined Mobile Network Architecture**
Source Activity:	WP 1 - Software Defined Mobile Network Architecture
Main Editor:	Tero Lötjönen
Authors:	All partners contributed, see the list of main authors below
Status / Version:	First release / 1.0
Date Last changes:	05.12.2014
File Name:	SoftwareDefinedMobileNetwork_Architecture-D1.1.doc

Abstract:	This document presents the project view on the software defined mobile network architecture.

Keywords:	Architecture, SDMN

Document History:	
7.1.2014	Document created
6.2.2014	Improved version to BSCW, under version control
6.11.2014	Version for the WP1 internal review
5.12.2014	Approved first release

Table of Contents

Authors

Partner	Name	Phone / Fax / e-mail
AALTO	Jose Costa-Requena	
		Phone: +358 5 0577 0142
		e-mail: jose@netlab.tkk.fi
Nokia Networks Finland		
	Tero Lötjönen	
		Phone: +358 40 5747842
		e-mail: tero.lotjonen@nsn.com
	Pekka Korja	
		Phone:
		e-mail: pekka.korja@nsn.com
	Jukka Salo	
		Phone:
		e-mail: jukka.salo@nsn.com
Nokia Networks Hungary	Zoltan Vincze	
		Phone: +36 20 977 7797
		e-mail: zoltan.vincze@nsn.com
Technical University of Budapest – Mobile Innovation Centre		
	László Bokor	
		Phone: +36 1 4633420
		e-mail: bokorl@hit.bme.hu
	Zoltán Faigl	
		Phone: +36 1 4633420
		e-mail: zfaigl@mik.bme.hu
Bull SAS		
	Gérard Jacquet	
		Phone:
		e-mail: gerardjacquet@bull.net
	Olivier Jard	
		Phone:
		e-mail: olivier.jard@bull.net
Coriant	Juha-Petteri Nieminen	
		Phone: +358 40 41312215
		e-mail: juha-petteri.nieminen@tellabs.com

INGENIA	Raul Moreno
	Phone:
	e-mail:

TT Argela		
	Aydın Ulaş	
	Phone:	+90 212 707 1259
	e-mail:	aydin.ulas@argela.com.tr
	Serdar Tan	
	Phone:	
	e-mail:	

AVEA	Engin Zeydan
	Phone: +90 216 987 6386
	e-mail: engin.zeydan@avea.com.tr

Technical University of Chemnitz	Thomas Bauschert
	Phone:
	e-mail:

Montimage	Edgardo Montesdeoca
	Phone: +33684509637
	e-mail: edgardo.montesdeoca@montimage.com

CWC	Suneth Namal
	Phone: +358 41 7282646
	e-mail: gkarunar@ee.oulu.fi

EXFO	Kari Hyväri
	Phone: +358 40 3010317
	e-mail: kari.hyvari@exfo.com

Executive Summary

This document summarizes the project view of Software Defined Mobile Networks (SDMN) architecture. The document includes the mapping of the multiple research contribution from different partners into SDMN architecture. This mapping results in multiple components that utilize SDN as basic technology to deploy the required functionality on areas such as transport, load balancing, security, monitoring, QoS/QoE, resource optimization, etc. Some of the researched components have a considerable impact on the mobile network architecture; some are rather optimizations and virtualisations of the existing functionalities; and some concern additional functionalities needed to support the introduction of SDN and cloud computing (e.g., monitoring and security).

The proposed SDMN architecture options help transform the current rigid and disparate mobile networks into scalable and dynamic ecosystems. SDN is considered the enabler technology for the Software Defined Mobile Network making it is a key element for obtaining more flexible networks that can dynamically adapt to the needs of operators, content and service providers. These various requirements defined for future networks have lead to the definition of a set of basic assumptions and uncertainties , presented in this document, that need to be addressed and validated. The requirements and basic assumptions will be better served with an evolved architecture that maintains the evolution path with legacy systems.

The centralized mobile network control plane in the cloud enables optimal usage of resources and the use of the emerging cloud platform native services. The SDN controlled user plane (mobile network and transport) provides great potential for the optimization of topologically distributed forwarding elements.

List of terms, acronyms and abbreviations

Active communication	In 3GPP, (PS) active communication is defined by the existence of one or more Activated PDP contexts that generate IP traffic to/ from servers or/and end users. For Evolved Packet System (EPS) term EPS bearer context [1] is used. Active communication is required to perform an Activity in the Use Case.
Cloud computing	A model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.
Controller	Logical realization of the control plane where a centralized SW element manages/controls the data plane.
End-to-End (E2E) Architecture	An architecture encompassing all NFs and resources owned by an operator
Hypervisor	Also called virtual machine manager (VMM), is one of many hardware virtualization techniques allowing multiple operating systems, termed guests (VMs), to run concurrently on a host computer. The hypervisor presents the guest operating systems with a virtual operating platform and manages the execution of the guest operating systems.
Independent Software Vendor (ISV)	An ISV makes and sells software products that run on one or more computer hardware or operating system platforms.
LTE network elements	The network components required in LTE control plane such as MME, P/S-GW, PCRF, AAA, etc.
Multi-tenancy	Enables sharing of resources and costs across a large pool of users thus allowing for: Centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.), Peak-load capacity increases (users need not engineer for highest possible load-levels), Utilization and efficiency improvements for systems that are often only 10–20% utilized.
Network Function (NF)	A functional building block within an operator´s network infrastructure, which has well-defined external interfaces and a well-defined functional behavior. Note that the totality of all network functions constitutes the entire network and services infrastructure of an operator/service provider. In practical terms, a Network Function is today often a network node or physical appliance.
NF forwarding graph	A graph specified by a Network Service Provider of bi-directional logical links connecting NF nodes.
NF set	A collection of NFs with unspecified connectivity between them
Network virtualization	The process of combining hardware and software network resources and network functionality into a single, software-based administrative entity, a virtual network. Network virtualization involves platform virtualization, often combined with resource virtualization.
Network Function Virtualization Infrastructure (NFVI)	The Network Function Virtualization Infrastructure (NFVI) contains all the hardware and software components that constitute the environment in which VNFs of the MNO are deployed, managed and executed. The NFVI includes resources for computation, networking and storage.
Resource management	Resources can be physical resources, logical elements, processes in the cloud, etc. Both user plane and control plane resources.
Resource pooling	The provider’s computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources, but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, and network bandwidth.

Software Defined Networking (SDN)	A term used for networks in which the control plane is decoupled from the data plane and made the control plane remotely accessible and remotely modifiable via third-party software clients. SDN requires some method for the control plane to communicate with the switch data path. One such mechanism is OpenFlow protocol.
Virtualization	Hardware virtualization or platform virtualization refers to the creation of a virtual machine (VM) that acts like a real computer with an operating system, but is separated from the underlying hardware resources. Technology allows servers and storage devices to be shared and utilization be increased. Applications can be easily migrated from one physical server to another.
Virtual Machine (VM)	A program and configuration of part of a host computer server. Note that the Virtual Machine inherits the properties of its host computer server e.g. location, network interfaces. A completely isolated guest operating system installation within a normal host operating system. Modern virtual machines are implemented with either software emulation or hardware virtualization or (in most cases) both together.
Virtualised Network Function (VNF)	An implementation of an executable software program that constitutes the whole or a part of an NF that can be deployed on a virtualisation infrastructure.
VNF Forwarding Graph (VNF-FG)	A NF forwarding graph (of logical links connecting NF nodes) where at least one node is a VNF through which network traffic is directed for the purpose of creation a set of network functions. (ETSI NFV term for Service chaining)
VNF set	A NF (Network Function) set where all the NFs are VNFs.
Virtualization levels	Different levels of virtualization; 1) transport network, 2) full LTE network including transport and network elements

AAA	Authentication, Authorization and Accounting
AF	Application Function
AKA	Authentication and Key Agreement
ALTO	Application-Layer Traffic Optimization
AMNS	Automated Mobile Network Slicing
ANDSF	Access Network Discovery and Selection Function
AP	Access Point
API	Application Programming Interface
APN	Access Point Name
CAPEX	Capital Expenditure
CDN	Content Delivery Network
CDR	Charging Data Record
CES	Customer Edge Switching
CGF	Charging Gateway Function
COTS	Commercial off-the-shelf
CQI	Channel Quality Indicator
C-RAN	Cloud Radio Access Network
DB	Database
DMM	Dynamic Mobility Management
DoS	Denial-Of-Service
DPI	Deep Packet Inspection
ECM	EPS Connection Management
EMM	EPS Mobility Management
EMS	Element Management System
eNB	Evolved Node B (eNodeB)
EPC	Evolved Packet Core
ePDG	Evolved Packet Data Gateway

ESM	EPS Session Management
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FCA	Flow Control Agent
GRX	GPRS Roaming eXchange
GTP	GPRS Tunneling Protocol
HIP	Host Identity Protocol
HO	HandOver
HSS	Home Subscriber Server
IDS	Intrusion Detection System
IETF	Internet Engineering Task Force
IFOM	IP Flow Mobility and Seamless WLAN Offloading
ISAAR	Internet Service quality Assessment and Automatic Reaction
IMS	IP Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
IPX	IP eXchange
IPS	Intrusion Prevention System
JSON	JavaScript Object Notation
LMA	Local Mobility Anchor
LTE	Long Term Evolution
MAC	Medium Access Control
MAG	Mobile Access Gateway
MDM	Mobile Device Management
MIH	Media Independent Handover (IEEE 802.21)
MIP	Mobile IP
MM	Mobility Management
MME	Mobility Management Entity
MNO	Mobile Network Operator
MPLS	Multiprotocol Label Switching
NF	Network Function
OCS	Online Charging System
OF	Open Flow
OFCF	Offline Charging Function
OPEX	Operating Expenses
PCEF	Policy and Charging Enforcement Function
PCRF	Policy and Charging Rules Function
P-GW	Packet Data Network (PDN) Gateway
PHB	Per Hop Behavior
PLMN	Public Land Mobile Network
PMIP	Proxy Mobile IP
QoE	Quality of Experience
QoS	Quality of Service
RAT	Radio Access Technology
REST	REpresentational State Transfer
SCTP	Stream Control Transmission Protocol
SDM	Software Defined Monitoring
SDMN	Software Defined Mobile Network
SDN	Software Defined Networking
SIEM	Security Information and Event Management

S-GW	Serving Gateway
SON	Self-Organizing Network
SPAP	Security Policy Administration Point
SSL	Secure Socket Layer
TDF	Traffic Detection Function
TLS	Transport Layer Security
TNO	Transport Network Operator
TRILL	Transparent Interconnection of Lots of Links
UE	User Equipment
UFA	Ultra Flat Architecture
VLAN	Virtual Local Area Network
VM	Virtual Machine
VMNO	Virtual Mobile Network Operator
VoLTE	Voice over LTE
VPN	Virtual Private Network
WLAN	Wireless Local Area Networks

1. Introduction

This document collects each SIGMONA partner's research topics and defines a consolidated common view of the mobile network architecture evolution. This evolution comprises software defined networking and cloud computing to enable rapid provisioning via reconfigurable computing resources (e.g., network infrastructure, storage, applications and services) using standardized components.

1.1 Main drivers, problem statement and general trends for the architecture evolution

The main drivers, for the mobile core (EPC) network evolution to serve the future challenges and define the way towards 5G networks, are the need for high capacity and the low latency. There will be multiple different types of applications with different requirements, which emphasize the need for dynamic scalability of the network functionality. The means for the efficient network resource operability and management seems to be even more important than the future network element costs.

Mobile network operators are facing growing challenges due to the explosive growth in data traffic caused mainly by the prevalence of smartphones and streamed audio and video services. In this new context, the operators need to manage the increase in traffic load, and meet rising consumer and enterprise expectations for excellent performance while providing ubiquitous broadband connectivity. Operators must also roll out new services and applications rapidly to maintain a competitive edge. Slow service rollouts are no longer acceptable. Finally, in every competitive market there is constant pressure to become more efficient; in other words, to maintain or improve performance at a lower operational cost. Existing mobile networks struggle with limitations; such as, stationary and expensive equipment, complex control protocols, and customized and varying configuration interfaces. The SIGMONA project's main goal is to study and apply SDN (Software Defined Networks) technology and principles within the mobile networking environments (referred to as SDMN – Software Define Mobile Networks) to be able to address these current limitations.

Network Functions Virtualization (NFV) and cloud computing are evolving from the typical IT data center applications to new areas. Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services). The main goals of network virtualization are the aggregation of distributed resources for optimized sharing and the utilization of shared pools of configurable computing HW resources for scalable on-demand network access. Cloud computing enables the ability to host network functions; such as, resource, policy, mobility, security, traffic management and monitoring as-a-service within the cloud.

In computation and storage environments there are several emerging technologies and enablers that could bring feasible processing potential to be utilized for the mobile network networks applications. Potential use cases in mobile networks include virtual operator concepts, network sharing/slicing principles and core network element user plane functionality cloudification for the NW control overlay functions.

Software Defined Networking (SDN) is a term used for networks in which the control plane is decoupled from the data plane, and made remotely accessible and modifiable via third-party software clients. SDN requires some method for the control plane to communicate with the switch data path. One such mechanism is the OpenFlow protocol.

Principal benefits of SDN are quite obvious in the area of cloud computing networking, but the objectives in the mobile network area need further study.

1.2 Scope of the document

The target of this document is to present the SIGMONA partners' research in the areas of cloud computing and software defined networking applicable to mobile core and transport networks. It then consolidates these separate topics into different architecture areas to construct a global view of the potential evolution.

This document analyses the different architectural options for integrating SDN and cloud computing principles into mobile networks and compares them in terms of pros and cons. The options mainly differ in the extent to which SDN principles can be applied to mobile specific functions and transport network functions.

The architecture work includes scenarios for cloud resource management and traffic management within the cloud. In the coming phase, the project will investigate the performance and scalability limits of selected architectures via validation and proof-of-concept platforms.

There are different layers and point of views for the future mobile architecture evolution considered here are:

- Mobile (3GPP) network architecture

- Transport network and security architecture
- Cloud architecture, providing the environment and resources (processing, storage...)
- Orchestration architecture: Management and resource provision in a cloudified environment
- NW topology view (where the different functions are located and how this is expected to evolve)

Some of these layers can be nested together and they partly overlap each other. The work presented in this document did not try to handle thoroughly all of these areas and interdependencies. It uses the layer approach to focus on each concrete functionality and considers assumptions relevant to the other layers to obtain a coherent and comprehensible picture of the required architecture evolution.

2. Main SDMN Basic assumptions

This chapter gives a description of items to be considered when defining the SDMN architecture. In the project's initial phase, the architecture evolution basic assumptions were defined including counter assumptions and uncertainties. This was deemed better than trying to list all the requirements in detail since it allowed to better comprehend the advantages and inconveniences of introducing new concepts and technologies that will most likely shape the future of mobile networks. A summary of the main identified topics is presented in the following sub-sections.

2.1 Migration: Compatibility with the legacy systems vs. Clean-slate deployment

The integration of SDN and network function virtualization into LTE should minimize the changes in network elements, thus providing a seamless migration based on operator needs. This allows incremental updates of network elements in certain parts of the network while keeping legacy elements in other parts of the network. Multivendor compatibility is assumed to be the underlying challenge also in SIGMONA. When specifying the new SDN/virtualization/cloud technology-based systems for mobile telecommunications the clean-slate approach, or the full deployment at once, as the opposite of the afore-mentioned is not the choice to be followed in SIGMONA. Besides the legacy support the migration should enable adoption of new technologies being currently defined such as Device-to-Device (D2D) communications.

2.2 Virtualization and running network functions in cloud

LTE EPC functionalities will emerge to virtualization and operators intend running the functionalities in the cloud (at least the control plane functionalities). This should be enabled without preventing the interoperability with existing LTE network elements and related systems (i.e. charging and roaming) and meeting the other properties of the current networks, like delay budget. LTE virtualization should enable dynamic management of computing resources and bandwidth to handle cost effectively the traffic demand.

2.3 Resilience: Fault tolerant SDN based mobile networks

SDN based network should be fault tolerant and fall back to standard behavior (i.e. last best known flow tables/actions or preconfigured flow tables/actions to be used as default) in case of failure. The SDN network should not be prone to misconfiguration and avoid single point of failure based on controllers. The SDN based mobile networks should fulfill similar reliability to current mobile networks of 99,999% availability. The flexibility of the SDN concept might allow finding a good tradeoff between complexity and reliability. The controller can move redundant modules of the network to reinforce reliability in parts as needed, thus providing sufficient reliability with reasonable complexity.

2.4 QoS provision in virtualized mobile core networks in SDN-based forwarding paths

The usage of policy control is optional in operator networks, therefore it depends on the operators' decision whether it will be applied in the future. Till this time the RAN part of the network represented the bottleneck, hence QoS guarantees had to be applied in RANs. With respect to the support of service guarantees by the transport networks, over provisioning is the mostly applied method.

Our assumption in SIGMONA project is that dynamic, service data flow-based policy control will be more and more needed by mobile network operators due to the increasing diversity of services and the related policy rules. Hence, in general, the QoS provisioning mechanisms specified by 3GPP, such as EPS bearers or PDP contexts and policy control by PCRF should be kept also in case of virtualization of mobile core and transport network.

The service-chaining concept requires network function forwarding graphs both through virtual and traditional transport network segments. Operators need to be able to control logical and physical interconnections, configure traffic class conditioning and forwarding behaviors (capacity, priority, packet loss, delay, shaping, dropping etc.), and to map traffic flows to appropriate forwarding behaviors.

2.5 SDN based mobility management versus 3GPP and MIP mobility

Deployment of virtualization and software defined networking technologies are foreseen mainly in the transport level, and user/data plane procedures will presumably remain intact. However, there are ongoing researches focusing on the possibilities of using SDN technologies for UE-level mobility management: OpenRoads and SDN-TRILL are evolving toward flow-level SDN mobility protocols executed by SDN switches and controllers based on pre-configured or dynamic policies.

Our assumption with respect to mobility management is that SDN-based transport networks may provide new potential for micro-mobility. E.g., they may replace GTP-based tunneling within the SDN domain of a controller. On the other

hand, macro-mobility events, typically occurring above SDN domain level can be based on distributed mobility management solutions adapted to SDN (e.g., handover to untrusted non-3GPP networks).

2.6 Locator and identity assignment to UEs in SDMNs

Our assumption in SIGMONA project is that the practice of IP address distribution to users will not change, even if the S/P-GW control plane and DHCP servers and relays will run as virtual network function, at changing physical locations in the cloud. Many IP mobility management solutions support identity/locator split, in order to introduce long-term identities for the participants. Such solutions have their benefits in addressing the peers on application-layer.

The counter-assumption is that new IP address assignment concept is needed for SDMNs to better adapt to the real location of virtual network functions, which allocate IP address for the UE, and the UEs.

2.7 Security and Traffic Management Synchronization coordination

Network security must be considered parallel to the traffic management and as an integral part of the network management. Traffic and security synchronization is necessary so that security of links, nodes and UEs is in place with changes in network topology, infrastructure elements, link failures and node mobility etc. On the other hand with deploying new security procedures or lapses in controller security or security breaches in the network should not affect the overall network traffic.

In SDN, synchronization becomes very critical since a security lapse of the centralized controller can effect setting the (traffic) flow rules on other data path elements managed by the same controller. Hence our assumption is that network security and traffic procedures and policies must be synchronized.

2.8 Optimized security setup constraints to reduce delays

As a security concern, if there are false rules inserted in the switches and it is not inspected and/or recovered within certain time, it might lead to severe security lapses in terms of data theft and inefficient use of network resources. Our assumption is that timely inspection of the flow rules in data-path elements must be incorporated to the controller. Added to that, the security procedures must not take long enough to hinder seamless connectivity.

2.9 Managing the security functions of physical and virtual elements and interfaces

Security in virtual and SDMN will be managed using centralized controllers. This necessitates that the controller must be secured and security of virtual and physical network elements and interfaces needs to be managed and assured. Virtual and software defined network techniques make it easier to modify and configure network functions using centralized controllers, making it not necessary to intervene directly on different network elements. This makes a controller a critical element in the network that needs to be secured, guaranteeing high availability at all times.

The security management is referring to the way of controlling/managing the security function in the physical or virtual network interfaces (i.e. doing the packet inspection) and what is the relationship with the virtualized elements (i.e. SDN controller, or virtual FW running on the cloud). The virtual security components need to identify to which virtual network the packets belong to in order to apply the required security management/policy.

It is assumed that besides securing the controller, the virtual network elements and interfaces are secured, and both MNOs and Virtual MNOs can define and manage their security policies.

2.10 Security cooperation between edge nodes and operators towards global trust

This assumption considers that cooperation between edge nodes and operators will provide global trust across operators. The current architecture of the Internet makes it expensive to attribute evidence of malicious behavior to a given host due to the possibility of IP address spoofing.

This assumption proposes leveraging the edge nodes as a connection and security broker for the hosts that it serves and establishes the foundations for trusted communications. As a result, the concept of trust aims at delivering enhanced security in future networks in order to curb anti-social behavior allowing for more educated decisions as the edge node has a more complete vision of the network.

2.11 Regulation for more competition vs. Let the market forces decide

There is a broad agreement that the EU mobile wireless market is underperforming relative to other advanced economies (like U.S.). A reason for underperforming is the market fragmentation which prevents the operators in EU from capturing the economies of scale and scope. To shape the course of the industry the Regulator will promote the deployment of such new technologies, which enable new market structures, and hence, more efficient use of investments, new entries and competition, and innovation.

The counter-assumption is that the industrial players deploy the new technologies and open, or do not open the business to new players without any reaction from the Regulator side.

2.12 Privacy and trust regulation driving force

In the cloud environment, where the data centers may locate in different countries and regions, instead of companies' own sites, there is a significant tension between the financial and efficiency benefits the cloud services offer and the risks that such services may pose to an individual's privacy or personal data. The challenge is to balance the interests of different stakeholders to arrive at a pragmatic approach to regulation that is consistent, clear and proportionate.

For supporting the deployment of the new telco clouds and the financial and efficiency benefits that they may bring, the policy-makers and regulators will pay attention and put efforts to carry out a comprehensive reform of the EU data protection rules for ensuring privacy and assuring information security. The common rules across regions may be difficult to agree upon, however.

The counter-assumption is that the industries, without any attention by the regulators, can develop principles and practices that reflect consensus on the best approach to privacy.

2.13 Network monitoring adapted to network virtualization

Network monitoring facilitates verification and validation of SLAs, managing performance (QoS) and user experience (QoE), troubleshooting, and assessment of optimizations and use of resources. On the one hand network virtualization sets new requirements for mobile network monitoring but on the other and it also provides means for implementing advanced network monitoring solutions. NFV/SDN enables the integration of cloud infrastructure that can provide high degrees of freedom regarding the placement of measurement points and the flexible control of traffic flows. An advanced and effective QoS monitoring solution should comprise both a distributed (SDN / NFV-based) QoS measurement system and a centralized evaluation system.

2.14 Service provisioning and optimization orchestrator entities

In SDN networks control applications have full view of network configuration; this together with status information provided by network monitoring and data collection systems enables mobile network orchestrator application to optimize service (e.g. latency) and/or resource usage easier than traditional networks that need to rely on signaling. The orchestrator can carry out via the control applications controlling multiple network elements, potentially from multiple vendors. This enables to introduce new services by writing or modifying the orchestrator whereas in traditional networks, all equipment needs to be upgraded to support the new service type. It is assumed that SDMN will not be implemented in a clean-slate approach, legacy and SDN network management solutions will coexist over time. In order to exploit the potentials of SDN it is required that legacy and SDN management solutions cooperate (e.g. by introducing abstraction and automation layer for legacy network part, etc.).

2.15 Availability of resources

Heterogeneity of resources and dynamicity of them are a challenge for network configuration awareness by means of availability of resources. Both physical and virtualized resources operators require to have updated information on available inventory, topology, capacity, resilience and optionally security assurance assessment. The resource availability awareness is the foundation for resource management process, service provisioning and optimization, traffic management as well as security protection and monitoring process.

2.16 Cost reduction impact of LTE network virtualization

Virtualization of the LTE network is assumed to provide cost savings from standardized network elements and higher capacity utilization, and SDN provides for easier network management due to the separation of data and control planes. However, virtualized network elements may increase the need for more computing power, more complex network management and create more complex value networks. The net benefit of SDN in LTE networks should be examined further in SIGMONA.

3. Architecture model

There was generated couple of generic architecture models, in order to align the various partner research topics under the similar definition, to allow consolidation and to find out the commonalities in the introduced reference points etc.

3.1 Architecture model background

There are various industry activities ongoing around the topics related to virtualized and software defined mobile network. A part of these has been taken as references for the network architecture work in the SIGMONA project. The aim of this chapter is to further define the SIGMONA work scope related to these other industry consortiums and their approach to the architecture definition.

3.1.1 ETSI ISG NFV

The ETSI Industry Specification Group Network Function Virtualization is an industry consortium with over 200 member companies, including over 30 network operator companies [2]. NFV is defining the end-to-end architecture which focuses on making the deployment and management of virtualized network functions different to the current physical network appliance or network function. The functional model identifies the main functional groups, which are required in order to realize NFV. The Figure 1 shows high-level functions and their interfaces. For a detailed description of each function, see ETSI E2E Architectural Framework document [3].

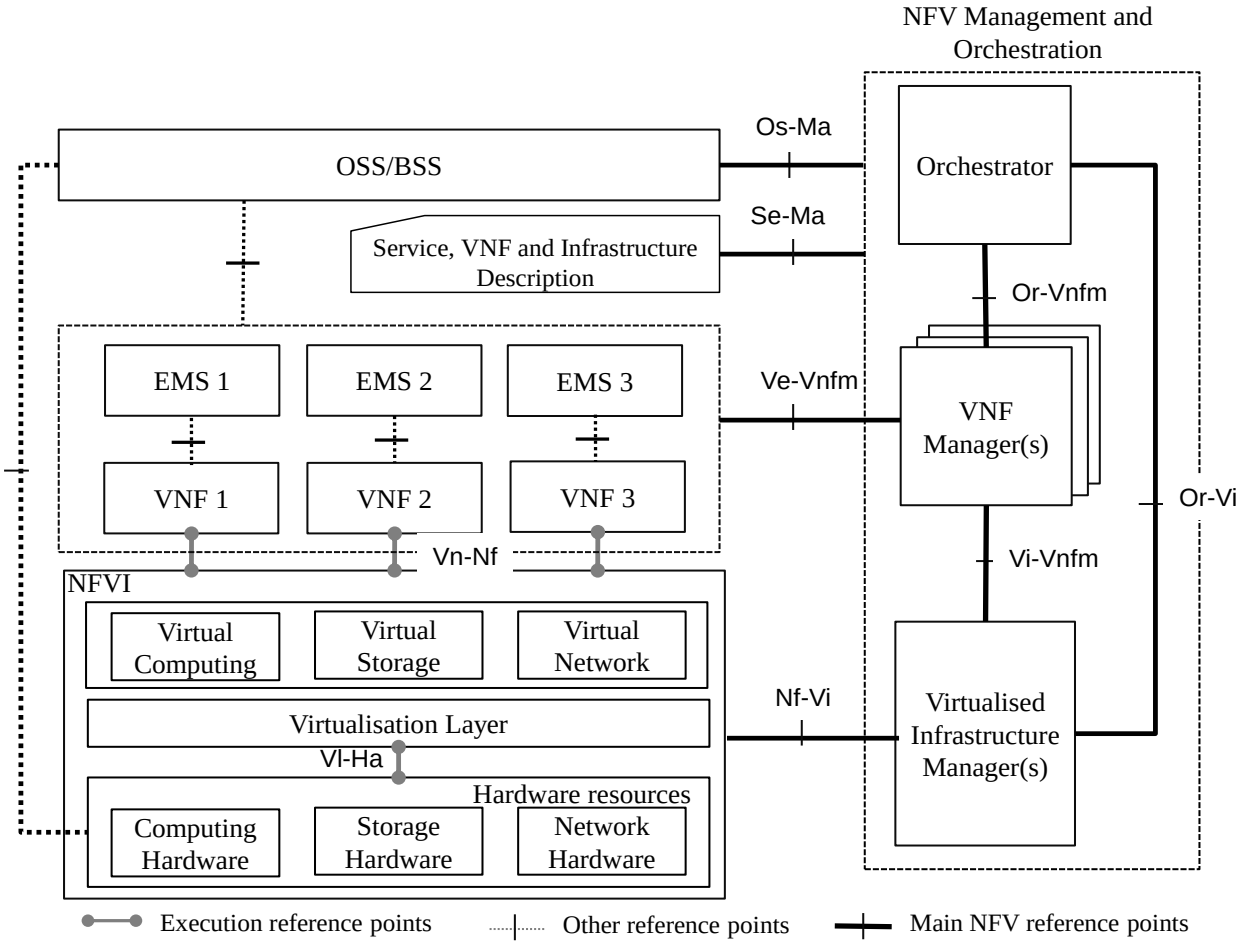

Figure 1: NFV E2E Architecture Overview

NFV is about creating virtual network functions and managing and orchestrating them on a virtualized infrastructure. NFV approach could be defined as “Evolution of network elements”. The ETSI ISG NFV documents reached stable status in June 2014, and they were moved to change control. The draft documents are available in the ETSI open area [4]:

- Infrastructure overview
- Hypervisor domain
- Compute domain
- Infrastructure Network domain
- Infrastructure Interfaces and abstractions
- Software Architecture: VNF Architecture
- Management and Orchestration (End of June)
- Performance and portability best practice
- Security problem statement

Phase 1 is expected to be completed by December 2014.

NFV is highly complementary to SDN. These topics are mutually beneficial but are not dependent on each other. Network functions can be virtualized and deployed without an SDN being required and vice-versa [3].

The OpenStack community has a mission to produce an open standard cloud computing platform for both public and private cloud providers regardless of the size [5]. OpenStack is a suite of software tools that is seen as one of the means to provide the orchestration layer for NFV infrastructure, aimed at building and managing cloud networks. But the challenge is that OpenStack alone might not be enough for a carrier-grade NFV orchestration as telecom applications span over multiple virtual machines, especially when it comes to a carrier-grade NFV deployment.

3.1.1.1 Relation to SIGMONA work scope

The Management and Orchestration framework provides the provisioning of virtualized network functions (VNF), their operations, the configuration of the virtualized networking functions and the infrastructure these functions run on. The virtualized networking infrastructure comprises functional blocks from e.g. the hypervisor, compute and infrastructure networking domains. The virtual network functional block contained in the NFVI as defined in Figure 1 is relevant to the SIGMONA work scope. However the NFVI must be considered in the context of the whole NFV due to the architecture depicted in Figure 1.

3.1.2 Open Networking Foundation

SDN is about creating, orchestrating and managing logical network topologies; virtualized connections between network functions and VNFs. SDN and ONF approach could be defined as “Evolution of networking”. The ONF Wireless Mobile WG is focusing on the application of SDN and OpenFlow to mobile packet core (3GPP EPC), wireless (backhaul) transport and enterprise unified (wired and wireless) access. The mobile packet core project team is working on the architecture and call flows and potentially required OpenFlow extensions for SDN based EPC, mobility management and service chaining. The architecture should have been finalized 1H 2014.

3.1.2.1 Relation to SIGMONA work scope

ONF-based SDN architectures inherit a number of benefits for mobile and wireless environments, including their wireless access, mobile backhaul, and core networking segments.

The paradigm of flow-based communication in SDN architectures fits well to provide efficient end-to-end communications in multi-access environments, when different radio technologies, like 3G, 4G, WiMAX, Wi-Fi, etc. are simultaneously available for users. SDN is able to provide fine-grained user flow management aiming to improve traffic isolation, QoS/QoE provision and service chaining.

Centralized control plane allows for efficient resource coordination of wireless access nodes, which makes possible to implement efficient inter-cell interference management techniques.

The fine-grained path management in SDN networks provides various optimization possibilities based on the individual service needs and independently from the configuration of the underlying routing infrastructure. In mobile and wireless environments it is really useful as users are frequently changing their network points of access, the used applications and services vary in bandwidth demands depending on the nature of the content to be transmitted, and considering that wireless coverage areas are providing a naturally chaining environment.

Virtualization of network functions efficiently abstracts services from the physical infrastructure. Multi-tenancy permits each network slice to possess its own policy, whether that slice is managed by a mobile virtual network operator, over-

the-top service provider, virtual private enterprise network, governmental public network, traditional mobile operator or any other business entity.

There are potentially needed some extensions to OpenFlow due to requirements of the mobile network adaptation.

3.1.3 3GPP

The 3rd Generation Partnership Project (3GPP) is a collaboration agreement that was established in December 1998. The collaboration agreement brings together a number of telecommunications standards bodies which are known as 'Organizational Partners' (ARIB, CCSA, ETSI, ATIS, TTA, and TTC).

The original scope of 3GPP was to produce globally applicable Technical Specifications and Technical Reports for a 3rd Generation Mobile System based on evolved GSM core networks and the radio access technologies. The scope was subsequently amended to include the maintenance and development of the Global System for Mobile communication (GSM) Technical Specifications and Technical Reports including evolved radio access technologies (e.g. General Packet Radio Service (GPRS) and Enhanced Data rates for GSM Evolution (EDGE)).

The project covers cellular telecommunications network technologies, including radio access, the core transport network, and service capabilities - including work on codecs, security, quality of service - and thus provides complete system specifications. The specifications also provide hooks for non-radio access to the core network, and for interworking with Wi-Fi networks.

3GPP specifications and studies are contribution-driven, by member companies, in Working Groups and at the Technical Specification Group (TSG: Radio Access Networks (RAN), Service & Systems Aspects (SA), Core Network & Terminals (CT) and GSM EDGE Radio Access Networks (GERAN)) level.

The 3GPP technologies from these groups are constantly evolving through Generations of commercial cellular / mobile systems. Since the completion of the first LTE and the Evolved Packet Core specifications, 3GPP has become the focal point for mobile systems beyond 3G.

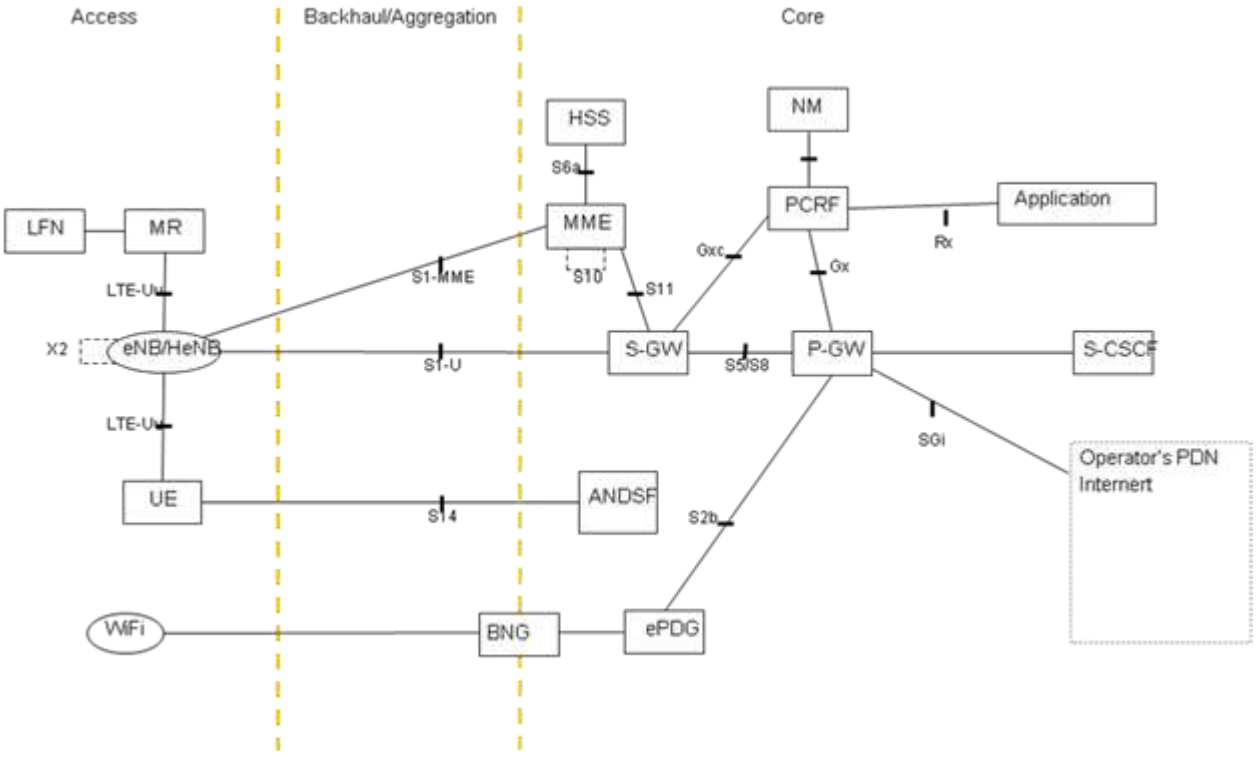

Figure 2. A simplified (LTE focused) 3GPP non-roaming architecture

Although these Generations have become an adequate descriptor for the type of network under discussion, real progress on 3GPP standards is measured by the milestones achieved in particular releases [6]. New features are 'functionality frozen' and are ready for implementation when a Release is completed. 3GPP works on a number of Releases in

parallel, starting future work well in advance of the completion of the current Release. Although this adds some complexity to the work of the groups, such a way of working ensures that progress is continuous & stable.

3.1.3.1 Relation to SIGMONA work scope

Studies on enhancements to 3GPP EPC utilizing SDN/OpenFlow and NFV principles have been proposed, and recently 3GPP agreed to start a 1-year study called "Study on Network Management of Virtualized Networks". The study focuses on work performed at ETSI NFV. Currently no work impacting the 3GPP network architecture has been planned.

A 3GPP study called "Flexible Mobile Service Steering (FMSS)" has been ongoing since January 2014. It is related to Service Chaining (in (S)Gi-LAN), and focuses on identifying use cases and requirements. It is likely that normative work on this topic will be started in the near future, and be included in 3GPP Release 13.

3.1.4 IETF

IETF (Internet Engineering Task Force) has several working groups in the areas of the relevant research to this project.

One of these is IETF SFC Working group (Network Service Chaining). Working group was started late December 2013 and is chartered with five deliverables:

- Problem Statement
- Architecture, initially single-domain
- Generic data plane encapsulation
- Control Plane requirements
- Manageability

New BoFs are proposed for the related areas [7]:

- ACTN (Abstraction and Control of Transport Networks). Status: non-WG Forming
- VNFPOOL (Virtualized Network Function Pool)
- APONF (Application-based Policy for Network Functions)

The Internet Research Task Force called Software Defined Networking Research Group (SDNRG) was chartered on 14 January, 2013. Its areas of interests are

- Classification of SDN models, including definitions, taxonomies and relationship to work ongoing in the IETF and other SDOs
- SDN model scalability and applicability
- Multi-layer programmability and feedback control systems
- System Complexity
- Network description languages, abstractions, interfaces and compilers, Including methods and mechanisms for (on-line) verification of correct operation of network/node function.
- Security

This research group also started to work on the definition of SDN layers and terminology. Recently an Internet-Draft has appeared in this topic. The proposed architecture model for SDN is mainly based on ONF SDN-architecture, but provides a different view, as illustrated in Figure 3.

The main components of the model are the following [8]:

- Network device: a device that performs one or more network operations related to packet manipulation and forwarding. Device can be physical or virtual.
- Application (App): a standalone piece of software, which performs a function by utilizing the underlying services. It does not offer any interfaces to other applications or services
- Service: a piece of software that performs one or more functions and provides one or more APIs to applications or services of the same or different layer.
- Forwarding plane: the collection of resources across all network devices responsible for forwarding traffic.
- Operational Plane: the collection of resources responsible for managing the overall operation of individual network devices.

- Control plane: the collection of functions responsible for controlling one or more network devices. The control plane interacts primarily with the forwarding plane and to a lesser extent with the operational plane, a lesser extent with the operational plane.
- Management plane: the collection of functions responsible for monitoring, configuring and maintaining one or more network devices or parts of network devices. The management plane is mostly related with the operational plane and less with the forwarding plane.
- Application Plane: the collection of applications and services, which program network behavior.

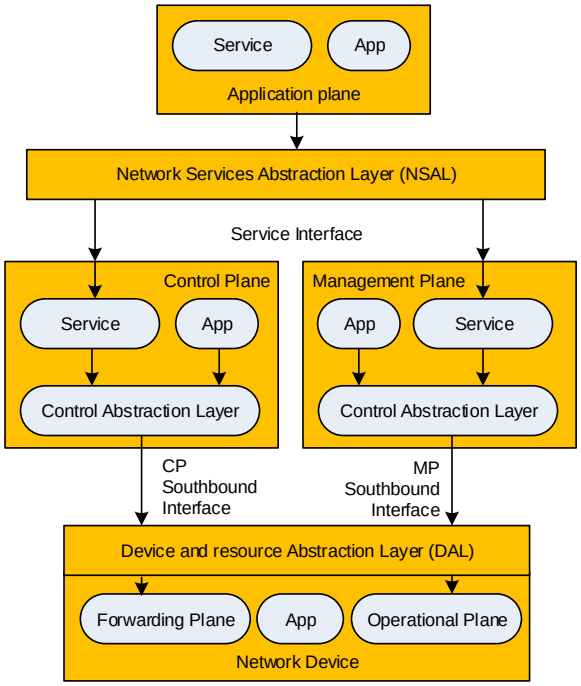

Figure 3. IRTF SDNRG view of SDN Layer Architecture.

3.1.4.1 Relation to SIGMONA work scope

The SDNRG IRTF had in the past meetings co-located with IETF meetings [9]. The topics of the presentations cover a wide range of areas, related to network, resource, traffic management, security. Therefore, the topics can serve as input for the different research topics in SIGMONA. The next meeting will be at IETF91 in November, 2014.

3.2 Overview of the used architecture models

This section discusses overall SIGMONA project Software Defined Mobile Network reference architecture and how the research topics in the project relate to it and what kind of new interfaces are introduced. There is quite a variety of different areas of the mobile and cloud network architecture in the scope of this project and there has been used the general reference models to reflect the different topics.

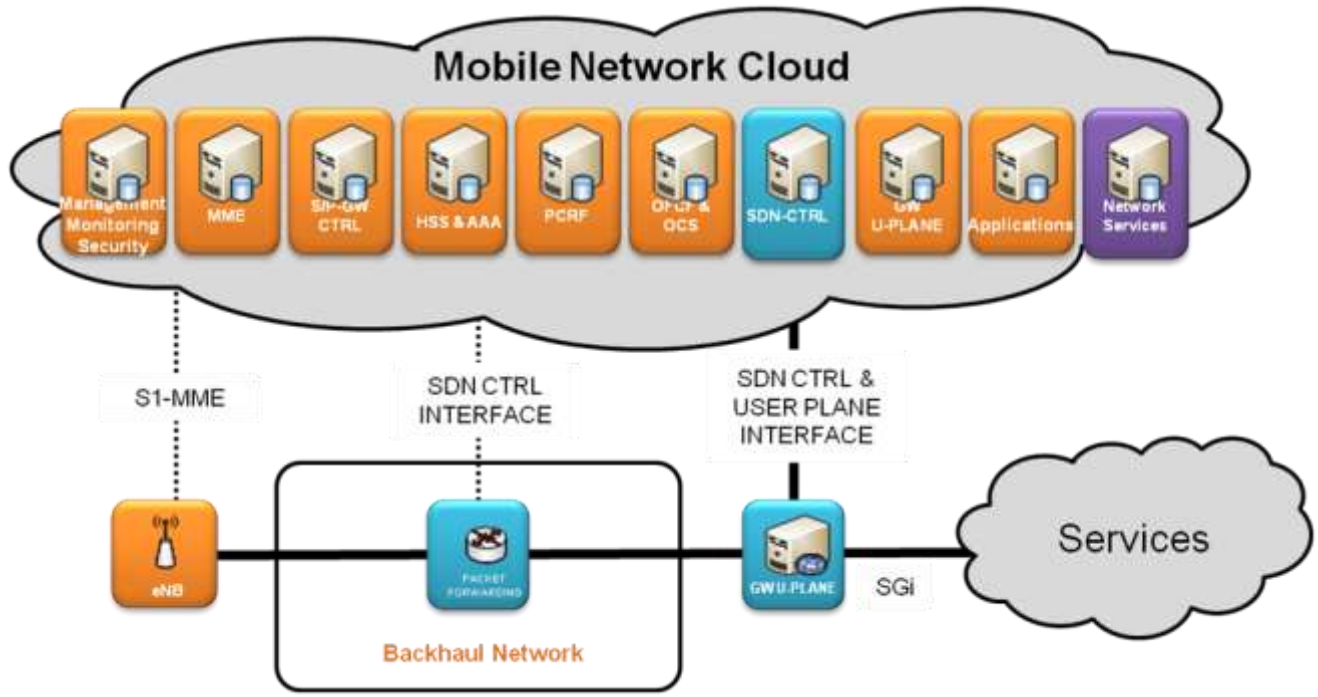

Figure 4. SIGMONA project Software Defined Mobile Network reference architecture.

In addition to this mobile network cloud view, the consolidated research topics have been analyzed in the SDN defined architectural model and the key functionalities in the different layers have been identified. The target in this has been to harmonize the functional areas and to define the topics in the same model scope.

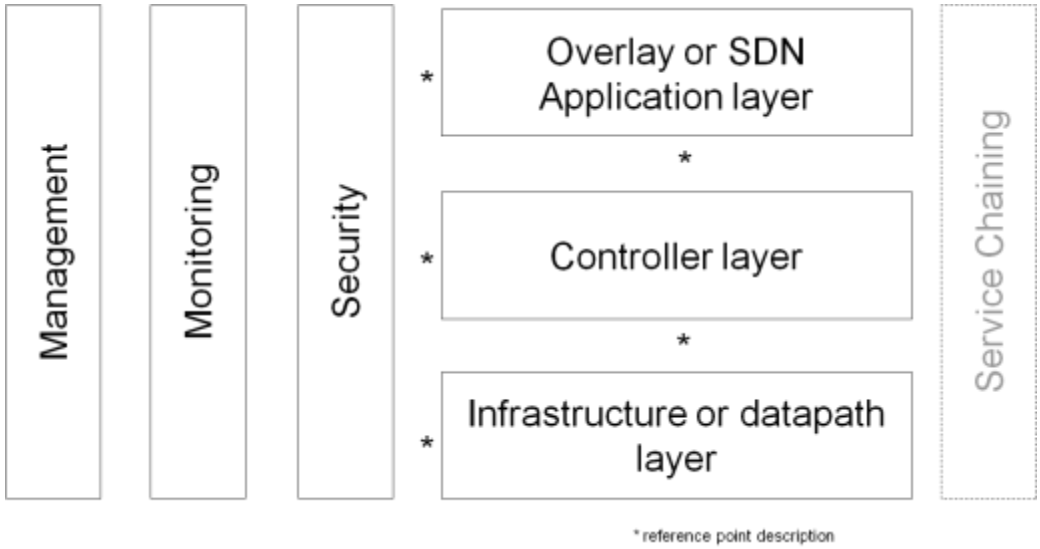

Figure 5. Layered SDN reference model, based on ONF definition [10]

The ”Controller layer” intended not to limit to only in a ”SDN Controller”, but may include other functions as seen feasible. A research topic may focus on only some parts of the model. For example, only the ”Controller layer” and the ”Datapath layer” may be relevant for a topic. Management, monitoring and security could be either a property of a single layer (like ”Controller layer”) or an e2e topic of its own spanning the whole model.

The project research topics are categorized under the following areas and further consolidated to form a common view on the different areas of the architecture and their impact to the potential evolution.

3.2.1 Mobile Core Network Cloud

Mobile Core Network Cloud is based on architecture framework defined in ETSI NFV Architectural Framework [3] section 5.2 High-Level NFV Framework. Some of the functional blocks illustrated within Mobile Core Network Cloud such as Mobility Management Entity (MME) can be seen as Virtualised Network Functions (VNFs) defined in section 7.2.2 [3].

> NOTE: "Network Services" illustrated within Mobile Core Network Cloud above may overlap with definition of "Network Services" described in ETSI NFV [3] chapter 6.
>
> NOTE: Placeholder for cloud federation (identity/security).

In the Mobile Core Network Cloud deployment logical layers Management, Monitoring, Security, Overlay/SDN Application Layer, Controller Layer,(virtualized) Infrastructure/Datapath Layer and Service Chaining may co-exist simultaneously.

> NOTE: Logical Architecture Model (chapter 4) provides a Controller centric view on network functionalities whereas Mobile Core Network Cloud aims at depicting a deployment view of a network model.

3.2.2 Management, Monitoring and Security

For Management, Monitoring and Security following solutions are defined

- Network Performance and Resource Monitoring in virtualized mobile networks (section 4.3)
- QoS and QoE Monitoring (section 4.3.1)
- Secure mobile data offloading and Traffic Management over SDMN (section 4.5.1)
- Joint Routing and Resource Management (section 4.5.1.2)
- Service Chaining in SDMN (section 4.5.1.6)
- Network security management, firewall and CES security (section 4.6.2)

3.2.3 3GPP defined functional entities

- MME, HSS & AAA, PCRF, OFCS & OCS: 3GPP defined interfaces towards non-virtualised functions outside Mobile Core Network Cloud such as S1-MME towards eNB. SDN Control interfaces within Mobile Core Network Cloud and towards entities outside Mobile Core Network Cloud may be deployed depending on the use case.
- S/P-GW Control Plane, GW User Plane: May be decoupled within the Mobile Core Network Cloud. GW User Plane may be deployed within Mobile Core Network Cloud and (non-virtualised) outside Mobile Core Network Cloud (section 4.1)
- Consolidated control plane: Once S/P-GW Control Plane and GW User Plane are decoupled the subscription contexts maintained within MME and GW Control Plane may be consolidated into an optimized entity within the Mobile Core Network Cloud (section 4.1).

3.2.4 SDN Control

Multiple SDN Control entities (per use case) may exist in the Mobile Core Network Cloud. SDN Control entity may provide a single functionality or multiple grouped functionalities together as listed below

- SDN Control for decoupling of GW Control Plane and GW User Plane (section 4.1).
- Backhaul controller: May interact with other functions in the Mobile Core Network Cloud such as 3GPP defined functional entities.
  - Automated mobile network slicing and RAN Control (section 4.2.2)
  - SDN based mobile backhaul (section 4.2.1)
- Network Performance and Resource Monitoring (section 4.3.1)
- Mobility Management
  - Wi-Fi network mobility with SDN switching (section 4.4.2)
  - Traffic Optimization, QoS and mobility management (section 4.4.1)

- Traffic Management
  - Secure mobile data offloading over SDMN (section 4.5.1)
  - Joint Routing and Resource Management (section 4.5.1.2)
  - Service Chaining in SDMN (section 4.5.1.6)
- Security Management
  - Synchronized Network Security and Traffic Management (section 4.6.2)
  - Network Performance and Resource Monitoring (section 4.3.1)
  - Network security management, firewall and CES security (section 4.6.2)

SDN Controllers are use case specific and are not required to be deployed simultaneously.

SDN Control entities either in same or different domain (Mobile Core Network Cloud) may interact indirectly via Mobile Core Network Cloud provided Application Programming Interfaces (APIs) but are not required to interact directly. In the future, optimized cloud architecture may be defined for direct SDN Controller interaction.

3.2.5 Applications

End user applications both in the scope of (current) 3GPP domain such as IMS/VoLTE and outside the scope of (current) 3GPP domain provided within Mobile Core Network Cloud. Applications are not in the scope of SIGMONA project.

3.2.6 Network Services

Network Services are part of the Mobile Core Network Cloud but fall outside the scope of 3GPP domain. Network Services may interact with other Mobile Core Network Cloud functional components such as SDN Control and/or 3GPP defined functional entities

In SIGMONA reference architecture following Network Services are defined:

- Automated mobile network slicing (section 4.2.2)
- Network Performance and Resource Monitoring (section 4.2.2)
- Wi-Fi network mobility with SDN switching (section 4.4.2)
- ALTO (Server) for Traffic Optimization, QoS and mobility management (section 4.5.1)
- Joint Routing and Resource Management (section 4.5.1)
- Performance Analysis & Verification (section4.3.1)
- Service Chaining in SDMN for NAT, IPS/IDS, DPI and Redirection (section 4.5.1.6)
- Resource and Traffic Management (section 4.5)
- Network slicing (section 4.2.2)
- Network security management, firewall and CES security (section 4.6.2)

3.2.7 eNB

3GPP defines eNB functionalities in e.g. TS 36.300. eNB interacts with MME via S1-MME control interface.

On the user plane there are options to deploy

- 3GPP defined GTP tunneling solution
- Further transport optimizations such as based on TRILL (IETF) (section 4.1).

3.2.8 Backhaul Network

Multiple SDN controlled packet forwarding elements may exist within the backhaul network (section 4.2). Packet forwarding entities are controlled by one or multiple controllers deployed either as standalone controllers within backhaul domain or in the Mobile Core Network Cloud.

3.2.9 GW User Plane

GW User Plane deployed outside Mobile Core Network Cloud provides user plane of the 3GPP defined S/P-GW functionalities for which control and user plane are decoupled. The GW User Plane is controlled by SDN controller residing in the Mobile Core Network Cloud via SDN control interface. The GW User Plane can be based on

- 3GPP defined GTP tunneling solution

- Further transport optimizations such as based on TRILL (IETF) (section 4.1).

OpenFlow is used as SDN Control interface protocol, however other protocol may be used as well.

GW User Plane may act as a forwarding entity towards packet processing taking place in GW User Plane residing in Mobile Core Network Cloud.

3.2.10 Services

Typical operator provided subscription based access to packet services such as non-roaming and roaming based voice (IMS/VoLTE) interconnection, internet and/or corporate access.

4. Partner topic mapping to architecture model

In this chapter the mapping of the multiple research contributions from different partners into the generic SDMN architecture models is presented.

4.1 EPC virtualization and evolution scenarios to Software Defined Mobile Network

4.1.1 Description of the research topic and the scenarios

The virtualization of LTE network elements enables the decoupling of data plane handling from control and allows using off-the-shelf network switches as part of the transport network while moving all the mobile specific control functionality to the cloud. The LTE virtualization allows the consolidation of several control elements into single networking module that benefits from SDN controller functionality to further separate the data from control and use SDN switches as the switches as shown in the figure below as enhanced MME (eMME).

The mobile core network needs to scale to up to hundreds of millions of users. Each user's subscription context data maintained within the core network, however, is largely independent from each other, enabling massively parallel processing of user's subscription context data. While in the current product generation the functional split between the different parts of the user's mobility context management (e.g., MME, S-GW, P-GW) has been the main basis for the split into different vendor products, each product has had to manage scaling to millions of users internally. In the cloud era it seems more feasible to use cloud-based scaling methods for such combined control element, each of which has the capability to manage all the functional needs of a single user's state management.

There are 4 different scenarios presented as potential architecture evolution:

Scenario 1. Virtualized vEPC with traditional GTP tunnel between eNB - S/PGW: Deployment of the vEPC, based on current standard network elements where each of them are running on different NVFs. Each of the different NW functions (i.e. MME, S/P-GW and FW) will be running on their own virtual machines in the cloud. The existing services and solutions can be smoothly migrated to the cloud environment and this enables optimization when the target content is located in the cloud. Also for the potentially needed service chaining functionality (operated in the cloud), the user plane tunnel termination in the cloud would be optimum solution. For the user plane data handling there might be needed some packet processing optimization methods.

Scenario 2. GTP tunnel eNB - S/PGW and SDN backhaul with L2 tag based switching: Integrate the SDN controller with the eMME. In this scenario we use SDN to add L2 tagging to the GTP packets so we can perform traffic engineering in the backbone. UE data packets are switched from the eNB to the S/P-GW across the core network and several paths. Load balancing between OFS#1 and OFS#2 links is possible based on the VLAN/MPLS identifiers. QoS can be provided using the L2 tag QoS bits.

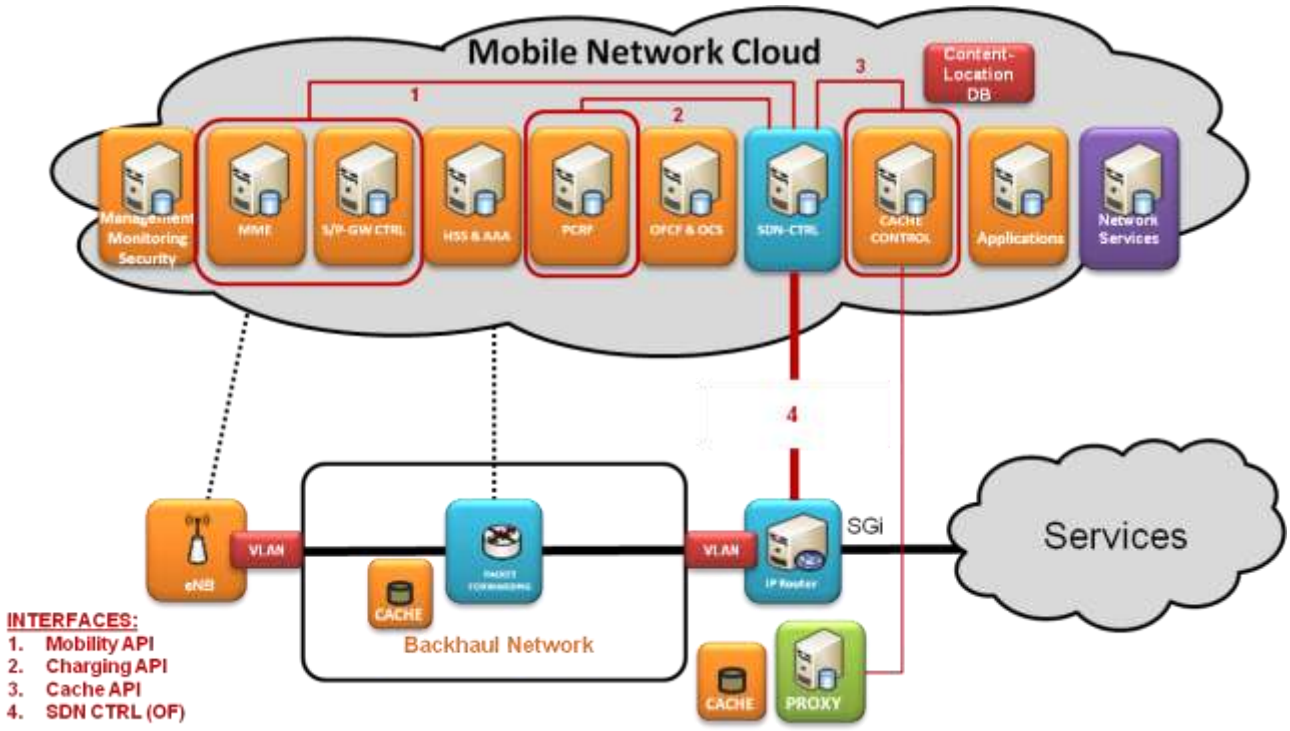

Figure 6. eMME integrated with SDN controller and cache control (scenarios 2 and 4)

Scenario 3. Consolidation of the core control plane functions. This is in part enabled by the further separation of the mobile network GW control and user planes from each other with the SDN principles [11]. The defined north-bound OpenFlow controller API for session flow control enables the further consolidation of the control plane functions of the mobile core network.

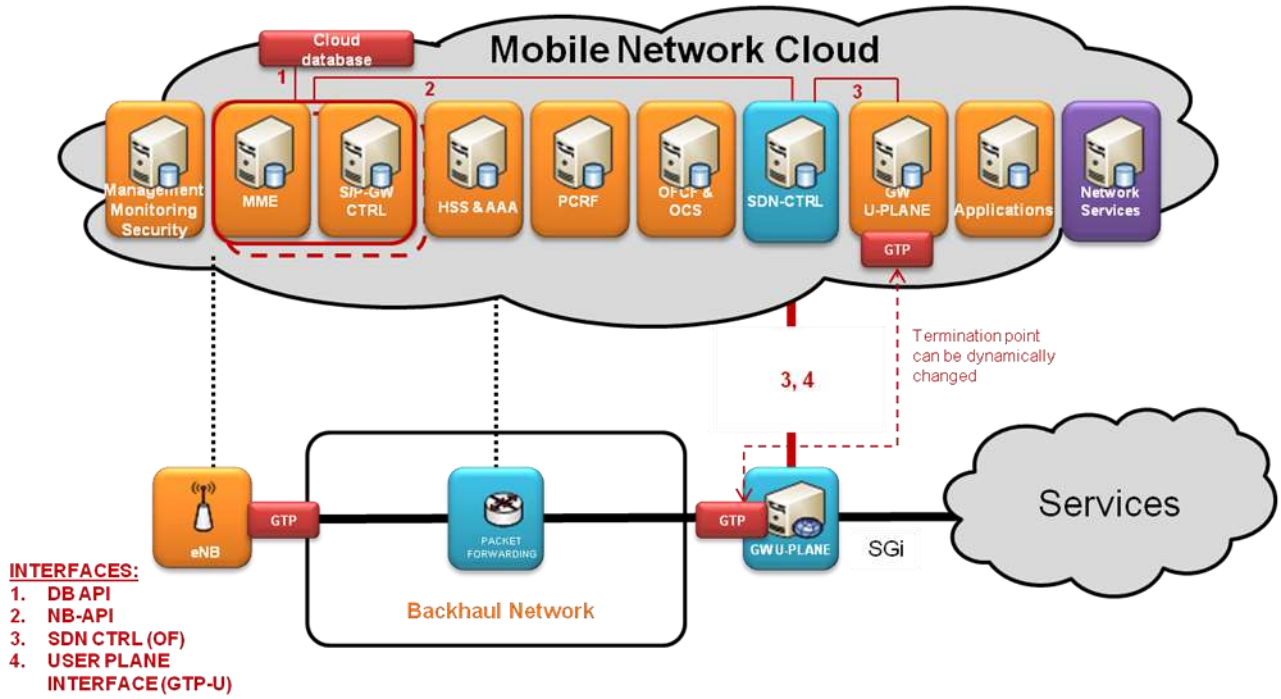

Figure 7. Consolidate Control Plane combined with OpenFlow controlled GW user plane (scenario 3)

This work concentrates in defining and demonstrating the consolidated control plane shown in the figure above. The UE specific context data is maintained and signaled in current architecture within different EPC control plane elements. The

consolidated control plane entity combines all control-functionality relating to a user's context and reachability control data, inheriting parts of functionality from the MME, S-GW, and P-GW. The remaining functions will be located into OpenFlow controlled user plane gateway. It is envisioned that the consolidated control plane consists of multiple finite state machines as transaction handlers that can receive and transmit standards-compliant control signaling, but that will communicate using internal APIs within the entity itself. This may enable signaling latency optimizations that in part help offset the additional latencies due to cloud-based data stores and management systems. The resource utilization can be improved by moving the user context statefulness from the processing elements (during the stabile state) to the cloud context database.

Other targeted benefits are related to improved cloud environment optimized scalability potential simplification of the signaling transactions, improved monitoring and serviceability impacts due to the new architecture.

Scenario 4. SDN controlled backhaul providing the mobility services without the GTP tunneling. Here the usage of SDN replaces completely the data plane part of standard NW elements such as S/P-GW. The usage of SDN has some uncertainties in terms of mobility management and support for mobile specific tunneling technologies such as the GPRS Tunneling Protocol (GTP).

The main research topics addressed would be i) compare distribute solutions such as TRILL [12] to manage mobility versus SDN controlled mobility, ii) identify the network elements that benefit from relocation and scaling out based on traffic demand and study using multiple data centers for moving or initiating new instances of virtualized network elements and iii) Improving network throughput by reducing overhead after removing mobile specific tunneling (i.e. GTP) and use lower layer tunneling such as VLAN or MPLS. Furthermore, in order to support the use of cloud, the overall cloud architecture (number, size and location) of computing infrastructure must be studied.

Main use cases:

1. Separation of control and user plane of the EPC GW
2. EPC control plane functionality consolidation
3. Integration of SDN in 3GPP architecture
4. Impact of mobility in SDN based mobile networks

Basic Assumptions:

- Assumption on Migration: Compatibility with the legacy systems vs. Clean-slate deployment,
- Assumption on mobility management: SDN based mobility management versus 3GPP and MIP mobility,
- Assumption on security in the edge: Cooperation between edge nodes and operators towards global trust.

Main targeted benefits:

Consolidate LTE virtual network elements into single component (eMME) that integrates SDN functionality. Reduce tunneling overhead in data plane and increase throughput. Use off-the-shelf switching network components without mobile functionality. Dynamic allocation of resources in the cloud across data centers according to the traffic demands.

4.1.2 Mapping to layered SDN model and reference points

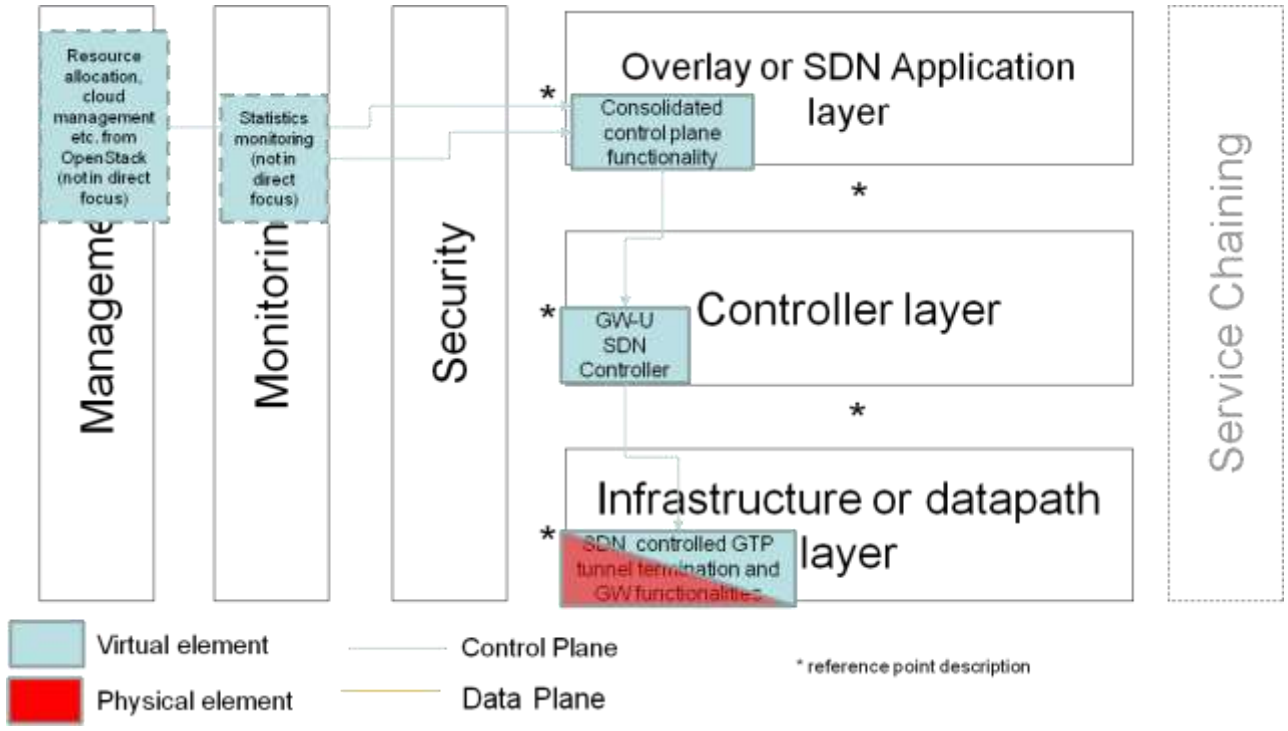

Figure 8. Virtualized EPC with SDN controlled scenarios

- SDN Application layer
  - **Network slicing manager** (scenarios 2&4)**:** The application layer will interact with MME and SDN controller to instantiate multiple slices of the controller layer for different operators. The PCRF will also be implemented as application that interacts with SDN controller to enforce certain policies and collect charging information from the controller. This application will manage the instances of the virtualized LTE network elements required for each VMNO to manage their slice of the network. This application will also act as load balancer and cloud mobility manager to initiate new instances of LTE network elements and move them in the cloud across different data centers to scale up or down depending on traffic demands.
  - **Consolidated control plane** (scenario 3)**:** Currently the same UE context data is handled and stored within MME, S-GW and P-GW, and keeping the info consistent causes some transactions/signaling messages
    - In the proposal all state transactions related to a set of UEs would be processed in the same VM
    - User Control Entity consolidates all the control functionality relating to a user's context and reachability control data
    - Inheriting control parts of functionality from the MME, S-GW, P-GW, and potentially PCRF elements
- Controller layer
  - **MME integrated with SDN controller** (scenarios 2&4)**:** The MME maintains the current 3GPP standards and continues supporting S1-MME interfaces with eNodeB and S11 interface with S/P-GW, S2 interface with HSS. The 3GPP standards allow interoperability with legacy 3GPP devices such as eNodeB, S/P-GW, HSS and PCRF. In addition the MME integrates the SDN controller functionality that allows the MME to interact with SDN enabled data plane
    - Use case: Integration of SDN with 3GPP architecture
    - Basic Assumption: Compatibility with legacy systems
  - **GW-U SDN Controller** (scenario 3)**:** The 3GPP specific context data (GW "instructions") handling via specific NBI from the GW-C function and converts them to OF based flow commands (not in direct focus of this research)
- Datapath layer

- **Non-GTP data plane** (scenarios 2&4)**:** Standard off-the-shelf SDN switches will be used without GTP for data plane. The mobility management would be done without GTP, but using SDN functionality. The SDN switches can be installed in existing mobile network together with non-SDN switches where data plane can be with and without GTP.
- **SDN controlled GTP tunnel termination and GW functionalities** (scenario 3)**:** GTP tunnel termination and closely related functions on the physical fast path forwarding element or in the service cloud (not in direct focus of this research)

4.1.3 Main research questions for this area

1. What is the scalability improvement potential of the EPC control plane functions in the cloud environment?
2. Legacy network interoperability and roaming scenarios with this kind of architecture evolution?
3. Can SDN based transport replace GTP and fulfill mobility and QoS requirements
4. Is inter-cloud mobility feasible and fulfill the telco requirements
5. Usage of SDN technology to replace legacy GTP tunneling
6. Mapping of GTP functionality into Ethernet or MPLS tunneling
7. Usage of SDN technology as data plane of FW/CES
8. Feasibility of mobility across data centers of different virtualized LTE network elements
9. Cloud infrastructure architecture
10. What kind of simplification potential there is in the signaling transactions?

4.2 Network capability management and slicing

4.2.1 SDN based mobile backhaul

4.2.1.1 Description of the research topic and the scenarios

The main research topics are SDN based mobile backhaul solutions and virtualized (NFV) backhaul network elements. SDN based mobile backhaul systems decouple control plane from network elements to a (at least logically) centralized SDN controller. The main target benefits include network programmability and simplification of network operations. Network programmability is expected to enable new networking models, for example traffic forwarding from base stations to virtualized mobile core network elements that migrate from a location to another depending on load and other QoS situations. Simplification is expected to be achieved by two means; SDN enabled integration of backhaul network operations to overall automated mobile network orchestration software stack and elimination of complex legacy protocols.

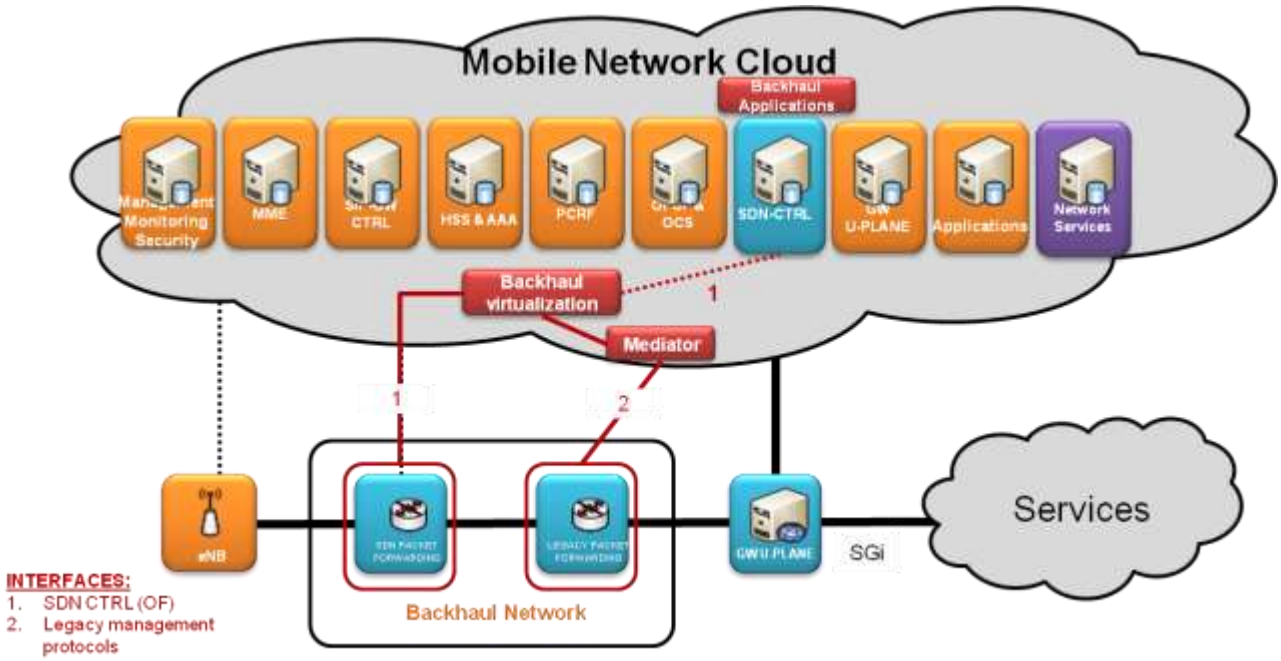

Figure 9. Mobile network SDN controlled backhaul

In SDMN, a transport network operator (TNO) is assumed to set up and maintain a mobile network backhaul that is shared by multiple mobile network operators (MNOs). By using network virtualization techniques, each MNO should be able to independently manage its own slice of the network. This allows reducing capital expenses and operational costs, through increased efficiency in network resource usage and management.

When a MNO incrementally attaches its own equipment (e.g. eNBs) into the TNO's network, the transport network service provisioning should be secure and automated, constrained by policies and agreements between the TNO and the MNO. MNOs participating in the shared network then maintain their own control plane (e.g., MME, PCRF, and HSS) and this is used to control the network slice they are allotted by a controller/virtualization hierarchy that is managed by the TNO. However, it may not be clear which kind of network topology the TNO is willing to expose to the MNOs, and what kind of slices or network abstractions a MNO is willing to see and control by itself (compared to the current practice where the network just provides "opaque" transport service between Customer Edge equipment, based on SLAs). There are also issues of how to establish and maintain the MNOs' dynamically changing network views, and how to segregate the MNO-specific resource allocations.

One of the basic assumptions in SIGMONA is that SDN technology will be incrementally deployed, and in the migration, backward compatibility with legacy networks must be maintained. Some of the current access network technologies (e.g., wireless mesh and MPLS), that are used for mobile backhauling, apply autonomous routing, load balancing, fault recovery and other similar functions that react rapidly to local changes. Moving these functions towards a centralized SDN controller could be counterproductive in terms of latency, reliability, and the amount of control traffic. This needs to be taken into account when constructing SDN-conformant hardware abstractions of legacy network partitions. One approach is to hide the internals of the legacy segments by presenting them to the SDN layer as SDN capable (virtual) switches. Then the SDMN can utilize all existing transport access technologies without need to hamper the (re)active local functions. Such aggregating network abstraction requires coordinated resource management by using a "mediator" function that translates between the SDN and legacy operation models.

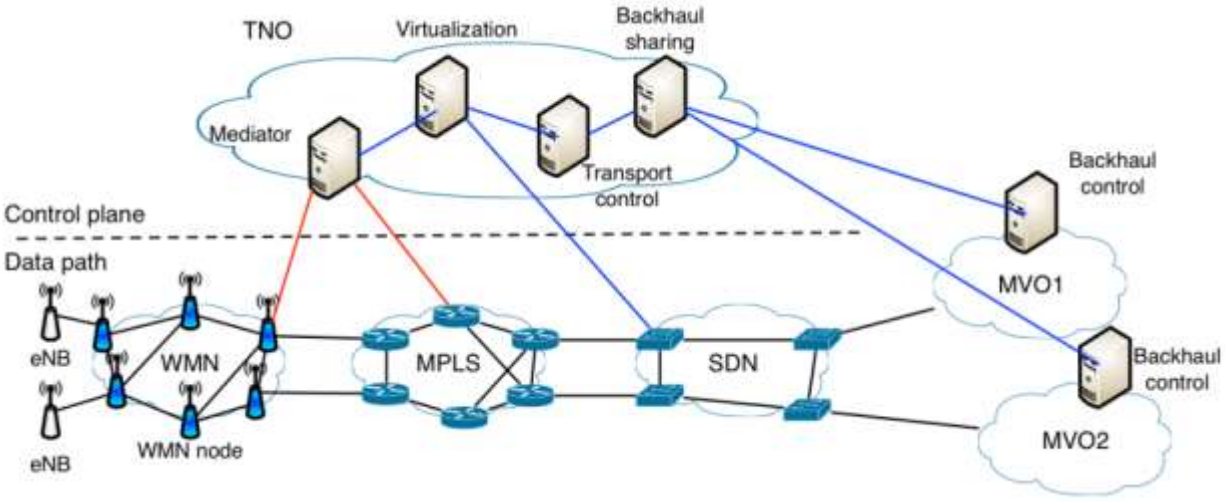

Figure 10. SDN-controlled mobile backhaul

NFV based network elements are different approach to backhaul network evolution. Expectation is that NFV will not turn out to be an optimal solution for simple packet transport switching tasks (such as MPLS label switching within a network), but NFV may have benefits at the edge of network. For example, mobile core site IP/MPLS PE router could be virtualized along with other core site functions (S/P-GW, MME, etc.). Or to take another use case, if base station site hosts networking functions such as caching in general-purpose hardware, then also cell site router could be potentially virtualized according to NFV model.

The main target benefit in NFV approach is the possibility to utilize general-purpose hardware also for backhaul tasks.

Related Use Cases:

- Provide reliability and resilience to mobile networks with SDN,
- SDN and Virtualization control architecture.

Basic Assumptions:

- Assumption on Migration: Compatibility with the legacy systems vs. Clean-slate deployment,
- Assumption on resilience: Fault tolerant SDN based mobile networks,
- Assumption on Quality of Service: QoS provision in virtualized mobile core networks in SDN-based network forwarding paths,
- Assumption on service provisioning and optimization,
- Global optimization; Control application has means carry out global optimization vs. Global optimization problem is not economically feasible

Main targeted benefits: Automation of the network service provisioning.

4.2.1.2 Mapping to layered SDN model and reference points

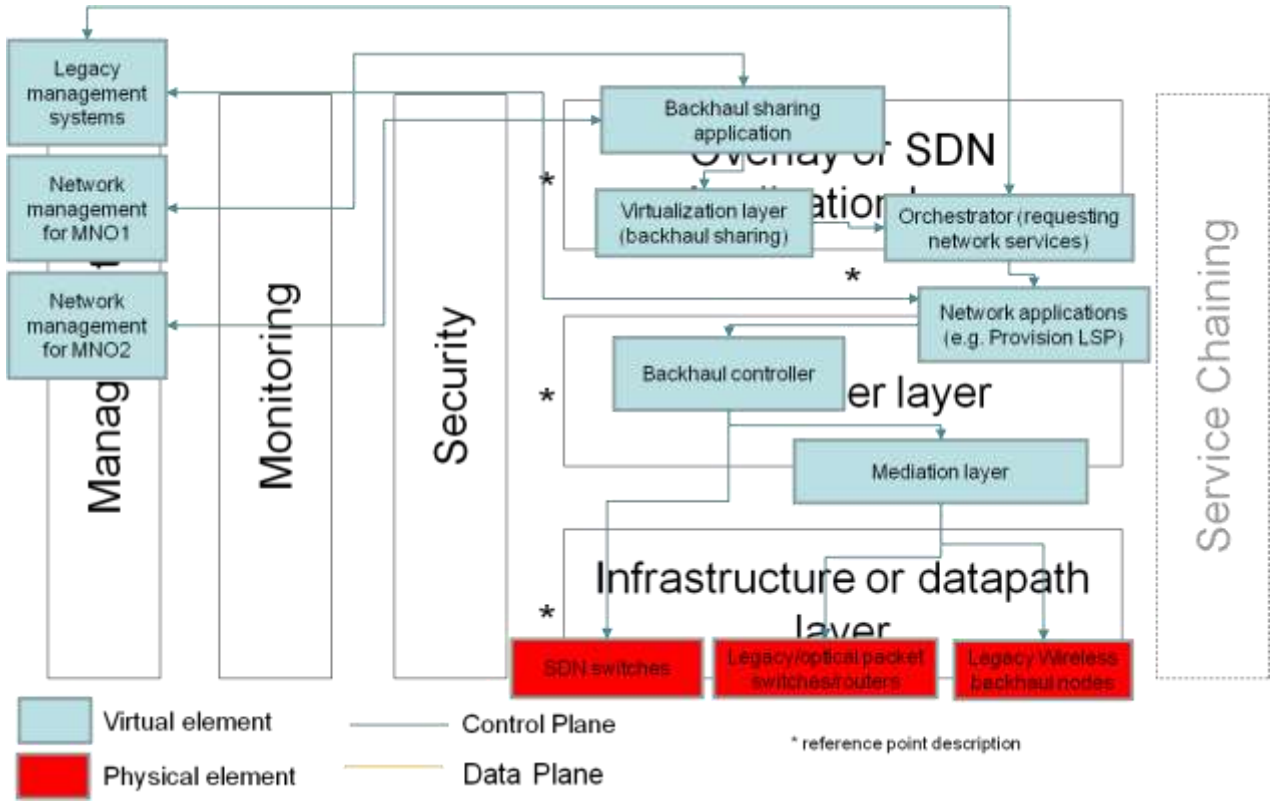

Figure 11. Backhaul e2e optimization and slicing

- SDN Application/orchestration layer
  - **Service orchestration:** This functionality is responsible for orchestrating overall SDN mobile network (comprising of radio network, backhaul network and core networks) functionality.
  - **Resource management:** This functionality is responsible for network virtualization, allocating resources to different users (e.g. operators) of the network.
  - **Backhaul sharing application:** Will provide interface for multiple MNOs for requesting and controlling backhaul connections.
  - **Virtualization layer:** Will provide secure isolated network slices that are used to create backhaul networks for MNOs; furthermore, it will orchestrate transport resource sharing
- Controller layer
  - **Backhaul controller core:** Common platform layer for mobile backhaul controller. Analogous to operating systems in servers etc. Will manage network resource slicing and sharing at NE level. Controls legacy networks via Mediation layer and true SDN switches directly.
  - **Network applications:** Responsible for controlling the specific network functions (e.g. IP VPN creation, LSP creation, …) in mobile backhaul network. Run on top of controller core.
  - Mediation layer will act as an interworking function between legacy networks and SDN; it enables SDN based control of legacy NEs; furthermore, it may hide legacy network structure and functionality by presenting legacy subnetworks as single SDN switch or simple network with virtual switches
- Datapath layer
  - The Datapath layer consists of legacy networks and NEs and SDN-ready routers and switches, including packet and optical devices.
- Management
  - MNO network management is responsible of requesting and managing backhaul connections.
  - Responsible for management tasks that are not performed by SDN controllers, e.g. Software downloads, fault management, etc.
- Monitoring

- Responsible for collecting data from network equipment, analyzing data and providing feedback to SDN controllers and orchestrators for performance optimization.

- Security

4.2.1.3 Main research questions

1. How SDN mobile backhaul network works when mobile core functions are virtualized and distributed to multiple datacenters?

2. How NFV technologies could be used in mobile backhaul? Would there be benefits in virtualizing also backhaul elements (e.g. gateway site routers) along with other elements (e.g. S-GWs)?

3. How existing mobile backhaul networks (e.g. IP/MPLS) can transition to SDN and how to construct SDN hardware abstractions of legacy network elements (WMN, MPLS, etc.)?

4. How the implementation of SDN protocols (such as OpenFlow) affects software architecture in network elements? How to co-operate?

5. How SDN technologies can be used to virtualize mobile backhaul network, e.g. for mobile network sharing and fixed - mobile convergence situations?

6. How to support plug-and-play network extension in, e.g., small-cell scenarios?

7. How to build a controller/virtualization hierarchy for multi-operator environments?

8. How to assess the dependability of an SDN-based backhaul?

4.2.2 Network slicing and resource allocation

The embedding of virtualized mobile networks (slices) within a cloud-based SDMN infrastructure is an extension of the resource management problem. Network and cloud resources are provided to create virtual mobile network slices according to the SLAs negotiated with the virtual network operators (VNOs). The setup of network slices and their dynamic adaptation versus the current traffic demand is performed by an own control function (the “slicing controller”) which interacts with the SDN controller and the overall resource management.

Algorithms for the optimal setup and dynamic adaptation of the virtual networks are designed. The optimization target is the minimization of the resource consumption (cloud resources and transport network resources) of the virtual network slices considering the traffic demands and the QoS requirements stated by the VNOs in their SLAs.

4.2.2.1 Description of the automated mobile network slicing

The main scope of this work is to define an advanced network management framework/architecture for virtualized mobile networks. Its focus will be on automated management of virtual mobile network slices that includes the provisioning, maintenance and continuous optimization of the slices. The work will not be limited to one technology, e.g., LTE/EPC but will cover heterogeneous radio access systems consisting of coexisting technologies. The research work will take a holistic view, i.e. radio, transport, core virtualization is considered. The framework will be designed to be able to operate also in partly-virtualized mobile networks (e.g., only part of the mobile backhaul is SDN based).

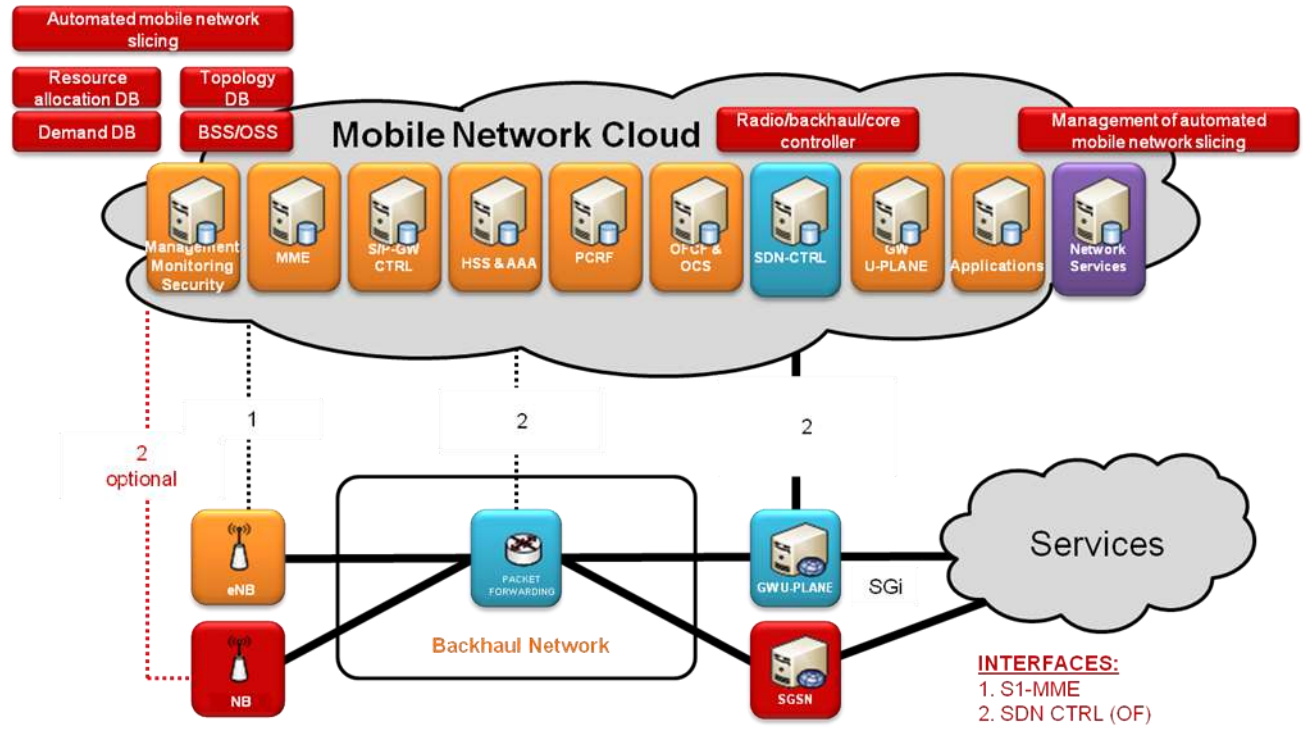

Figure 12. Automated mobile network slicing

In order to have a consistent network operation, the framework has to cooperate with other management functionalities such as service, traffic and mobility management and should harmonize their actions. The research shall analyze the possible cooperation and harmonization scenarios among the management functionalities. The framework is to consider cost and revenue aspects in its operation, i.e., it evaluates the received optimization triggers and planned optimization actions based on the expected revenue and prioritize them, or even skip ones with low expected revenue. This needs the interfacing of the BI/OSS tools and the framework; the research will analyze the possible ways to use them in the revenue evaluation.

The **key use-cases** for the automated mobile network slicing investigated in the research work are:

- Application/service defined mobile network
- RAT coexistence over shared MBH
- New methods / algorithms for the optimum setup and dynamic adaptation of virtual mobile core networks; implementation of the new methods / algorithms
- Evaluation for realistic network and traffic scenarios; performance evaluation of the new methods via simulation and lab tests

4.2.2.2 Mapping to layered SDN model and reference points

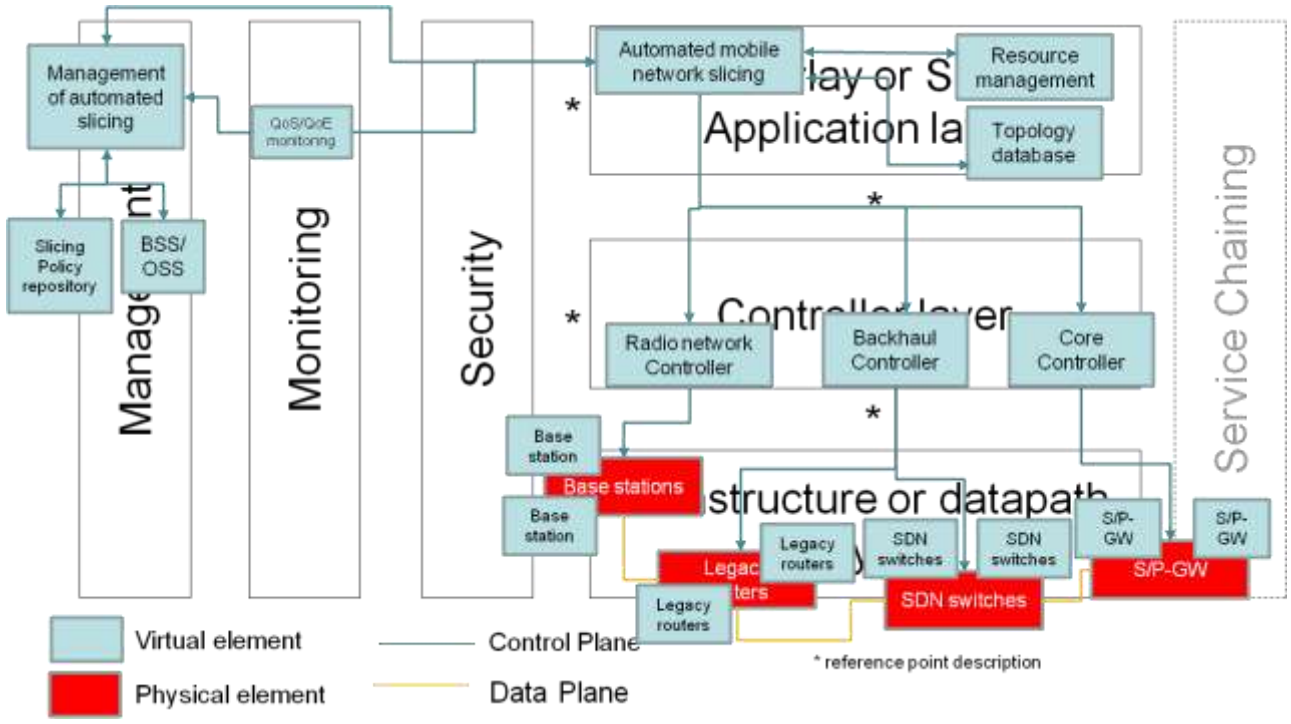

Figure 13. Automated mobile network slicing (AMNS)

- SDN Application layer
  - **Automated mobile network slicing (AMNS):** this functionality is responsible for creating/maintaining/tear down of the mobile network slices based on the instructions of the "Management of mobile network slices functionality"
    - UC: Automated mobile network slicing in SDMN
    - BA: Assumption on service provisioning and optimization
  - **Resource management:** provides services for making resource reservations and optimization based on the requested slice maintenance requests originated from the AMNS.
  - **Topology database:** provides services to the AMNS about the network topology.
- Controller layer
  - **Radio network controller**: responsible for controlling radio resources
  - **Backhaul controller:** responsible for controlling backhaul resources
  - **Core controller:** responsible for controlling the core resources
- Datapath layer
  - The datapath includes legacy and SDN-ready routers and switches. QoS shall be manageable in the datapath.
    - BA: Compatibility with legacy systems; Assumption on Quality of Service
- Management
  - Management of automated slicing: based on the received demands from the network operator, or from other operators it is managing the network slices of the mobile network
    - UC: Automated mobile network slicing in SDMN
    - BA: Assumption on service provisioning and optimization
- Monitoring
  - QOS/QoE monitoring: provides information on the service quality of the mobile network slices

4.2.2.3 Main research questions

1. How to align different mobile network management functions (e.g. mobility, resource, etc.) with the automated mobile network slicing?
2. What are the time-constraints for the automated mobile slice management operation?

4.2.3 SDN extension to mobile backhaul and RAN design

The overview of the proposed software defined LTE architecture is shown in Figure 14, which identifies the key architectural components as well as additional elements for SDN extension to mobile backhaul and RAN design. This architecture contains three main levels: Application level, RAN management/SDN controller level and Infrastructure layer.

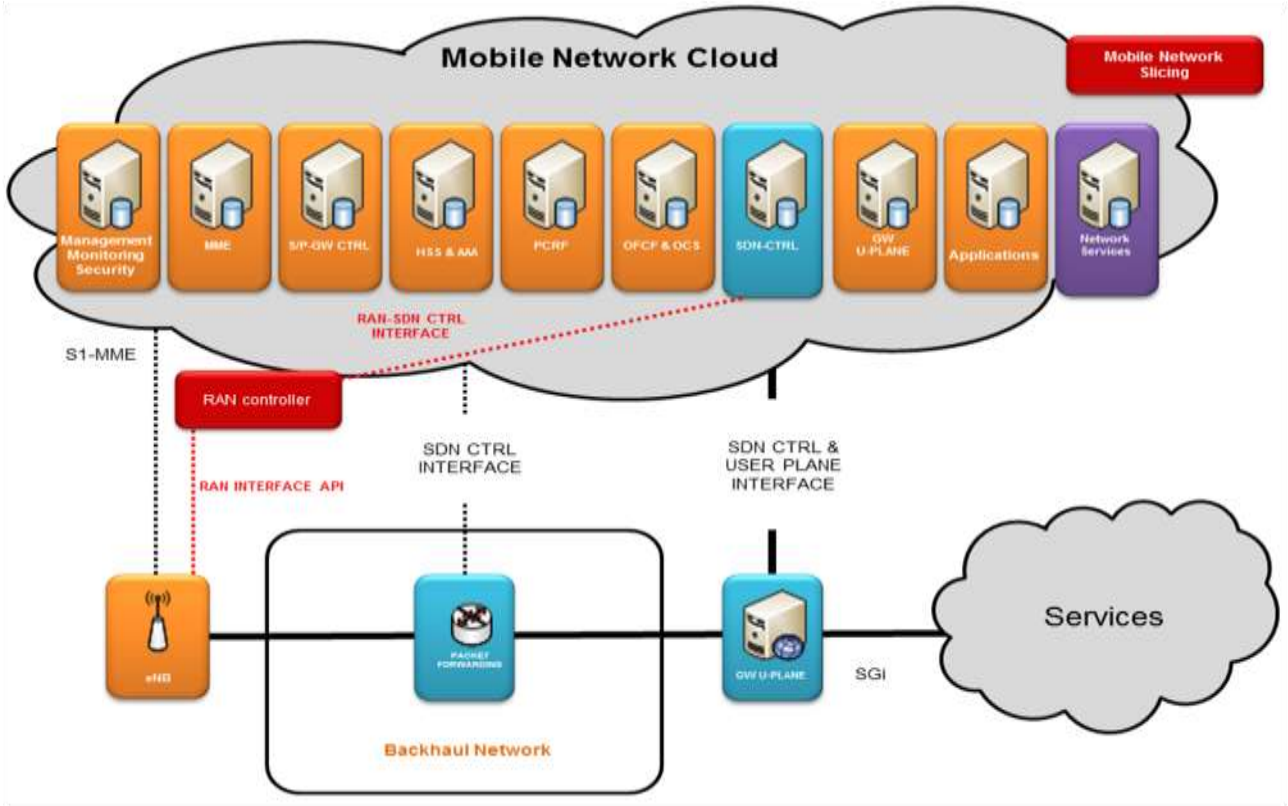

Figure 14. RAN controller enabling enhancements to RAN resource optimization

The application level includes network virtualization/mobile network slicing and running EPC elements such as MME and PCEF/PCRF as NFV applications. Their functionalities are also virtualized in the application layer, such as controlling the handoff decision of relevant eNBs, as well as policy charging execution based on subscriber and policy information of PCEF and PCRF. Moreover, the routing functionality is implemented as an application that also runs on top of the SDN controller.

The controller layer contains two main components: 1) RAN controller, which controls the monitoring and management of the radio access network and 2) SDN-CTRL, which is a central controller and controls the RAN controllers, backhaul and EPC components of LTE core network as shown in figure above.

The local RAN controller periodically collects data from RAN elements about the radio CQI (Channel Quality Indicator), congestion, traffic, interference, etc., at the time scales around milliseconds. The RAN controller is a separate network element, where RAN control plane functionality is centralized. This controller is able to communicate with the core network through northbound API to control the operation of the wireless interface. It will be able perform different tasks related to RAN related enforcements based on the optimization problem (ICIC, CoMP, scheduling, QoS, etc.). One RAN controller would serve several of eNBs (on the order of tens) and one centralized SDN-CTRL would

manage many RAN controllers with different application functionalities as an applications deployed on top of SDN-CTRL.

For the Northbound API named as RAN-SDN CTRL INTERFACE in Figure 14, RAN controller will be communicating with the SDN-CTRL through bidirectional REST API. For the southbound API, named as RAN INTERFACE API in Figure 14, RAN controller will be communicating with network devices with "special" enabled functionalities using a new OpenFlow-like protocol (created as separate plug-in in OpenDaylight framework).

The SDN-CTRL creates and dynamically optimizes the virtual elements in application layers by efficiently and fairly virtualizing and allocating resources according to the requirements. It also has the capability of establishing or modifying the rules in each virtual access element in the application layer, such as routing, virtualization and setting flow priorities. Network virtualization and mobile network slicing is managed as an application running on the SDN controller.

In the infrastructure layer, RAN controller controls the radio elements, the physical switches comprised of baseband processing remote radio heads. RAN controller's main tasks include radio resource management, scheduling, handover implementation, interference management, etc. It also pushes relevant rules to eNBs using the OpenFlow-like protocol. Such a layered architectural framework for controller layer can enable effective interference management, scheduling, bandwidth allocation and QoS management as well as provide reliable QoS guarantees for the RAN part of a mobile cellular infrastructure provider.

In the Infrastructure layer, the data path is divided into several components each with different functionalities such as SDN forwarders, radio remote heads/baseband processing units. This architecture allows for traditional eNB deployments as well as Cloud-RAN (C-RAN) implementation where the baseband processing and RF radio heads are separated and the baseband processing is moved to the cloud.

4.2.3.1 Mapping to layered SDN model and reference points

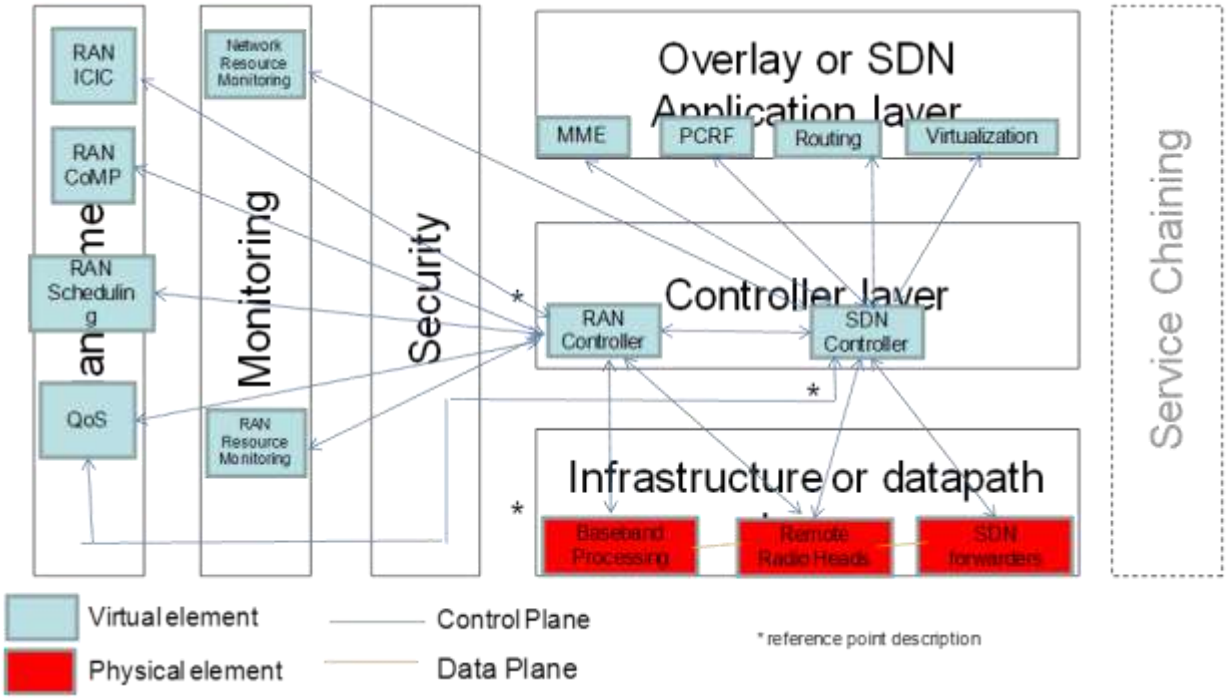

Figure 15. SDN extension to mobile backhaul and RAN design

- SDN Application layer
  - MME and PCRF/PCEF remains as functions controlling via the SDN functions
  - ROUTING: The routing on the core network composed of SDN forwarders runs on the SDN controller.
  - VIRTUALIZATION: Network virtualization and NFV is managed as an application running on the SDN controller.

- Controller layer
  - **SDN CONTROLLER**: The SDN controller manages the routing of the LTE flows. As such, a part of the SDN controller functionality resides on individual eNBs so that last mile forwarding via X2 is possible during handoff. All other routing decisions are handled by the main controller.
  - **RAN CONTROLLER**: The RAN controller handles the necessary coordination between multiple base stations for inter-cell interference cancellation (ICIC), coordinated multipoint (CoMP) transmission, joint scheduling etc. as defined by the 3GPP standard. This is handled via an SDN framework where the control plane (decision of resource allocation) is separated from the data plane (user flows) and a RAN controller pushes relevant rule to eNBs using a new OpenFlow-like protocol.
  - **RAN INTERFACE API**: This interface runs in parallel with S1-MME interface. RAN controller will be communicating with network devices with "special" enabled functionalities using this RAN INTERFACE API which is a new OpenFlow-like protocol. This southbound interface protocol will handle measurement and commands between OpenDaylight based RAN controller and the network elements (i.e. eNodeB, switch or router), will be used to monitor the resources in the RAN network elements as well as perform radio optimization handling (e.g. handover) applied from the RAN controller.
  - **RAN-SDN CTRL INTERFACE**: This interface ensures communication between RAN controller and the SDN-CTRL through a northbound REST API of RAN controller.
- Datapath layer
  - SDN FORWARDERS: The data plane of the network is composed of SDN forwarders. There is no need for specialized forwarding hardware in this case.
  - BASEBAND PROCESSING / REMOTE RADIO HEADS: The architecture allows for traditional eNB deployments as well as Cloud-RAN (C-RAN) implementation. In C-RAN, the baseband processing and RF radio heads are separated and the baseband processing is moved to the cloud.
- Management
  - RAN ICIC / CoMP / SCHEDULING: Resources of individual eNBs (and sometimes a group of eNBs) are allocated amongst users to maximize user throughput and QoS while minimizing interference of a flow to other flows nearby by the RAN Controller.
  - QoS: The QoS is managed jointly by the RAN Controller and the SDN Controller so that an end-to-end QoS maintenance is possible.
- Monitoring
  - RAN RESOURCE MONITORING: Channel quality feedback from each active user is received by an eNB every 1 ms. This feedback is then forwarded to the RAN Controller.
  - NETWORK RESOUCE MONITORING: The SDN Controller queries the SDN Forwarders regularly to assess the traffic load of the individual links in the core network.

Main targeted benefit(s): Target is to improve the dynamic optimization of the RAN parameters and the resource usage both in RAN and Core Networks.

4.2.3.2 Main research questions

1. How to develop new OpenFlow-like protocols that push relevant rules to eNBs as well as develop RAN Controller Northbound APIs?
2. How to develop jointly optimized SDN-CTRL and RAN controller applications with different functionalities?

4.3 Network resource monitoring and modeling

4.3.1 Performance and resource monitoring in virtualized mobile networks

Network monitoring is required for the verification and validation of SLAs, managing performance (QoS) and user experience (QoE), troubleshooting, assessment of optimizations and use of resources. The introduction of network virtualization in mobile networks sets new requirements for network monitoring and therefore creates needs for advanced network monitoring solutions. The monitoring solution shall be suitable for large-scale networks.

On the other hand the monitoring can also profit from the flexibility obtained from NFV. The cloud-based SDMN infrastructure provides high degrees of freedom regarding the placement of measurement points and the flexible control of traffic flows. The QoS monitoring solution comprises both a distributed (SDN / NFV-based) QoS measurement system and a centralized evaluation system.

There are different functional aspects to be studied:

Information extraction: Understanding how to deal with virtualization to obtain information on traffic flows, profiles, and properties by means of extracted protocol metadata and measurements.

Scalability and performance issues: The design of the monitoring architecture and the location of the observation points need to be done in such a way as to assure scalability, and different monitoring use cases need to be studied to obtain the best balance between performance, cost and completeness of the results. Furthermore, hardware acceleration and packet preprocessing technologies need to be integrated and controlled by applications and functions to obtain highly optimized solutions.

Analysis of different control and user plane traffic flows over the network domains and new interfaces between SDMN and existing networks and identification of related flows in different network domains.

Dynamicity: Changes in virtualized networks and applications become more easy and frequent. Monitoring solutions need to be able to adapt to these changes.

Performance, resource and security monitoring viewed as complementary: Monitoring can provide the knowledge necessary to assure the network's QoS and security. To be able to detect certain types of security issues, performance analysis is necessary. On the other hand, security issues will have impact on the performance.

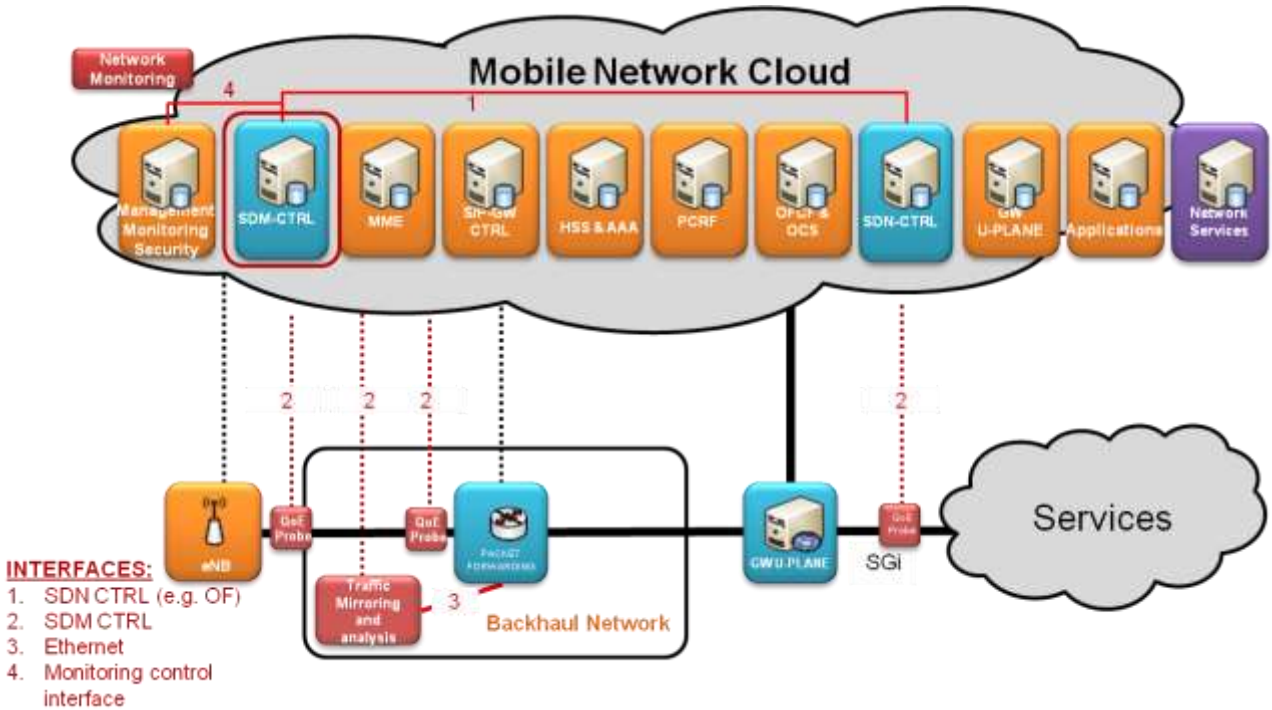

Figure 16. Performance and resource monitoring in virtualized mobile networks

4.3.1.1 QoE Monitoring

Target is on the design and implementation of a distributed (SDN / NFV-based) QoE monitoring system without client participation tailored to the cloud-based SDMN infrastructure. QoE enforcement is performed by traffic management mechanisms. Several variants of the QoE monitoring system with different degrees of centralization and function

virtualization will be developed and implemented. A performance analysis is carried out to compare the different variants. For validation a field trial is planned. The best solution will be adopted in a demonstrator setup. Furthermore, QoS to QoE mapping algorithms for selected Internet applications (e.g. video streaming, voice over IP services) are developed and implemented.

4.3.1.2 Traffic Modeling

Traffic models describe the temporal and spatial dynamics of the traffic demand. Whilst some models for the time of day mobile Internet usage behavior can be found in the literature, more research regarding the modeling of non-user generated traffic shifts has to be performed.

In this work the publicly available traffic profiles of selected Internet applications, as well as of cross traffic between datacenters, are analyzed. From that, appropriate traffic models which describe the dynamics of these traffic patterns are derived. The models are used as input for the design and evaluation of the resource and traffic management methods.

Main targeted results are the models which describe the traffic dynamic of selected Internet or CDN-based services.

4.3.1.3 Mapping to layered SDN model and reference points

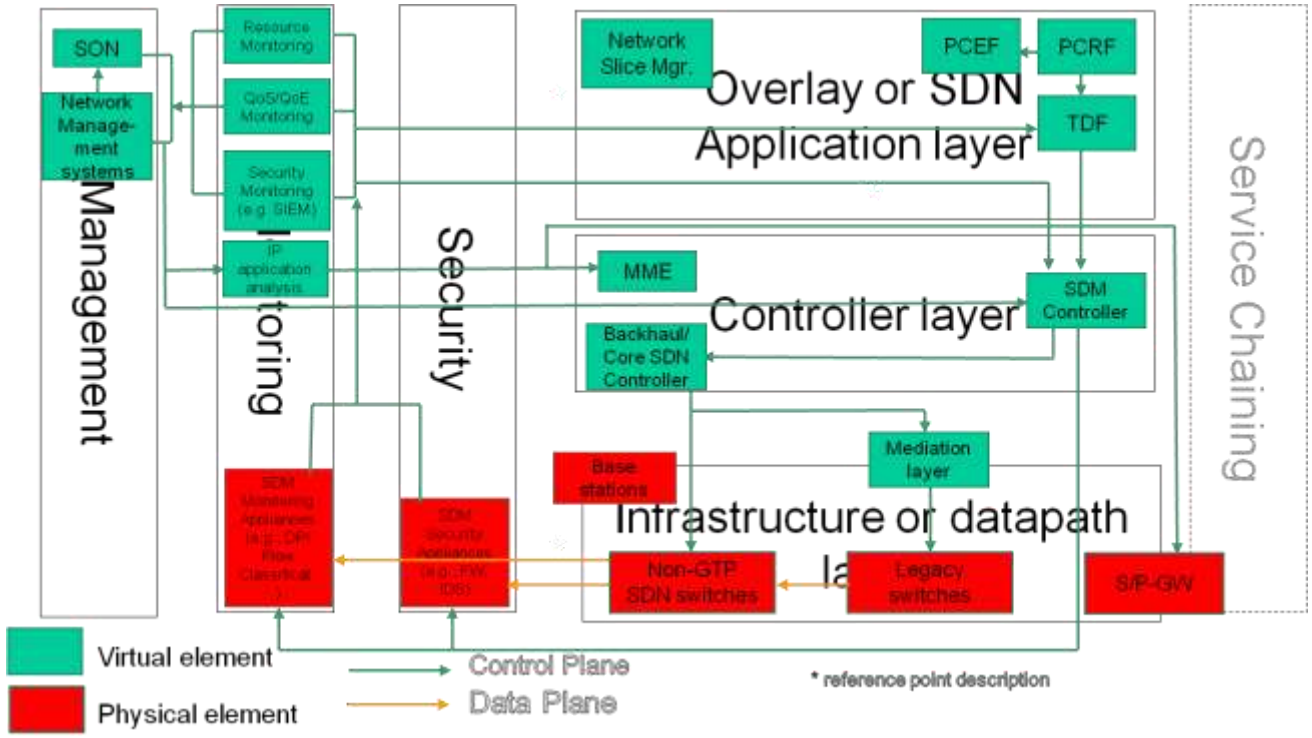

Figure 17. Virtualized EPC/Network performance and resource monitoring

- SDN Application layer
  - **Traffic Detection Function** (TDF): The PCRF instructs the TDF (DPI device) using Sd interface to look for specific application flows. The TDF will alert the PCRF using Sd when these flows are detected. Then the PCRF can instruct the PCEF to install a change rule using Gx. For this the TDF informs the SDM to set up the detection and the Monitoring application will notify the TDF.
- Controller layer
  - **Software Defined Monitoring** (SDM): The SDM controller allows controlling SDM enabled appliances to optimize the metadata extraction to what is needed by the network functions. The SDM needs to interact with the SDN to set up network taps or traffic mirroring for the monitoring and security appliances. The Management System and SON will manage the SDM controllers and define what policies need to be applied, depending on the state of the network provided by the monitoring applications.
  - **Backhaul/Core SDN Controller:** SDN controller used in case of SDN enabled data plane
- Datapath layer

  - **SDN switches** will be configured by the SDN controllers to send traffic or copy of traffic to the monitoring and security appliances.
  - **PGW, SGW, Base station**: according to 3GPP standards (some of elements may be virtualized)
- Management
  - Management system and SON algorithms will allow defining policies and rules that need to be enforced. They will also obtain feedback from the monitoring applications to supervise the network and manually or automatically adapt the policies and rules to changes in the network.
- Monitoring & Security
  - **QoS, resource and security monitoring** applications will analyze the metadata provided by the monitoring and security appliances. In turn these applications will inform the network functions (e.g., management, SON, TDF) that require information on the state of the network and application flows. Eventually, they will inform the SDM controllers, if they act as decision points.
  - **LTE User and Control plane monitoring:** Call and session analysis combined with IP application analysis with correlation.
  - **Performance Monitoring**: DPI based analysis for QoS/QoE measurements
  - **Security monitoring**: DPI based analysis to detect and eliminate unwanted traffic, including functions for threat detection

Main use case(s): Network performance and resource monitoring.

Related basic assumption(s): Network monitoring aware of virtualization.

Main targeted benefit(s):

Scalable network monitoring targeted to the needs of applications and network functions. This can be referred to respectively as Application Centric Network Monitoring and Software Defined Monitoring. Furthermore, combining performance and security monitoring and analysis will bring benefits to both security and QoS management.

Create technical concept for a novel distributed (SDN/NFV-based) QoE monitoring system and new algorithms for QoS/QoE mapping, implementation of these algorithms.

4.3.1.4 Main research questions

1. How mobile network virtualization affects network monitoring?
2. How benefits of the virtualization and cloudification can be exploited in network monitoring solutions?
3. How to utilize performance measurements in SDMN?
4. Effects of separation of the control and user planes with SDN techniques
5. Monitoring of Open Flow control traffic
6. Challenges to correlate monitoring information over several monitoring points
7. Traffic capturing possibilities
8. Scalability and performance of the monitoring solution

4.4 Mobility management

SDN and NFV techniques provide new opportunities for flexible mobility management. Current 3GPP networks make use of mobility management protocols, which are based on different IP tunneling options (MIP, PMIP and GTP). Further study is needed to understand the real advantages of the SDN-based mobility management, compared to the existing Distributed Mobility Management solutions.

In this and two following chapters the following Architecture layer and reference point numbering has been used:

- Potential architecture layers:
  - 1: Infrastructure or datapath layer
  - 2: Controller layer
  - 3: Overlay or SDN application layer

  - 4: Management
  - 5: Monitoring
  - 6: Security
- Reference points:
  - 1-2: infrastructure or datapath layer – controller layer
  - 2-3: controller layer – overlay or SDN application layer
  - 1-4 / 1-5 / 1-6: infrastructure layer – management / monitoring / security
  - 2-4 / 2-5 / 2-6: controller layer – management / monitoring / security
  - 3-4 / 3-5 / 3-6: overlay or SDN application layer – management / monitoring / security

4.4.1 Terminal based Mobility management and related tunneling

The user and flow mobility management concepts can be grouped into two main groups: (1) purely SDN-based approaches, and (2) post-distributed mobility management (post-DMM) solutions. The mobility management method described in Section 4.4.1.1 belongs to purely SDN-based approaches, while the methods in Sections 4.4.1.2 and 4.4.1.3 are two solutions, which integrate the existing distributed mobility management solutions into SDMN.

4.4.1.1 OpenFlow based mobility management for SDMNs

This use case extends OpenFlow based technologies for efficient LTE/EPC mobility management in SDMNs with heterogeneous access environment, utilizing IEEE802.21 MIH / ANDSF support for vertical handovers and proactive operation. This provides handover optimization and multi-access support capabilities.

The SDN controller is coupled with the 802.21 MIH / ANDSF enforcer functions in order to manage and control mobility independently of the access technology, optimize path based on the wireless and wired link conditions, provide seamless flow-level handovers, eliminate the tunneling overhead and reduce the overhead.

The UEs provide interfaces to control, query and manage access links with the IEEE 802.21 / ANDSF functionality. Network information server instances (ANDSF / MIH MIIS) provide static information about the access environment. PCRF handles authorization and policy decision: considering that a mobility event will occur in the future, proactive operation is naturally supported. The dynamic information can also be applied.

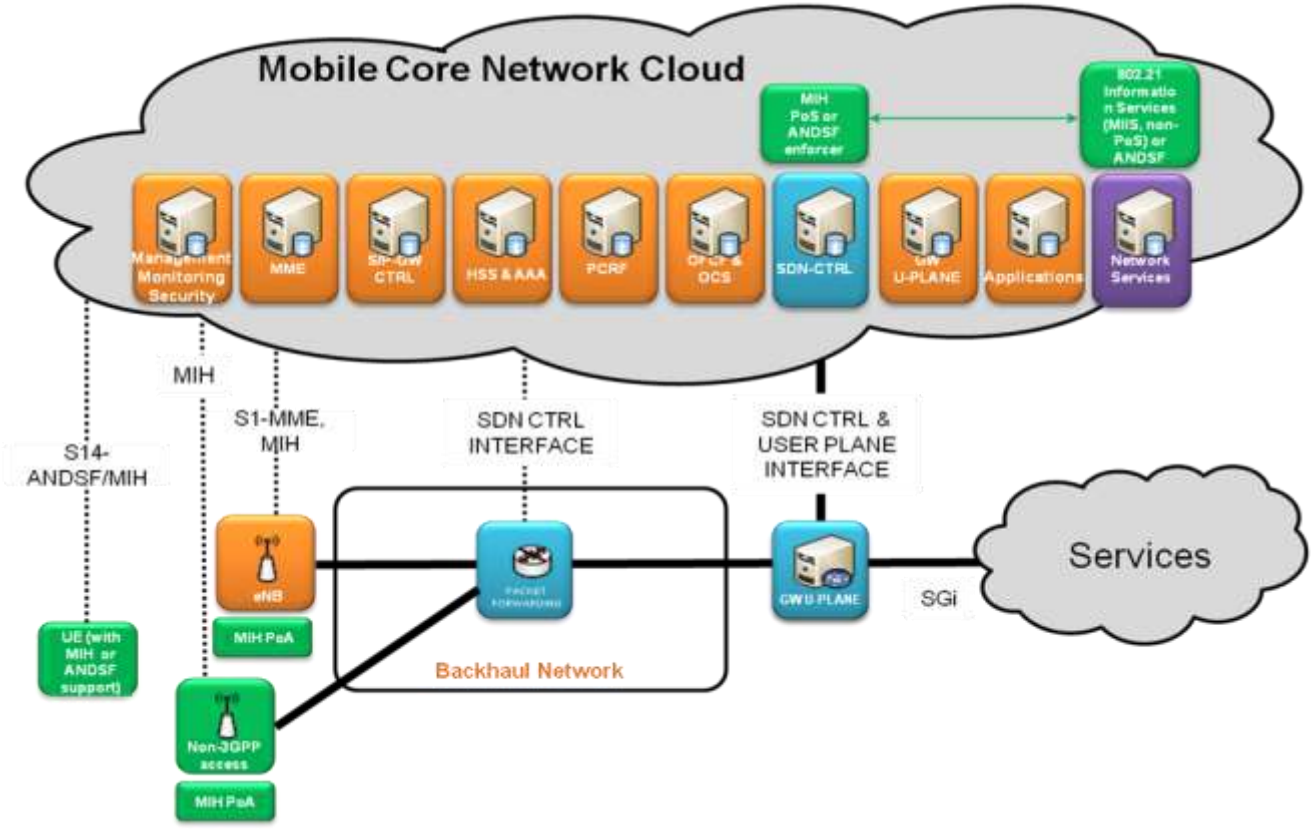

Figure 18.Terminal and OpenFlow-based mobility management for SDMNs

By the integration of centralized and/or distributed anchors (of post-DMM solutions) with the SDN forwarding functions QoS/QoE driven mobility management and support of complex mobility scenarios will be supported. Purely SDN-based mobility management solutions provide mobility transparency to higher layers even without applying additional tunneling.

4.4.1.1.1 Mapping to layered SDN model and reference points

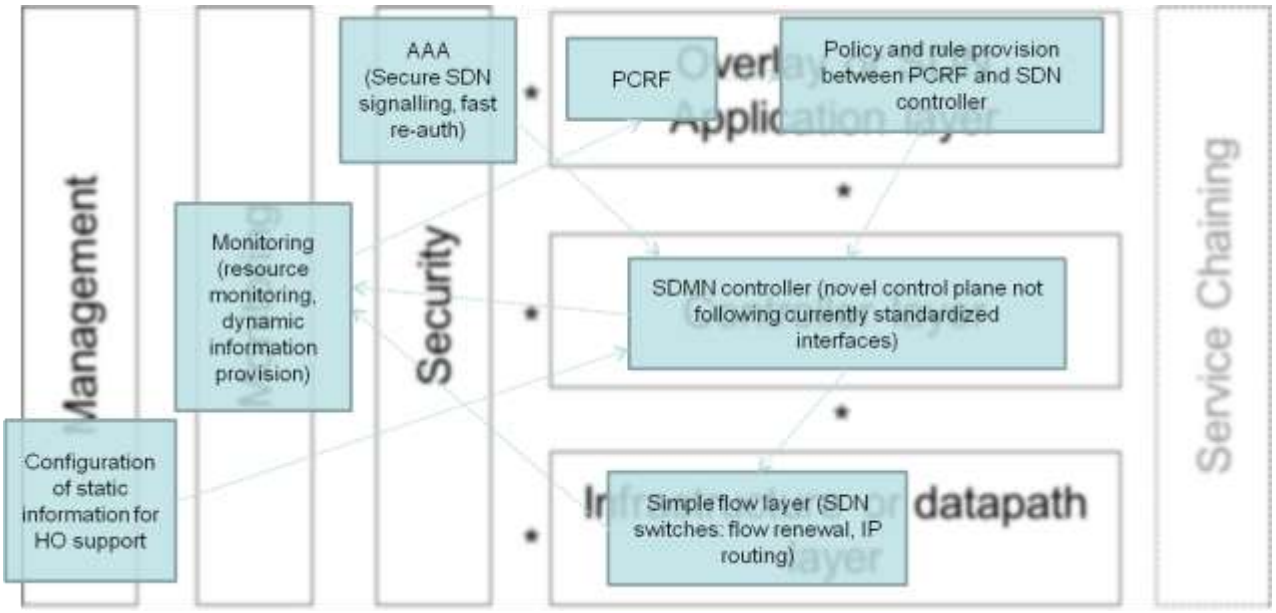

Figure 19. SDN-based mobility management

Use case	Touched Architecture layers	Notes
SDN-based mobility management	1: 2: 5: 6: 3: 4:	SDN switches extended with mobility management capabilities SDMN Controller (novel controller with special functionalities to support SDN level mobility management) Resource monitoring and information provision to the controller for advanced HO support Secure SDN signaling and cooperation with the controller for (fast) re-authentication Policy and rule provision between PCRF and SDN controller for network-initiated/supported HOs Configuration of static information for HO optimization (e.g., in the ANDSF server)

Use case	Reference points	Notes
SDN-based mobility management	1-2: 1-5, 2-5: 2-3: 2-4: 2-6:	OF: FlowMod Monitoring Dynamic policy rule provision for HO support Static information provision Secure signaling, fast re-authentication

4.4.1.2 Post DMM: Extension of HIP-based DMM solutions

This research topic deals with the integration of Host Identity Protocol-based Ultra Flat Architecture (UFA HIP) concept [13] into SDMNs, in order to enable scalable mobility management. Host Identity Protocol (HIP) supports secure IP connectivity using IPSec security associations, access-technology agnostic user access authorization, network security, IP mobility management, multihoming, IPv4/IPv6 interworking for any IP-based application. Its disadvantage is that end-to-end SAs are inappropriate in operator-controlled networks, service continuity is not supported in case of inter-GW handovers, and until this time, no killer application has been realized for its widespread deployment. In case of end-to-end HIP, a terminal-based mobility management solution, the network has no ability to control and decrypt IPSec communication, which encumbers, e.g., traffic control, mobility management, adequate accounting, lawful interception by the operator. Additionally, a terminal-based control may cause unnecessary high network and computational overhead on the user equipments (UEs) and in the access networks, in an environment where radio resources are expensive.

UFA HIP introduces a hop-by-hop traffic forwarding approach, i.e., it divides the E-E SAs into two segments–one between the UE and the GW and the other between the GW and the peer of the UE. I.e., the trust model is changed, users trust the (authenticated) network for not using user data with malicious intent. The split of SAs leads to the following problem during mobility of terminals in distributed networks: inter-GW handovers result in session mobility for the HIP host associations (HAs) and IPSec SAs on the network-side. UFA HIP solves the inter-GW handover problem using proactive context-transfers by integrating IEEE 802.21 Media Independent Handover (MIH) protocol, HIP and L2-access authorization. It has been shown in [14][15] that UFA HIP has significant performance gains in terms of signaling load on user side, in the access networks and at the rendezvous service over the terminal-based architecture due to the application of delegation of signaling rights. On the other hand, the former architecture results in greater signaling overhead in the transport network connecting the gateways and at the gateways. Naturally, the decryption and encryption of user data on the GW side requires high amount of computational resources.

Another benefit of UFA HIP is the small handover latencies provided by the proactive behavior based enhanced handover suit. The solution strongly relies on cross-layer communication using MIH. It has been depicted in [16] that both the handover latency and the packet loss was significantly decreased in UFA HIP scenarios compared to basic HIP and MIPv6 use-cases.

An important constraint of UFA HIP concept is that all GWs must be trusted. UFA HIP concept must be deployed in closed scenarios, i.e., we must guarantee that HIP-enabled nodes should only communicate with HIP-enabled nodes due to the following security reasons. Allowing in connections of non-HIP enabled nodes (through HIP - non-HIP proxies), or allowing a HIP-enabled node to open ports on non secure channels can potentially lead to the bypass of IPSec firewall, and further attacks. Hence UFA HIP is appropriate only for secure, private VPN-like scenarios where both UEs and GWs are HIP-based. In such scenarios trust in all GWs is not the limitation, if GW functions are segregated from other parts of the MNO's infrastructure.

A natural evolution step of UFA HIP is to put GWs into the mobile network cloud, and use SDN techniques for appropriate GW selection, optimal path selection between the source and target GWs during and after handover events, efficient support of different levels of GW distribution, more lightweight UFA GWs due to their virtualization on top of high-performance hosts. The benefit of integration of IEEE 802.21 MIH/ANDSF with the controller could solve proactive handovers.

The main components needed by this mobility management solution are the following: SDN supported HIP signaling overlay in the cloud. MIH Media Independent Information Service (MIIS) in the cloud, HIP-compatible UEs, CNs and UFA GWs distributed in the cloud. HIP and MIH signaling must be implemented between the GW nodes and the cloud.

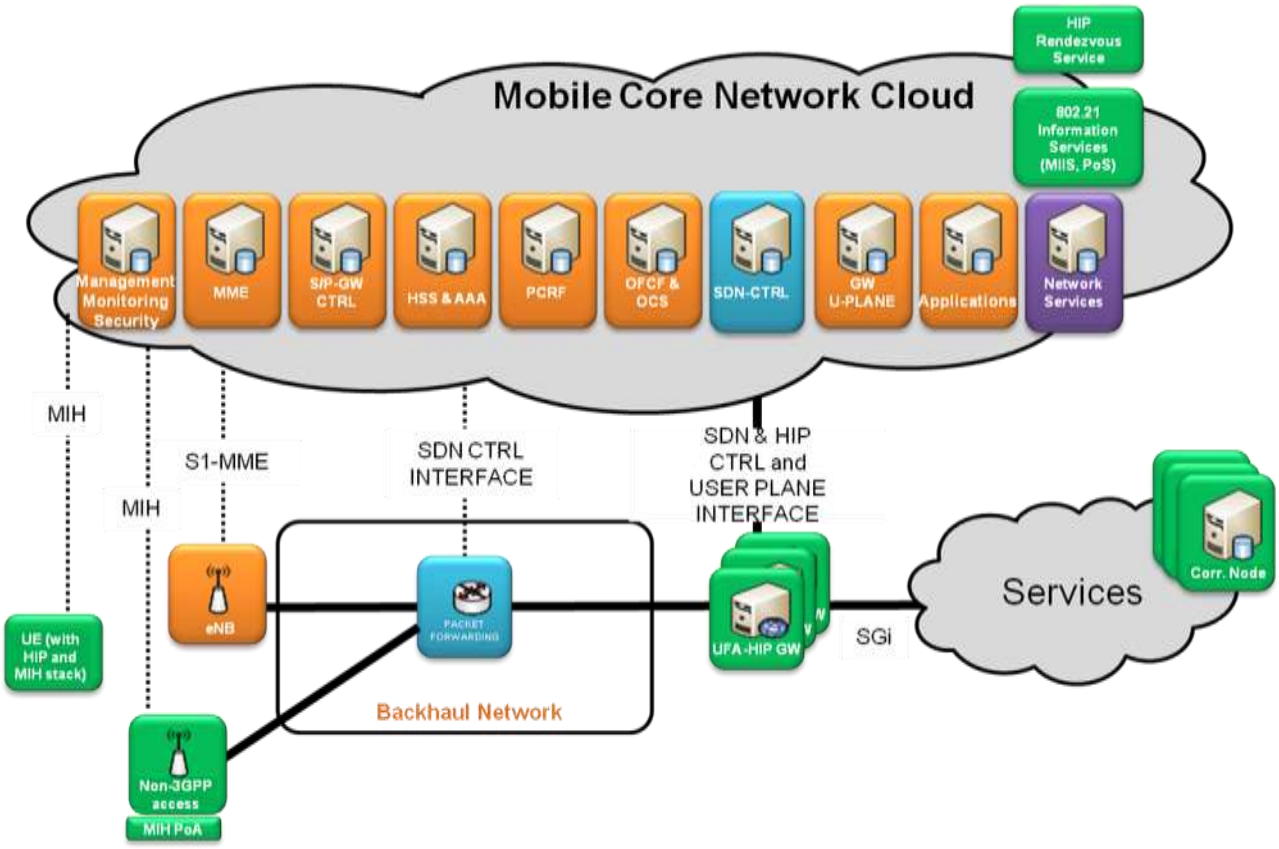

Figure 20. SDMN-aware UFA-HIP mobility management

4.4.1.2.1 Mapping to layered SDN model and reference points

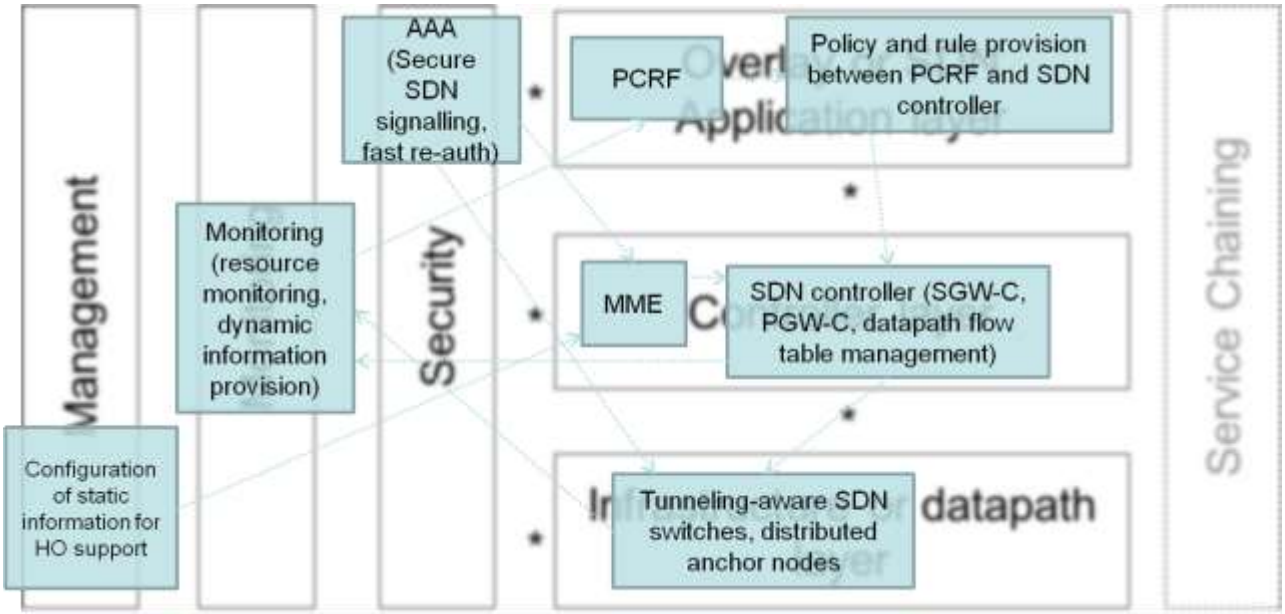

Figure 21. Post DMM principles with SDN control

Use case	Touched Architecture layers	Notes
Post-DMM	1: 2: 5: 6: 3: 4:	tunneling-aware SDN switches, distributed anchor nodes MME, SDN Controller (with special functionalities to support tunneling-based mobility management solutions) Resource monitoring and information provision to the controller for advanced HO support Secure mobility signaling and (fast) re-authentication Policy and rule provision between PCRF and distributed anchor nodes (and their controllers) for network-initiated/supported HOs Configuration of static information for HO optimization (e.g., in the ANDSF server)

Use case	Reference points	Notes
Post-DMM	1-2: 1-5, 2-5: 2-3: 2-4: 2-6:	OF: FlowMod Monitoring Dynamic policy rule provision for HO support Static information provision Secure signaling, fast re-authentication

Use cases:

- Impact of mobility in SDN based mobile networks: Post-DMM solutions for scalable mobility management
- Impact of mobility in SDN based mobile networks: SDN based extensions to LTE/EPC mobility management

Basic assumptions:

- Assumption on mobility management: SDN based mobility management versus 3GPP and MIP mobility
- Assumption on locator and identity assignment to UEs in SDMNs: the current practice will not change even if P-GW-C is virtualized vs. new identity/locator assignment solutions are needed

4.4.1.3 Post DMM: Proxy MIPv6 in SDMNs

The Proxy Mobile IPv6 (PMIPv6) in the SDMN topic investigates concrete evolution of the standardized mobility management protocol for the SDN-NFV architectures. Considering the recent efforts of shaping this protocol for distributed/flat architectures, the objective is to address the specification details for efficient integration into an SDMN infrastructure (e.g. integration with service chaining, coordination with the orchestrator, etc.). Furthermore, handling of the SDN control plane, emulation of network functions (Mobile Access Gateway – MAG, Local Mobility Anchor - LMA), and relocation of the LMA function on a per user basis will be investigated.

There will be developed a new SDN southbound protocol for fine-grained wireless performance monitoring. A SDN southbound protocol is able to provide low-level link technology information such as Received Signal Strength Indicator, wireless channel frequency, information about neighboring wireless access-points. There is needed new

specification of SDN software for the controller and on the infrastructure devices. There will be an implementation of the NFV architecture relying on the OpenStack Cloud infrastructure with the necessary orchestration capabilities.

Targeted benefit(s):

An important part of this topic objective is to design the PMIPv6 evolution that takes full advantage of the SDN-NFV concept. We thus expect higher flexibility in the solution deployment, reduced overhead cost (by mean of IP-in-IP tunnel removal for instance). The targeted benefits are related to the data plane performance. The latency caused by the control plane (mobility management procedure, routing establishment, etc.) and the resulting data plane throughputs are of paramount importance.

4.4.2 Network based Mobility management and related security concerns

Network based mobility management is another interesting area of mobility use cases. One of the first basic concepts in this field is the basic support for network mobility in Mobile IPv6. Many existing distributed mobility management solutions have network mobility extension, which means that the movement of a whole network, belonging to an IP address prefix, is supported.

4.4.2.1 Secure Wi-Fi network mobility with SDN and HIP based mobile switching

Wireless operators built networks with end-to-end proprietary Wi-Fi solutions provided by incumbent vendors. Operators and end-customers were locked into proprietary platforms, paying for capabilities they do not need, and waiting for vendor-provided upgrades to new features. This approach defined a generation of business models that burdened wireless operators and end-customers with capital equipment costs, while failing to provide the agility required meeting the rapidly changing demands of mobile users and applications. Thus, SDN Wi-Fi is getting research attention in the field of software defined networking. Mobility management provided by OpenFlow based technologies for LTE/EPC with heterogeneous access environment is studied. The client authentication and backhaul security for mobile switches are based on the OpenFlow in the multi-access environment.

For SDN to really work for the “entire” network, it must include the Wi-Fi part under the control. Wi-Fi is no longer a toy or something off on the side that is a nice to have. It is a required part of the overall business network and users demand that it be available all the time, transparent, and provides the same feature benefit of the wired network. The Wi-Fi network has a unique set of challenges that make the need for centralized control with distributed intelligence imperative:

1. Mobile devices are constantly on the move and users expect a consistent user experience regardless of from where they are connecting. Connectivity must be ubiquitous and the rules for connecting must be able to transparently and dynamically change as the user/device roams throughout the connected world.
2. Based on several factors, such as meeting compliance mandates, traffic engineering, security, and distributed office locations the best data path for Wi-Fi traffic changes. For some traffic it might be best to keep it localized to the access point (AP), while in other cases a centralized traffic approach would be the best way to go or in some cases the ability to simultaneously do both should provide this required flexibility.
3. As discussed above an open API into the centralized management system is necessary to allow the easy integration of any application required by the business. This programmability might be the integration of a mobile device management (MDM) system to manage mobile devices when they are not connected to the network, additional reporting tools for compliance, or any number of the hundreds of applications in use by businesses.

Mobile base stations, mobile routers, switches, mobile clouds and server migration are some of the compelling reasons why mobility is insisted in software defined networks. This would extend SDN benefits into mobile environments e.g., moving trains, buses, flights and other automobiles. The existing competing solutions with Mobile IP (MIP) still have problems related to “triangle routing” and drop of IP packets due to frequent handover when the host is away from Home Agent (HA). Mobility offers many dazzling opportunities that also bring with them some profound challenges related to security and privacy. We argue that mobile IP has security flaws against DoS, passive eavesdropping, insider attack, replay attack, tunnel spoofing and location privacy.

Security in IP based networks is widely tackled with a common set of protocols composed of secure file transfer protocol (SFTP), secure socket layer (SSL), and transport layer security (TLS). OpenFlow enables an acceptable level of security with SSL or TLS though; it does not support mobility. These limitations insist modifications to the current OpenFlow architecture. More specifically, below we describe the major problems of present OpenFlow version.

1. **Flow processing**: Change of address would disrupt flow processing from network switches. Therefore, they require fast and regular updates to flow tables.
2. **Secure session management**: Changing an IP address may also tear down active SSL/TCP sessions.
3. **Secure handover**: Problem of mutual authentication and cannot support mobility and certificate exchange is not preferable for fast moving OpenFlow clients
4. **Flow rule management**: Change of IP address to solve latter issue causes additional overhead, since flow rules must be updated frequently.

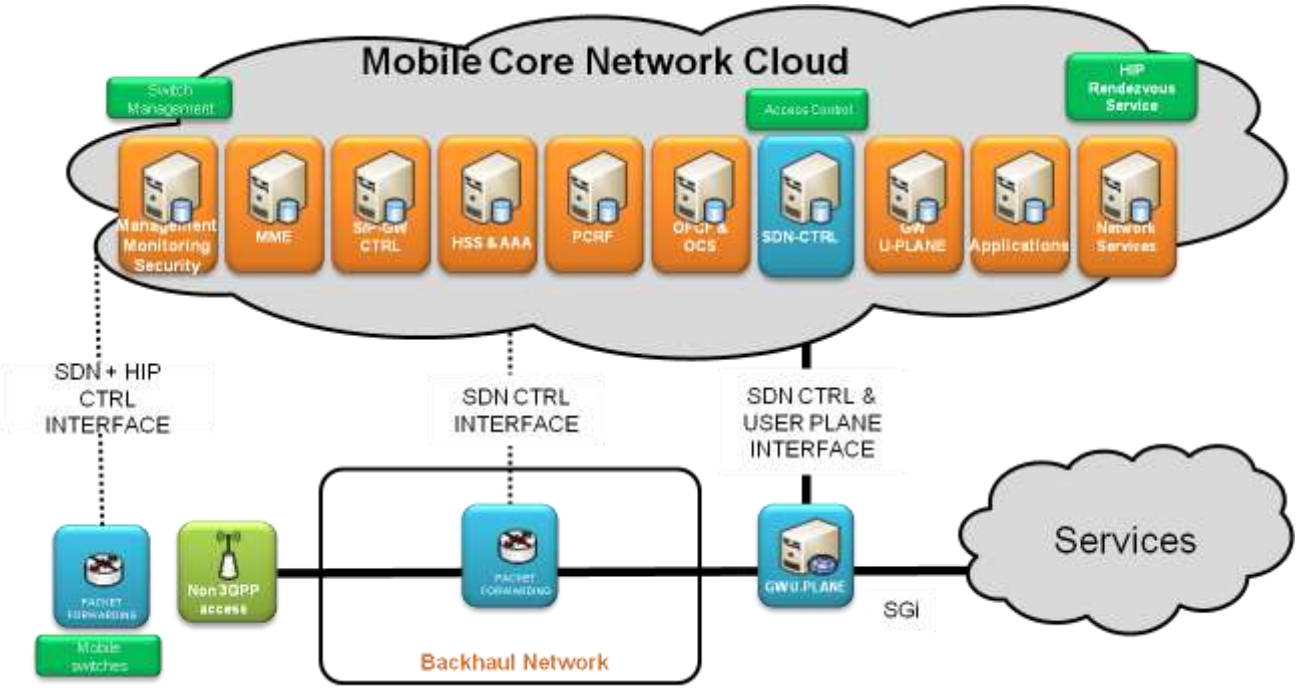

Figure 22. Secure network mobility with OpenFlow controlled Wi-Fi architecture

SDN based mobility scenario is presented, where the switches or routers are configured on the fly. The flow control agents (FCA) on the mobile switches and routers are responsible for updating the software controller of the new location information for location based services. Besides that, the flow rules could be built on top the newly introduced cryptographic global identifiers which do not change over time.

Thus, updating the flow rules due to mobility is no longer required. This would reduce the processing and control traffic overhead due to dynamic address configuration and thus, flow processing. Furthermore, it reduces overhead in the policy charging and rules function (PCRF) and improve the flexibility with QoS management and network configurations. Moreover, controller can keep records of HITs and authenticate them through controller signed certificates or using other techniques, such as DNSSEC and DHT-based verification.

4.4.2.1.1 Mapping to layered SDN model and reference points

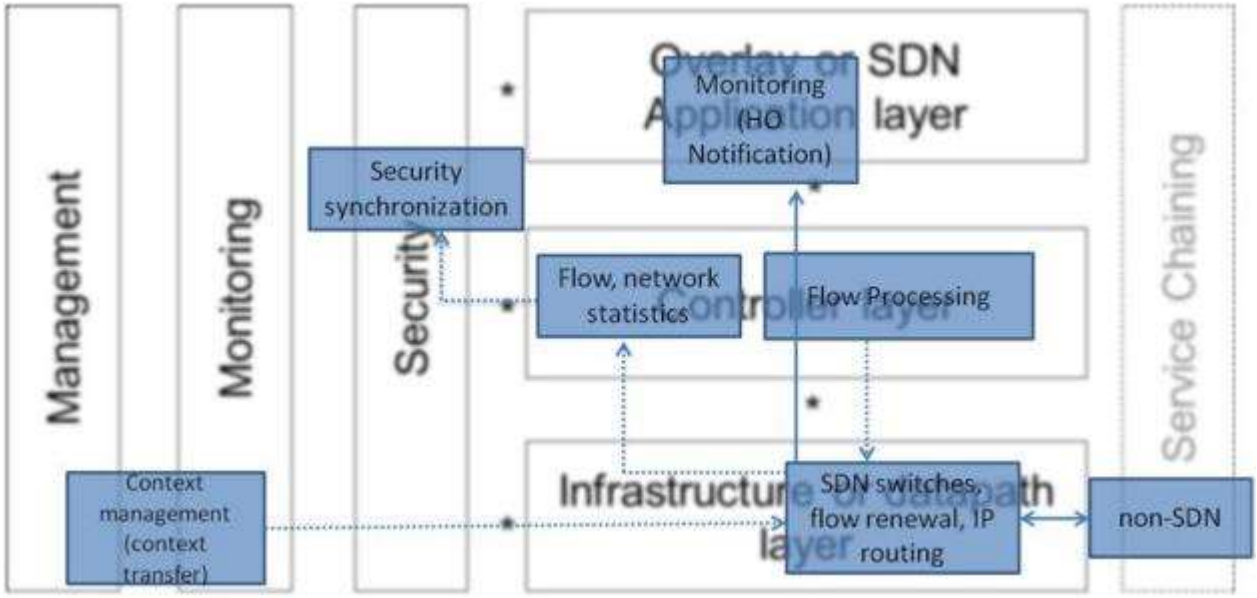

Figure 23. Switch/controller of the Wi-Fi mobility

Components:

- The SDN controller is coupled with Radius server and out-band signaling:
  - Mobility between different technological domains
  - Pre-configuration of flow paths to enhance mobility
  - Enabling seamless mobility between OpenFlow enabled access points and flow-level handovers
- Access points are OpenFlow enabled and provide standard IP Ethernet interfaces to controller and WLAN (802.11) interface to wireless users
  - EPC based reachability to Radius server for user authentication
- SDN controller gathers information from the mobile switches and compares them to assure security and seamless service continuity.
- Load balancing between different access technologies based on the network slices with FlowVisor.
  - Dynamic client based load balancing at the 802.11 environment.
  - Possibility extend it towards Cognitive network section

Use case	Reference points	Notes
Switch/controller Mobility (CWC)	1-2 1-5 2-5	OpenFlow packet processing (Packetln) Monitoring (Retrieve network statistics from the controller) Ex_: Floodlight support
SDN Wi-Fi	1-2 1-5 3-5	Secure signaling, fast authentication Mobile IP assignment

Use cases:

- Switch/Controller mobility in SDN environment: Identity/locator separation and in SDMNs

- Wireless SDN and seamless mobility: Application level development to track moving devices and management of seamless mobility.

Basic assumptions:

- Assumption on Quality of Service: QoS provision in virtualized mobile core networks in SDN-based network forwarding paths
- Assumption on mobility management: SDN based mobility management versus 3GPP and HIP mobility
- Assumption on locator and identity assignment to UEs in SDMNs

4.5 Resource and traffic management

SDN and NFV techniques provide new opportunities for flexible traffic engineering. Resource and traffic management are closely linked to each other. The resource management decides which and how many virtual network function instances are provided at which cloud datacenter location and determines the optimal traffic routing within the (possibly multi layered) transport network to reach these functions. The task of the traffic management is to react on traffic fluctuations (for which an adjustment of the cloud resources would not be reasonable) by adjusting the routing paths according to the operator's policy (e.g. load balancing target).

4.5.1 Traffic Management and optimization

In the SDMN architecture the routing paths are not determined by decentralized routing protocols (like in today's IP networks), but by the SDN controller. The traffic management to some extent forms an inner control loop (path adjustment while keeping the virtual network function instances constant) which is executed more frequently compared to the resource management (outer control loop). Therefore traffic management allows for a fine-tuned optimization. Due to the combined traffic and resource management the cloud based SDMN architecture provides more degree of freedom in optimization than traffic engineering in the traditional mobile network architecture.

This work comprises the development of optimization algorithms for the traffic management within the SDN-enabled transport network (in the mobile core domain) to react on traffic fluctuations and overload situations. The optimization target is to perform an optimal load balancing so as to achieve the best possible network quality (QoS) within the limits of the given network and cloud resources. Moreover the interaction with the resource management will be considered. As solving the optimization models for large-scale problems probably is not possible by applying exact solution methods, heuristic solutions will be considered as well. The evaluation of the algorithms is performed by simulation.

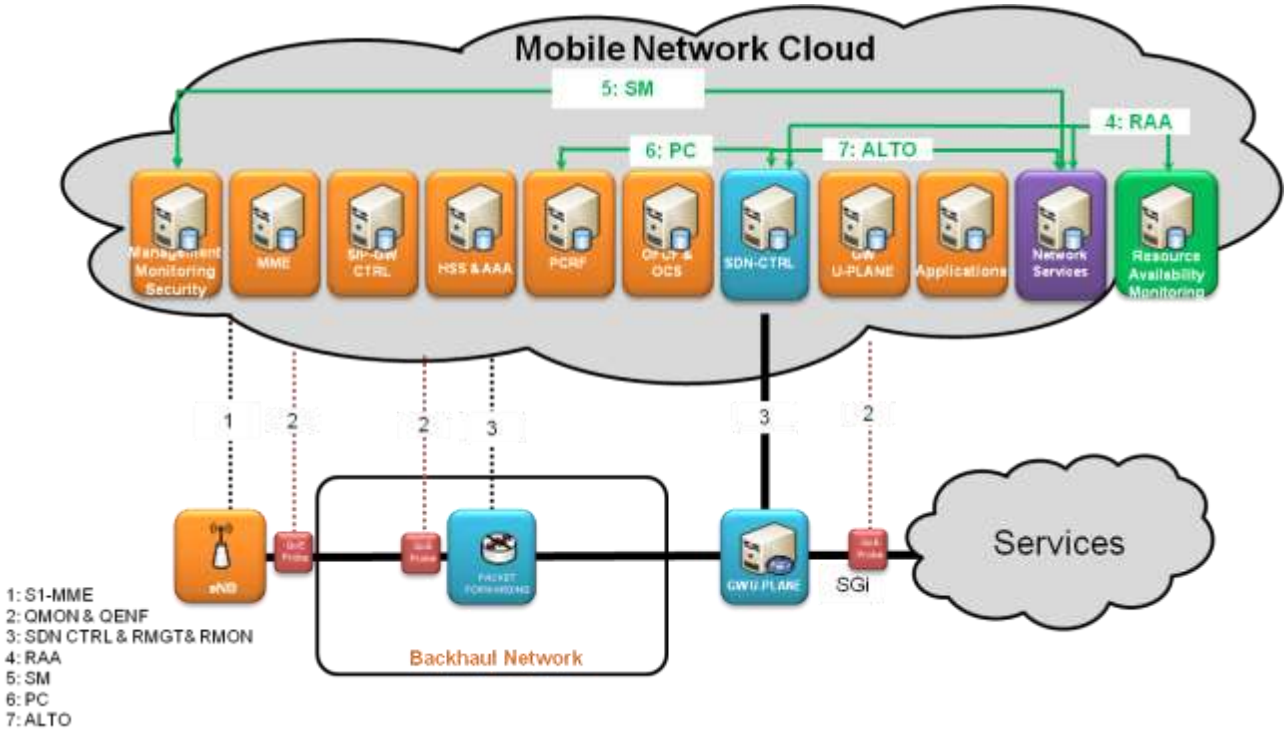

Figure 24. A consolidated view of traffic management (TM) and Resource management (RM)

Impacted and added components in the Mobile network cloud are: Resource Availability Monitoring, PCRF, SDN-CTRL and Network services. For the Forwarding plane the QoE probes are the main addition in this concept.

Impacted and added horizontal interfaces by the studied concepts are:

- Resource Availability Awareness (RAA): SDN controller REST API to Resource Monitoring GUI (RMon)
- Switch Management (SM): for joint routing and resource management
- Policy Control interface (PC): PCRF Rx, Gx [19] to SDN controller (REST API)
- Application-Layer Traffic Optimization (ALTO): ALTO Client-to-Server [18] , ALTO Network-to-Server interface from SDN network segments [17]

Impacted and added vertical interfaces:

- Resource management (RMGT): resource management, Qdisc configuration, etc.
- Resource Monitoring (RMON): retrieving information regarding network resources availability
- QoE/QoS monitoring (QMON): monitoring flow-level performance metrics
- QoS/QoS enforcement (QENF): enforces authorized QoS rules

Main target is to develop a new method for the optimization of traffic flows in the mobile core domain during network operation; implementation of the algorithm and evaluation for different realistic network and traffic scenarios

In the following sub-chapters these topics are described in more details.

4.5.1.1 SDN based application layer traffic optimization and QoS enforcement

This research work addresses two different areas: SDN based Application-Layer Traffic Optimization and DiffServ QoS enforcement architecture for SDN transport with the policy control.

Application-layer traffic optimization provides network, application and service endpoint awareness. ALTO-SDN integration could provide several benefits for (V)MNOs and CDNs (see targeted benefits below).

The main research topics addressed are:

i) integration of ALTO network information service into SDNs for network and service endpoint aware traffic management

ii) enforcement of QoS policies using DiffServ QoS architecture and PCRF in SDN-based network segments

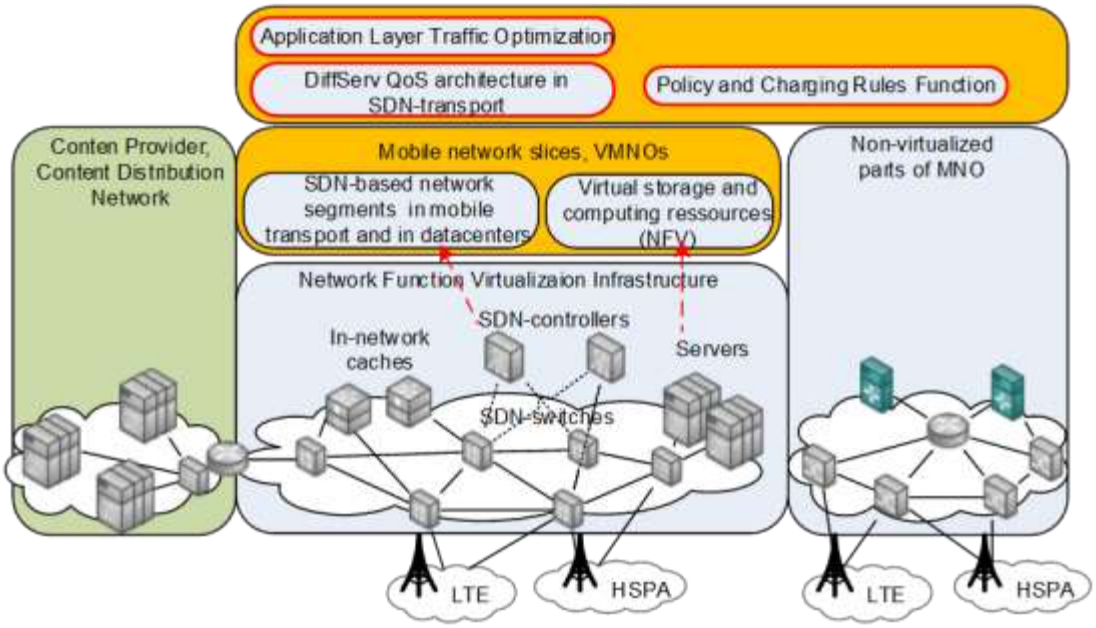

Figure 25. Integration of Application-layer Traffic Optimization and DiffServ QoS in SDMNs.

Main targeted benefits:

By the integration of ALTO network information service into SDNs ALTO becomes transparent for the UE or the service claimant entity (no deployment cost in the UE). Due to ALTO information, the ALTO client in the SDN controller can overwrite the initial peer selection decision of the service claimant entity (e.g. UE). Any flow can be dynamically selected for getting ALTO guidance and SDN provides built-in redirection mechanisms.

DiffServ QoS architecture can provide soft QoS guarantees on top of different Layer-2 technologies in a scalable manner. However, in current SDN standards (e.g., OpenFlow, OpenFlow Configuration) QoS management is still in early stage. Our research will focus on the discovery of QoS capabilities of SDN-transport (traffic classification, shaping, policing, dropping), in order to provide QoS guarantees in partly virtual network forwarding paths of MNOs and VMNOs.

Use cases:

- Macroscopic traffic management (traffic engineering) with SDN: Application-level traffic optimization in SDMNs
- Microscopic traffic management with SDN: DiffServ QoS in SDN-based transport

Basic assumptions:

Assumption on Quality of Service: QoS provision in virtualized mobile core networks in SDN-based network forwarding paths

4.5.1.1.1 Mapping to layered SDN model and reference points

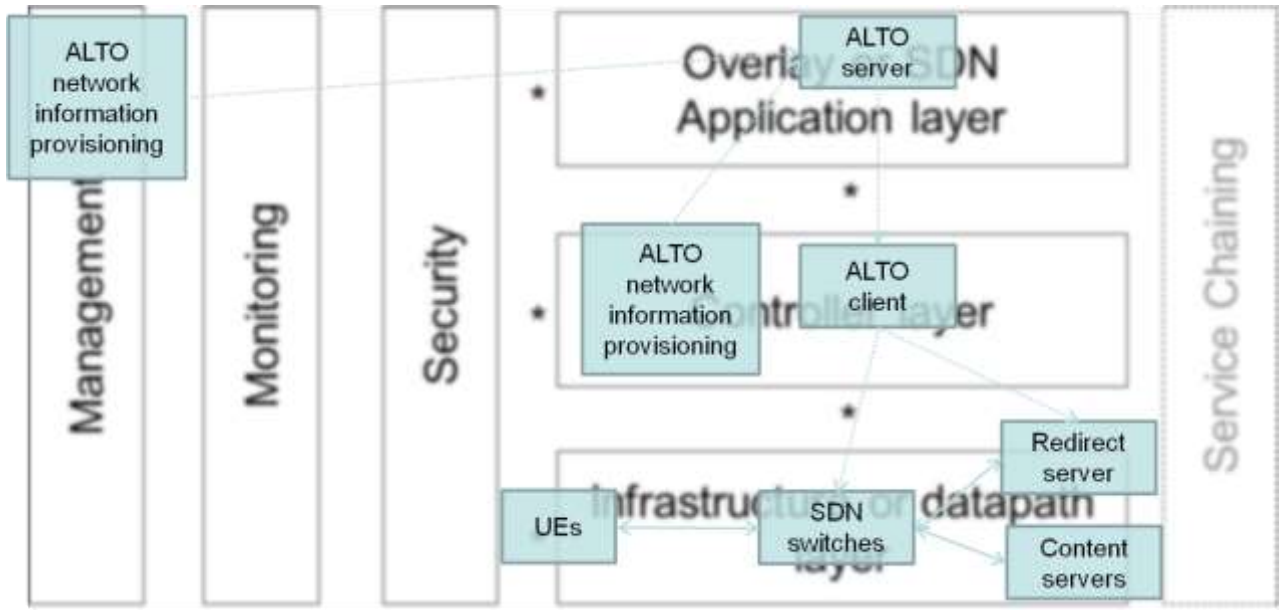

Figure 26. ALTO network information service integration into SDN

Use case	Touched Architecture layers	Notes
ALTO-SDN	2: 3: 4:	Flow programming, packet listener for ALTO service invocation ALTO guidance for peer selection (ALTO client), dynamic network information provision to the ALTO server ALTO network information provision

Use case	Reference points	Notes
ALTO-SDN	1-2: 2-3: 3-4:	OF: FlowMod, PacketIn, ALTO guidance, ALTO information provision ALTO information provision

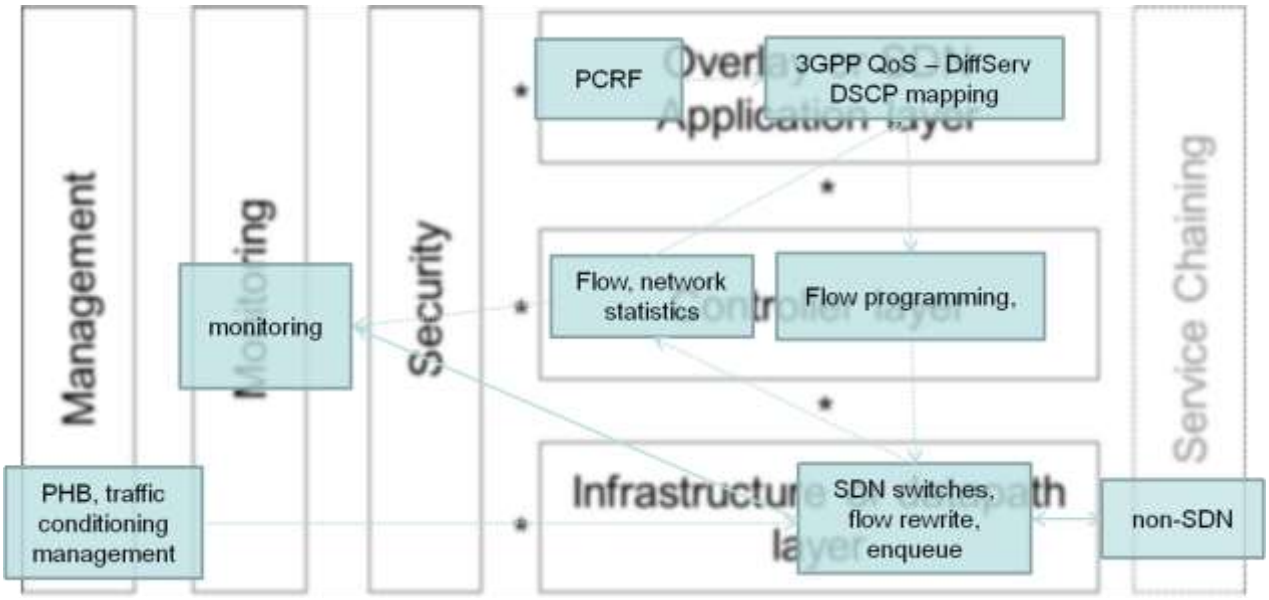

Figure 27. DiffServ QoS mapping to SDN

Use case	Touched Architecture layers	Notes
DiffServ QoS in SDN-based transport	1: 2: 4: (5):	Flow rewriting (DSCP), enqueue Flow entry addition / modification / removal Queue management, classification, DiffServ Per-hop behavior implementation Monitoring is needed in case of DiffServ with traffic engineering option

Use case	Reference points	Notes
DiffServ QoS	1-2: 2-3: 1-5, 2-5: 1-4:	OF: FlowMod (flow rewrite, enqueue, output), monitoring Dynamic policy rule provision for QoS mapping, flow mapping to traffic behavior aggregates Monitoring (e.g., sFlow), get SDN network and flow statistics from controller DiffServ PHB configuration

4.5.1.2 Coordinated traffic and resource management and efficient routing

Existing cellular networks suffer from inflexible and expensive equipment, complex control-plane protocols, and vendor-specific configuration interfaces. Thus, some efforts have been done to apply SDN concepts to cellular networks in order to simplify the design and management of cellular data networks. Compared with wired networks, cellular networks have some unique features and face significant scalability challenges. For example, because users are always moving in the cellular networks, there will be a large number of state updates generated from the data plane, which would create big pressure on a central controller. In addition, the average response latency would also increase sharply when a set of base stations communicate with one remote controller concurrently. As networks become more complex and traffic diversity increases, there is the apparent need to manage the traffic carried by the network.

The 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, controlling the bandwidth allocated to competing flows in a fair manner, minimizing latency and ensuring reliable operations when traffic patterns change or parts of the network fail. Joint resource and routing management has received a great deal of attention within the SDN community. The management can be categorized as centralized where a single entity performs allocation and this entity therefore requires all the global knowledge and necessary resources for that task. Then, the decisions are distributed over some or all of the physical nodes. Therefore, communication and cooperation protocols are required to coordinate the process and less than optimal results are expected. The allocation can be performed in statically such as the shortest distance path algorithm; traffic constraints based algorithm or dynamically such as by allowing adaptive changing on the basis of users' current demand and traffic of the networks.

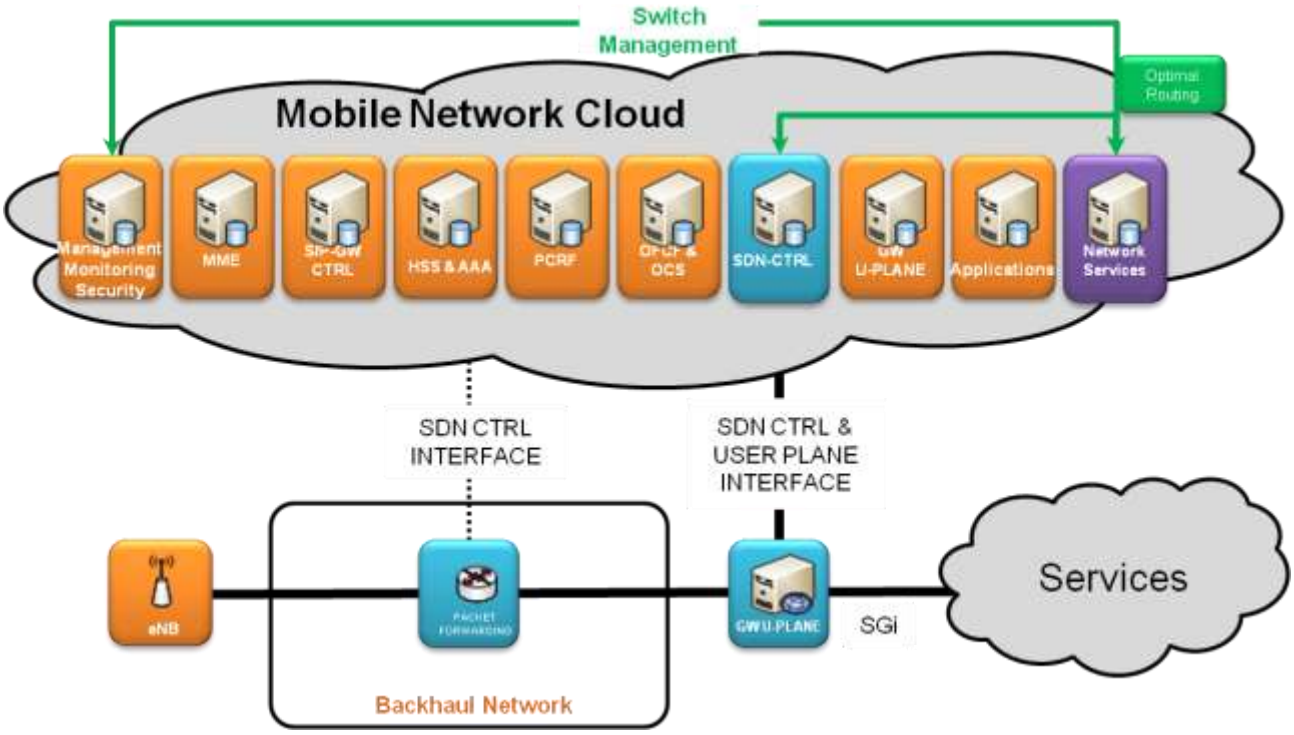

Figure 28. Joint Routing and Resource Management for optimal routing

This requires appropriate functions in Network Management, Monitoring, Security, SDN Control, Network Services and SDN switches.

For routing/resource management, most common approaches and techniques are heuristics. Besides, Mixed Integer Linear Programming (MILP) is also often used since the problem nature is NP-hard. Approaches using graph theory and multi-commodity flow is also examined but they may not take into account all of the requirements of the problem (especially node requirements). Other approaches are based on game theory, multi-agent systems, self-organization and policies.

Use case:

Optimized routing in the access, transport and core networks of SDN-based MNOs 2.1.8.1

Basic Assumptions:

- Assumption on Migration: Clean-slate,
- Assumption on network functionality: Network Functions in the cloud,
- Assumption on resilience: Cluster based SDN controller on active-passive mode,
- Assumption on Quality of Service,
- Assumption on mobility management: IP address distribution will not change in short or mid-term,
- Assumption on network monitoring: NM should be aware of network virtualization.

4.5.1.2.1 Mapping to layered SDN model and reference points

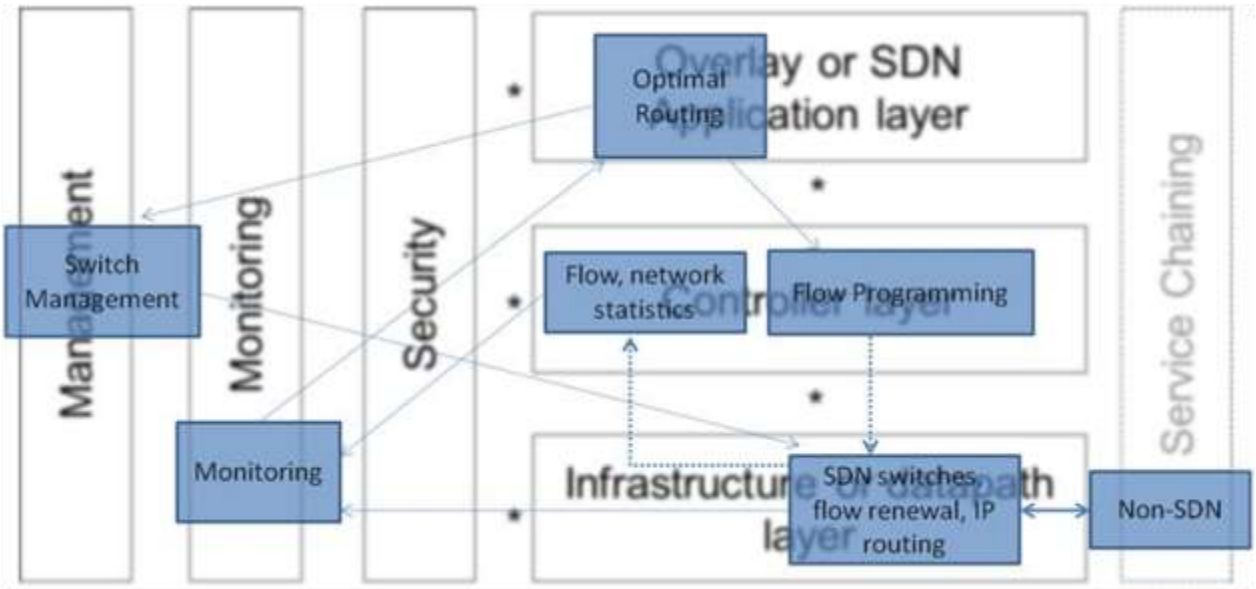

Figure 29. Joint Routing and Resource Management

Use case	Touched Architecture layers	Notes
Joint Routing and Resource Management	1: 2: 3: 4: 5:	OFConfig/OVSDB capable OF switches Flow programming and flow statistics collection Optimal routing application Switch management with OFConfig/OVSDB Resource monitoring for optimal routing calculation

Use case	Reference points	Notes
Joint Routing and Resource Management	1-2: 1-4: 1-5: 2-3: 2-5: 3-4: 3-5:	OF: FlowMod, PacketIn, …. Packet statistics OFConfig/OVSDB configuration Switch status Optimal routing guidance to Flow stats to monitoring unit Optimal routing guidance to switch management for OFConfig/OVSDB configuration on switch All monitoring info collected by optimal routing app from monitoring

4.5.1.2.2 Research approach

- To provide resource allocation techniques for mobile networks exploiting virtualization in regard of content-aware, operator centric, optimal treatment of user traffic with appropriate QoS/QoE.

- To apply the principles of control theory and/or game theory to further manage bandwidth and routing resources to varying traffic conditions.
- To analyze optimization possibilities for the existing traffic and resource management solutions/algorithms and to examine their suitability for heterogeneous mobile network architecture for designed semi-distributed algorithms.
- To investigate SDMN controller design for simultaneous transmission enabling network bonding.

4.5.1.3 SDN-controlled IP wireless mesh network for device-to-device communication

This research topic addresses the opportunity of offloading the Radio Access Network by the use of IP Wireless Mesh Network (e.g. Wi-Fi). Here we consider SDN control of IP communications in the edge networks, meaning that smartphones are SDN capable. It is commonly assumed that such a target would be handled by end-terminals themselves as a completely distributed system without the mobile network supervision. This research topic investigates whether the mobile network operator can keep control (supervision) of these communications to redirect traffic to a different access network (e.g., fixed).

The proposal targets to demonstrate establishment and control of an IP wireless mesh network by the use of NFV supervisor (through SDN control plane). The selection of an alternative access network gateway (for instance a fixed access point – a home premise box) could be the root point for the creation of an IP wireless mesh network with neighboring end-terminals.

A new SDN southbound protocol is developed for the fine-grained wireless performance monitoring. Implementation of the NFV architecture is relying on the OpenStack Cloud infrastructure with the necessary orchestration capabilities.

Targeted benefit(s)

The main targeted benefits of this concept are

- Data communications latency
- Overhead costs
- Level of reduction of traffic in the Radio Access Network

4.5.1.4 Secure mobile data offloading over SDMN

The proposal deals on Secure Mobile Data Offloading over SDMN. Due to the widespread diffusion of smartphones, tablets, and other mobile devices that we have experienced over the last years, the global mobile data is expected to grow exponentially over the next few years. This has become a major challenge for the mobile telecom operators over the world, who are experiencing severe problems in coping with the mobile data traffic generated by their users. As provisioning additional infrastructures bears significant costs both at the deployment and the management phases, the operators will need to decide to either drastically reduce the quality of service (QoS) for all the users, or block a significant fraction of the users to provide decent QoS to a few. Both alternatives are clearly quite far from an optimal solution. The research topic is aimed to pump the resource efficiency of the network, whether the system is coverage limited, capacity limited (or both).

- **Use Cases:** Resource management based on SDN. Macroscopic traffic management
- **Related Basic Assumptions:** Assumption on availability of resources.1.1.15. Assumption on Quality of Service: QoS provision in virtualized mobile core networks in SDN-based network forwarding paths 1.1.4. Assumption on managing the security. How to include security functions of physical and virtual elements and interfaces.1.1.9
- **Main Targeted benefits:** The main benefit for the system is in terms of contribution to resource efficiency, whether the system is coverage limited, capacity limited (or both). Resource management and availability will be covered on both virtual and physical resources.

4.5.1.4.1 Mapping to layered SDN model and reference points

In order to allow the traffic offload from the operator network, the proposed architecture includes some blocks to be implemented. These blocks must coordinate the dissemination of content. In order to reach this objective the proposed

architecture must deal with the content to disseminate, the list of clients interested in this content (subscribers), and service constraints to be met.

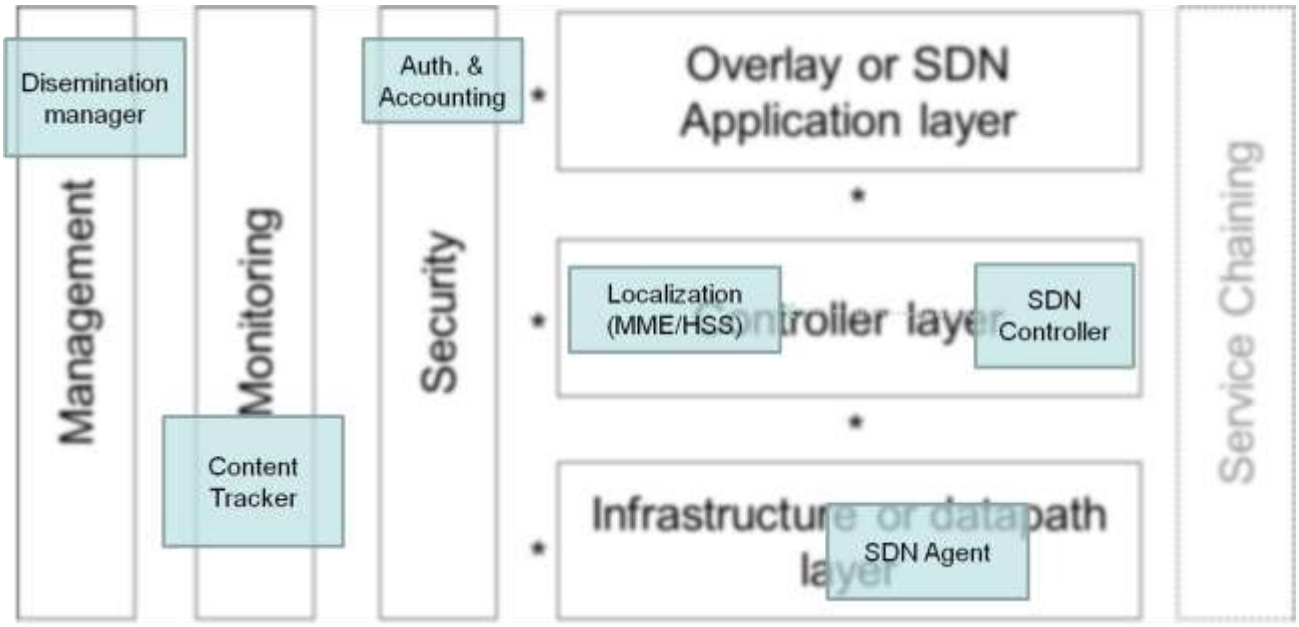

Figure 21. Secure Data Offloading Over SDMN

The proposed architecture must deal with offloading relaying on the following blocks:

- Management Layer:
  - Dissemination Manager: this block would be responsible for managing the offloading process. It will determine, from the list of subscribers that require the content and from localization and topological information gathered, the dissemination strategy to be applied (e.g. deliver a content to the subscribers A and B through the LTE network and ask them to rely the content in their neighborhood using Wi-Fi with specific routing policies).
- Monitoring Layer:
  - Content Tracker: this component receives acknowledgments sent by subscribers when they receive the content.
- Security Layer:
  - Authentication & Accounting: this module deals with the authentication, trust and credit management functionalities. It should also hold information about subscribers such as reputation or credentials status.
- Controller Layer:
  - Localization: deals with information on the localization of the subscribers (MME / HSS).
  - SDN Controller: Will work together with the Localization module to let it interact with the SDN dataplane.

4.5.1.4.2 Main research questions

- How to transfer secure mobile data offloading techniques from EPC core to SDMN?
- Can secure data offloading over SDMN be effective in order to avoid coverage or capacity limitations problems on the network?

4.5.1.5 QoS/QoE enforcement in virtualized ISAAR

The QoE Management framework ISAAR (Internet Service quality Assessment and Automatic Reaction) [20] can benefit from adding SDN functionalities. The quality monitoring of Internet services starts with identification of the respected flows within the traffic mix. Normally a lot of processing is needed for this detection, due to the Deep Package Inspection (DPI) required. To simplify that mechanism the OpenFlow (OF) matching rules can be used. Therefore, the OF controller has to be configured in a way which allows identifying measurable traffic by using the matching rules and teeing it out to one of the measuring points of ISAAR in real-time. This flow selective copy port mechanism allows for traffic steering of the relevant packets towards centralized measurement probes. On the other hand OF functionalities can be used for changing the per hop behavior (PHB) for each traffic flow.

Challenges:

- To realize a comprehensive traffic monitoring and enforcement with only few measurement points the traffic has to be crossed out and transported to that measurement points
- Current SDN implementations like Open flow are lacking of prioritization functions for specific traffic flows
- The signaling between the points of presences within the network and the measurement points has to be defined and the additional network load has to be analyzed
- Interaction with PCRF and the mapping of 3GPP QCIs to the DiffServ classes and other markings which can be used for traffic enforcement on flow level is also an important challenge.

Added modules are measurement and manipulation probes:

- QoE probe at different locations (with different processing power): extraction of performance related parameters and QoS/QoE metadata as well as teeing out specific traffic for monitoring in datacenters

Links to functions within the Mobile Network Cloud

- SDM CTRL: allows controlling the use of monitoring resources, recuperating traffic or metadata for analysis.
- MME: for location aware monitoring
- PCRF: for using PCRF capabilities for traffic manipulation
- Applications: for running QoE calculation, Rules design and decision units and every virtualized function of the QoE monitoring and manipulation framework

Interfaces

- ISAAR needs access to the marked entities shown in the figure above to collect needed data for monitoring and to initiate the enforcement
- New Interfaces should only be designed, if the existing ones are not sufficient. Therefore in Figure 24 Interface 2: QMON & QENF is mentioned which allows a direct communication between the monitoring and enforcement probes and the virtualized functions within the data center

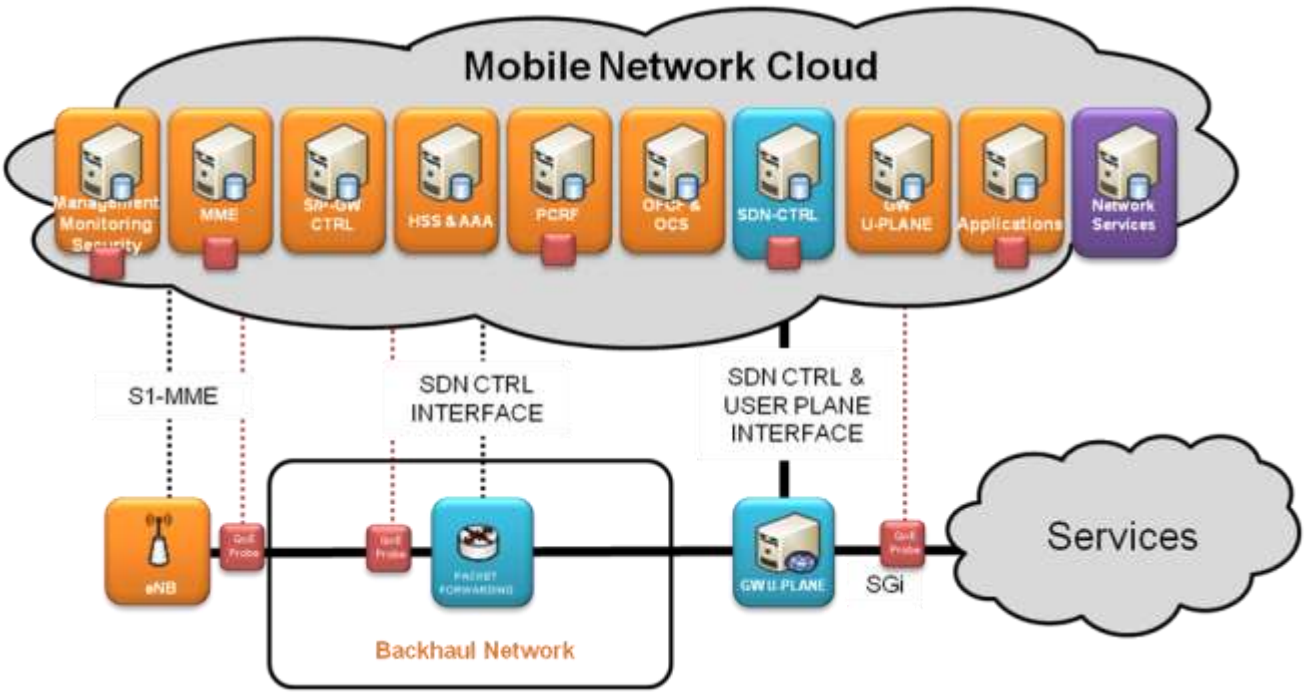

Figure 30. QoS/QoE enforcement in virtualized ISAAR concept

4.5.1.6 Joint traffic and cloud resource management for video traffic optimization with service chaining

The operator provided services have to become more efficient to handle competitive market and fast rollouts are required. In addition to provide those services there are many network applications such as firewalls, content filters, intrusion detection systems (IDS), deep packet inspection (DPI), network address translation (NAT), content caches, load balancers, wide-area network (WAN) accelerators, multimedia transcoders, logging/metering/charging/advanced charging applications, etc. already existing in today's networks. Such applications are generally referred to as middleboxes, because they are commonly executed along the traffic path, with their existence usually unknown by the end users. In fact, almost all traffic in mobile networks visit a pre-defined sequence of middleboxes en route to its destination today. Such a sequence is commonly referred to as a "service chain." Then, a service chain consists of a set of network services that are interconnected through the network to support an application such as VoIP, streaming video, e-mail, web browsing, etc. In other words, service chaining "orchestrates" the network flow for every offered application.

Creation of high performance service-chaining applications is essential to meet traffic demand, higher quality expectations from customers while reducing capital and operational expenses associated with their networks. To achieve this, only necessary middleboxes should be processed, should refrain processing unnecessary middleboxes. Detailed identification of each flow, and establishing an appropriate dynamic service-chain for that flow is the main problem.

The modifications of existing standard mechanisms been depicted in the following figure:

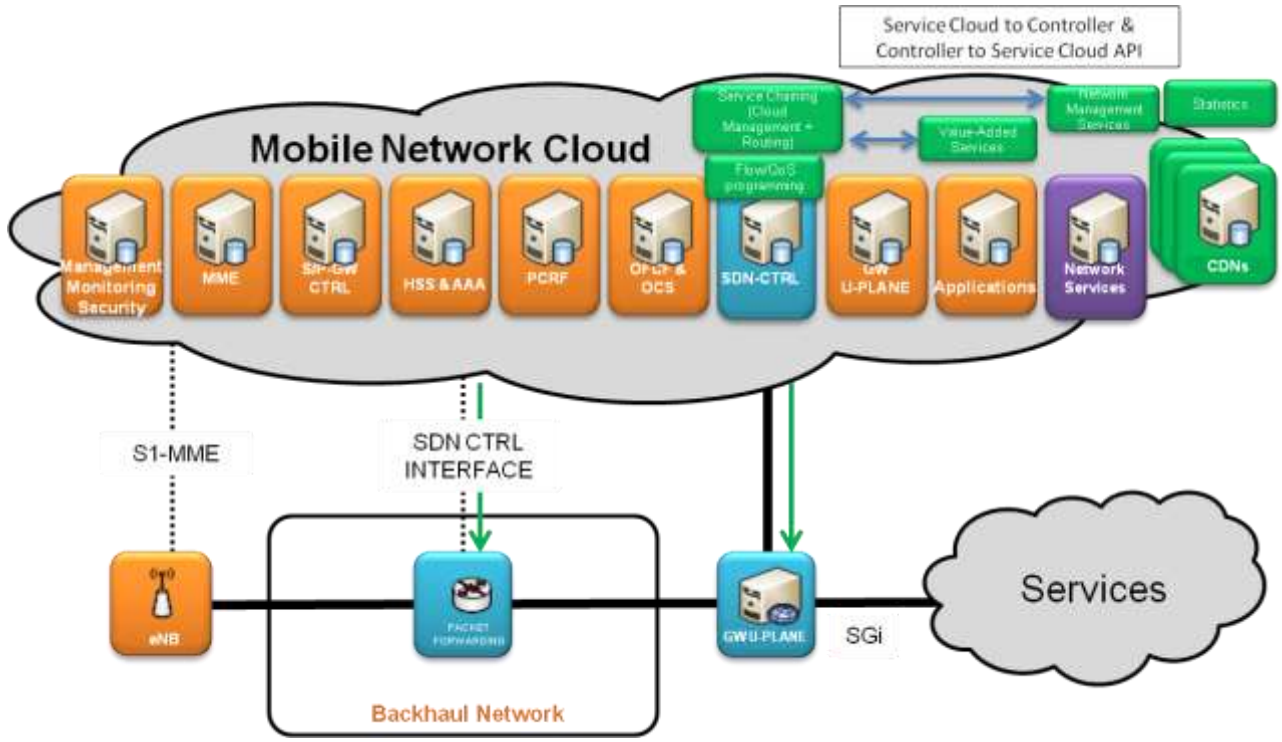

Figure 31. Joint traffic and cloud resource management for service chaining in SDMNs

There are challenges associated with configuring the network to bypass certain middleboxes for a given type of traffic flows. The existing solutions can be summarized as follows [21]:

- Single box running multiple services: In this approach all middleboxes are realized in a single router or gateway and new services are added with additional service cards to the node. This approach makes the integration of third party service appliances difficult. It also suffers from a scalability issue as the number of services and the aggregated bandwidth is limited by the router's capacity. An example to this approach is the Packet-Gateway (P-GW) in 3GPP networks.
- Statically configured service chains: In this approach, one or more static service chains are defined where each service is statically pre-configured to send traffic to the next service in its chain. This approach requires a large amount of service-specific configuration and is error prone, and is not flexible. It cannot support the steering at the granularity of per subscriber.
- Policy-based routing: In this approach a router furnished with policy-based routing (PBR) is used to route traffic to different services. Each time the traffic is returned to this router and then sent to the next service. This approach is not scalable as the router must be able to handle N times the incoming traffic line rate to support a chain with $N - 1$ services.
- Policy-aware switching layer: A policy-aware switching layer which explicitly forwards traffic through different sequences of middleboxes has been proposed in the literature recently. In this approach, each policy needs to be translated into a set of low level forwarding rules and be made available on all the relevant switches. The approach provides no explicit way to consolidate application and subscriber related rules and thus manual configuration is required to establish the set of low-level rules. The manual configuration is needed for each new flow and thus the approach is not scalable.

Given explosive traffic demand predictions, higher quality expectations from customers while a reduction in per user revenues, mobile network operators need to find ways of reducing capital and operational expenses associated with their networks and find new ways to monetize the afore-mentioned network services and applications. In this light, creation and rapid deployment and upgrading of cost-effective, high performance service-chaining applications are crucial. Thus, unnecessary processing of certain processing by one or more middleboxes should be avoided. Then the problem is detailed identification of each flow, and establishing an appropriate dynamic service-chain for that flow is

fundamentally important. For example, it may be desirable that upon detection of a long-lived video flow, DPI could be dynamically bypassed to save processing resources. Another example is that a transcoding appliance could be bypassed once the content type and codec have been identified. Scalable and dynamic steering capabilities could also enable new ways to monetize the middlebox deployment, where subscribers can buy QoE-aware network based services.

For the video delivery there is target to enable end-to-end traffic optimization within the SDN controller domain. From client apps to the cloud video application, the traffic path is built by the application in order to guaranty a known Service Level Agreement.

The principle is that the video application uses the SDN controller orchestration REST API in order to configure the wished path characteristics i.e. A End, B end, and a minimum bandwidth requirements. Latency and jitter optimizations could be also achieved, but are not mandatory in the current adaptive streaming solutions.

Another target is to extend SDN native horizontal scalability up to the SDMN cloud by the definition of Service Point Virtualization. Video application resources can be optimized based on its usage. Resource allocations or removals linked to network based flavor configuration can provide the ability to automatically adapt virtualized network components within the housing clouds (i.e. Load balancer, FW, routing…).

4.5.1.6.1 Research Approach

Service Provider SDN brings SDN capabilities to the network with policy-based and centralized control for improved network programmability and payload elasticity. NFV enables applications to share network resources intelligently and be orchestrated efficiently. It entails implementing network functions in software, meaning they can be instantiated from anywhere in the operator's network as well as in a consumer or enterprise customer premise. Finally, the cloud technology brings capabilities into the network with the flexibility and elasticity to deploy software applications wherever and whenever they are needed so that services and applications can be deployed, modified and withdrawn very rapidly.

In the future mobile service provider networks, the following steps may be used to provide an efficient, dynamic, scalable service chaining solution:

1. Following the SDN paradigm, centralize network management, analytics and configuration functionality to provide a single master that configures all networking devices.
2. Following the NFV paradigm, extract networking, analytics, security, QoE and billing services from the underlying hardware by creating service virtual machines so that independent scaling of each service becomes possible on the industry-standard x86 hardware in cloud.
3. Introduce a centralized SDMN orchestration controller that enables and monitors multiple networks, analytics, security, QoE and billing services to connect in series across devices within the network. A given service chain application is then simply a software program that runs on top of this controller. In other words, software is used to virtually insert services into the flow of network traffic. Networks can then dynamically respond to new information gathered about each flow as well as the network as a whole.
4. Develop service chaining applications that will optimize the usage of network and security hardware to deliver high performance.

For mobile service providers, service chaining has applications in two broad categories:

1. Network Application Services: Services that are necessary to maintain the security and health of the network as well as services that are used to log, meter and subsequently charge traffic. Examples include data leak prevention, fraud systems, authentication, billing, firewalls, DPI, etc.
2. Value-Added Services: Services that mobile service providers can develop for individual users to subscribe to. Examples include video optimizers, URL filters, parental controls, ad inserters, VPNs, VLANs, etc.

In an ideal solution, service chaining needs to be intelligent so that user identity, application, subscription, location, flow destination and context-aware dynamicity are possible. For example, not every response to a user request needs authentication, nor does every response need to be examined by the fraud system. Similarly, a user's flow need not visit a parental control middle box unless it is subscribed. Likewise, a streaming-video flow from a trusted source need not go through DPI upon detection. Therefore, evaluation, analysis, and further processing of inbound and outbound traffic coupled with the flexible, dynamic and scalable programmability of network and services control provide an attractive solution to network services chaining.

The service providers can use new service chaining techniques to generate revenue from applications. Until now third parties have delivered services such as video on demand over service provider networks, while service providers themselves have been unable to enter these markets, because of the complications involved in provisioning. But service chaining enables them to more efficiently embed applications and related services in the network itself, placing them at an advantage over the third party provider.

In this project work will concentrate on developing an OpenFlow-based SDMN architecture where the SDMN orchestration controller will control not only the flow forwarding hardware, but also the middleboxes that are distributed in geography in the Network-Enabled Cloud. A service chain application running on this controller will be developed that will jointly optimize:

- participating middlebox selection
- dynamic chain adaptation based on results from executed middleboxes
- corresponding end-to-end routing

for different user, subscription, application, geography, and context data so that user perceived quality is maximized, network congestion and middlebox overload probabilities are minimized.

Developing an SDMN service chaining orchestration to cover all possible scenarios is not realistic, therefore the focus in this project will be dedicated on the below described use cases.

4.5.1.6.1.1 Video delivery content optimization Use case: End-to-end video traffic optimized routing via SDN control

The optimized video route selection with the dynamic Flow programming, based on the current network status, provided by the Resource Management service. The scenario is divided in three main steps:

1. The End-User is connecting to a video delivery application servers to select a video to play on his/her terminal
2. The application asks the SDN interface to setup an optimized network path from one of the video content servers and the End-User's terminal
3. Video is played on the End-User's terminal, streamed by the selected delivery server

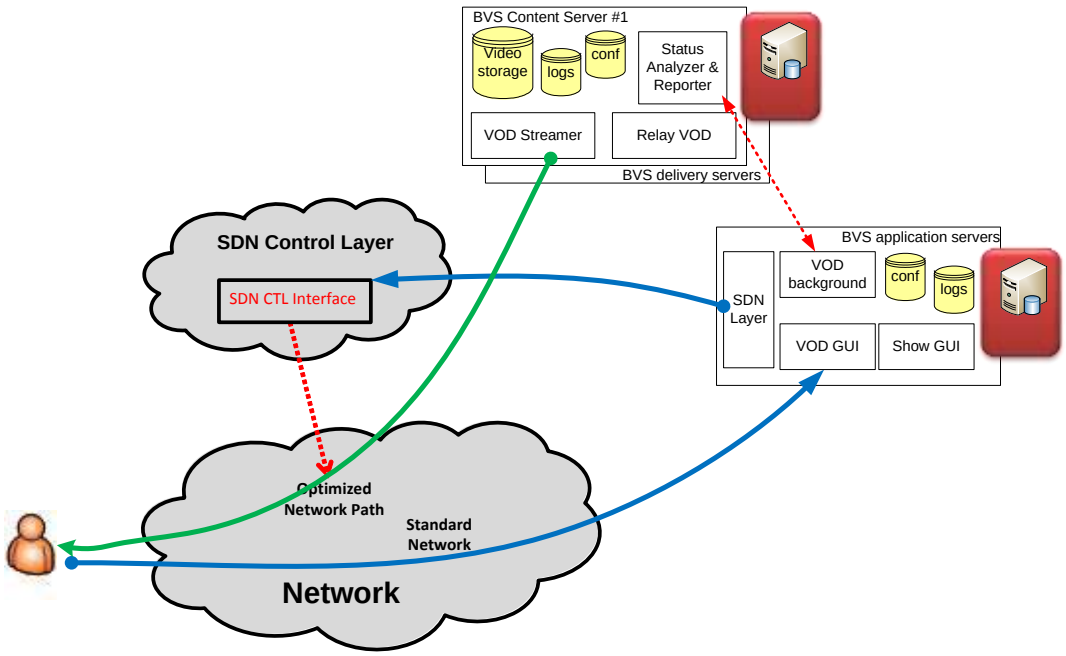

Figure 32. SDN optimized path use case, HTTP video streaming scenario

4.5.1.6.1.2 Service Chaining Use Case 1: Selecting video transcoder according to QoS requirement and routing efficiency.

Network providers may sell high quality video subscription to customers or may want to differentiate quality based on specific application (Youtube, Facebook etc.). Video transcoder in the service chain can be selected accordingly via the SDN controller.

4.5.1.6.1.3 Service Chaining Use Case 2: Ad insertion over video.

Network providers may want to insert ad over video according to user profile. SDN decides based on either user profile (subscription details, device capability) and/or network congestion map decides on a "service chain" for this flow. The different "video streaming" flows can be provided, one for a standard user and another for a premium user.

4.5.1.6.1.4 Service Chaining Use Case 3: Ad insertion over web content flow.

Network providers may want to insert ad over web content flow according to user interest. SDN decides on a "service chain" for this flow that contains DPI followed by Big Data Analytics that instigates Ad Insertion into the content. SDN pushes the necessary rules to the routers to facilitate the Service Chain.

4.5.1.6.2 Mapping to layered SDN model and reference points

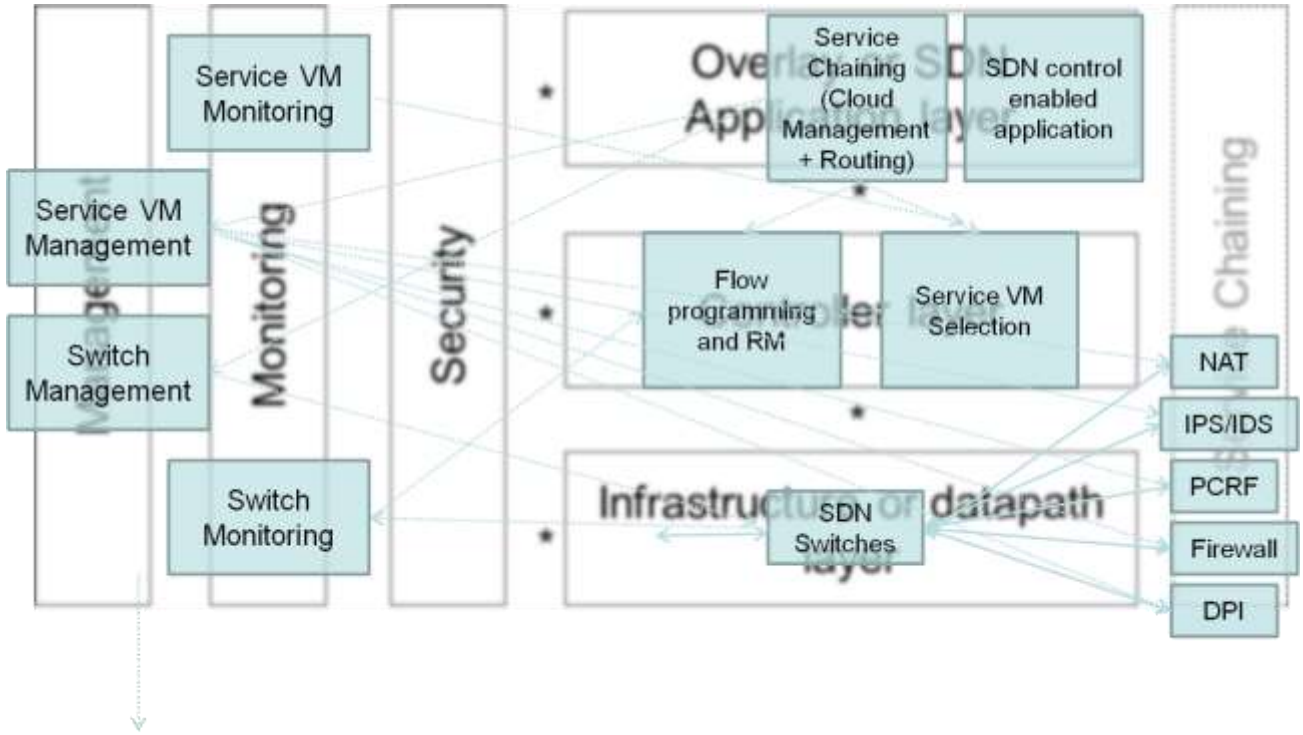

Figure 33. Architectural Model for Service Chaining and video delivery optimization in SDMNs

Use case	Touched Architecture layers	Notes

Service Chaining	1: 2: 3: 4: 5:	OFConfig / OVSDB capable OF switches, Service VMs Flow programming and Service VM Selection Service chaining application Switch management with OFConfig / OVSDB and Service VM management Resource monitoring for optimal routing calculation, VM Load monitoring for effective load balancing
SDN Advanced video delivery	1: 2: 3:	Flow programming, Resource Management service Video streaming path optimization SDN module
Virtualization and Telco cloud	3, 4, 5:	Network orchestration SDN layer

4.5.2 Network and cloud resource monitoring and management

Resource management and network resource availability awareness are further investigated. Physical and virtualized resources (managed by MNO and MNVO) will require to have updated information on available inventory and topology.

Resource models are required for describing the resource consumption of (virtual network) functions that are running as software processes (on VMs) on the servers within the cloud datacenters. The cloud resources are expressed in abstract units of processing power, memory consumption and switching capacity.

In this work package resource models are derived for mobile network functions as well as other functions (e.g. traffic monitoring, traffic manipulation, traffic caching). The models are used as important input for the design of the resource management methods.

4.5.2.1 Network resource availability awareness

Network resource availability awareness will be the foundation for resource management process, service provisioning and optimization traffic management, as well as security protection and monitoring process.

Resource availability awareness in SDMNs requires appropriate functions in SDN controller, SDN switches, present in the reference scenario figure, Figure 34. The scenario requires IP connectivity. Information regarding virtual and physical networks will be collected for status monitoring and to be presented in a GUI.

Related basic assumption: Assumption on availability of resources.

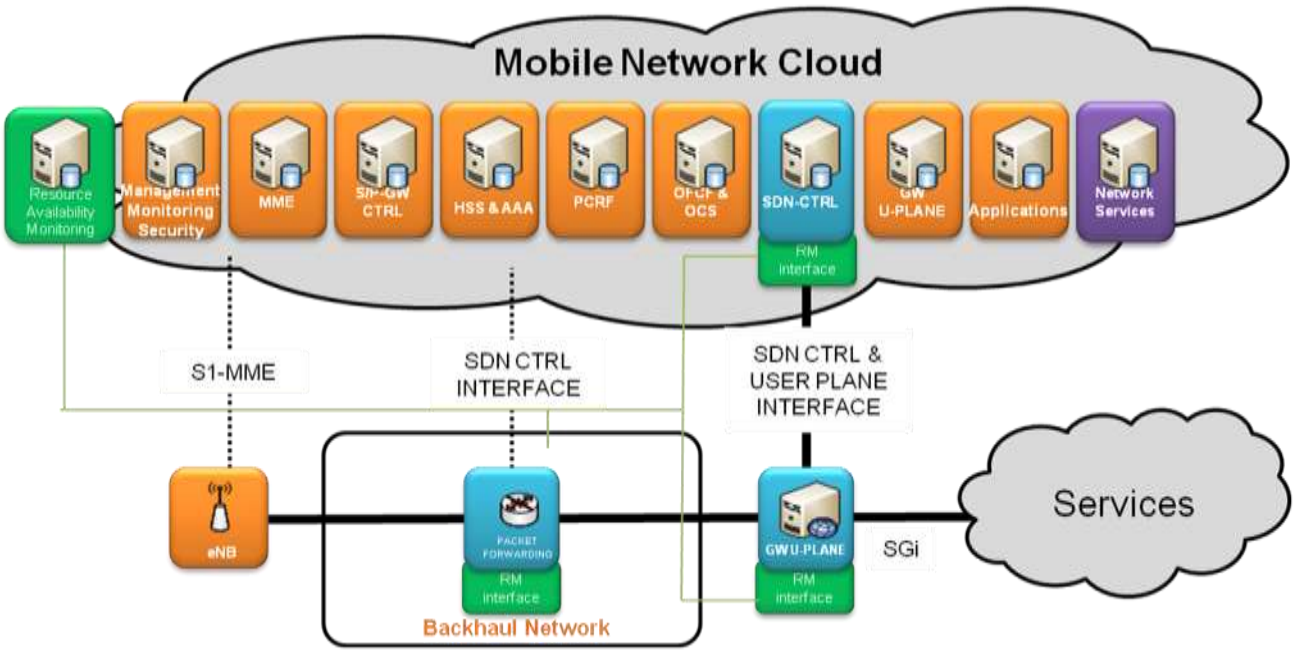

Figure 34. Resource availability awareness enhancements

4.5.2.1.1 Mapping to layered SDN model and reference points

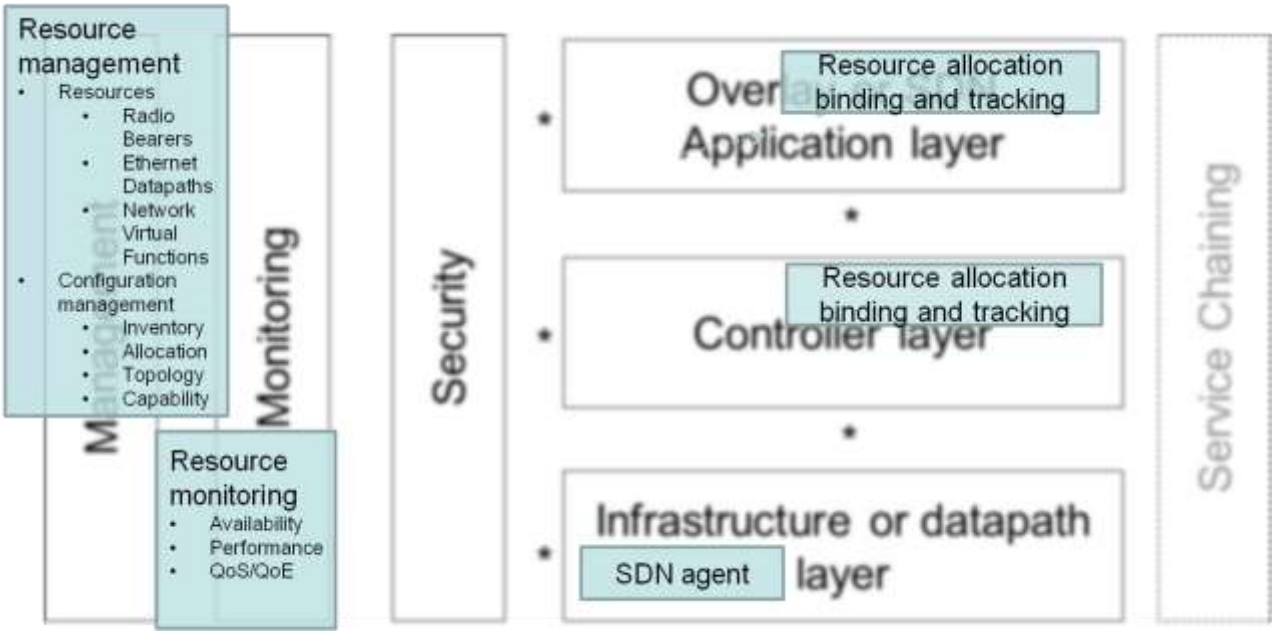

Figure 35. Resource availability management based on SDN

Considering different resource management Nextel will focus on network resource management, where at management and at monitoring level different network resources will be managed and monitored and at application, control and data path layer this network resources should be bound and tracked.

Use case	Reference points	Notes

Resource management based on SDN	2-1:	Resource provisioning/binding
	1-2:	Status and measurement logs
	2-3-2:	Network features
	2-4-2:	Resource provisioning/binding and inventory
	2-5:	Resource availability monitoring and state information
	5-4	Correlated information
	4-2/3	Policy provisioning

4.5.2.2 Resource management in the cloud

This topic is closely coupled with joint routing and resource management for optimal routing (described in Section 4.5.1.2), and joint traffic and cloud resource management for service chaining in SDMNs (described in Section 4.5.1.6).

The main focus areas are scalability, load balancing, multi-tenant platform support, and QoS aspects in the cloud infrastructure. Key performance indicators such as connection setup and teardown durations, finding the shortest or most efficient network path, latency, jitter, and packet loss are the main criteria for performance and quality of service (QoS) evaluations. The results and feedback will be used for method and algorithm optimizations and dynamic adaptation of virtual mobile networks, resource, policy, mobility, security and traffic management and monitoring are being defined and implemented.

Target is to adapt current network management platform in order to configure and control the different partner SDN test systems involved, using programmatic resource management and deployment. The management platform is fully horizontally scalable, both in the device management and event storage. The latest real-time BigData technologies are employed in order to cope with the immediateness and data volumes required. Being based on a high availability and scale out configuration, it will be able to cope with the most demanding programmatic configuration tasks, covering the most demanding usage scenarios.

There will be created a technical concept for the coordination of network and cloud resource management regarding the mobile core domain. Investigate the new methods / algorithms for the optimization of the resource consumption in the mobile core domain during the network operation. There will be implemented the algorithms and evaluation for different realistic network and traffic scenarios. Create models which describe the cloud resource demand (processing power, memory, and switching capacity) of selected virtual network functions.

Another target is to evaluate and verify these newly defined methods and algorithms by means of simulation and lab tests against realistic network and traffic scenarios. Our main focus areas are scalability, load balancing, multi-tenant platform support, and QoS aspects in the cloud infrastructure. Key performance indicators such as connection setup and teardown durations, finding the shortest or most efficient network path, latency, jitter, and packet loss are the main criteria for performance and quality of service (QoS) evaluations. Our results and feedback will be used for method and algorithm optimizations.

4.5.2.2.1 Mapping to layered SDN model and reference points

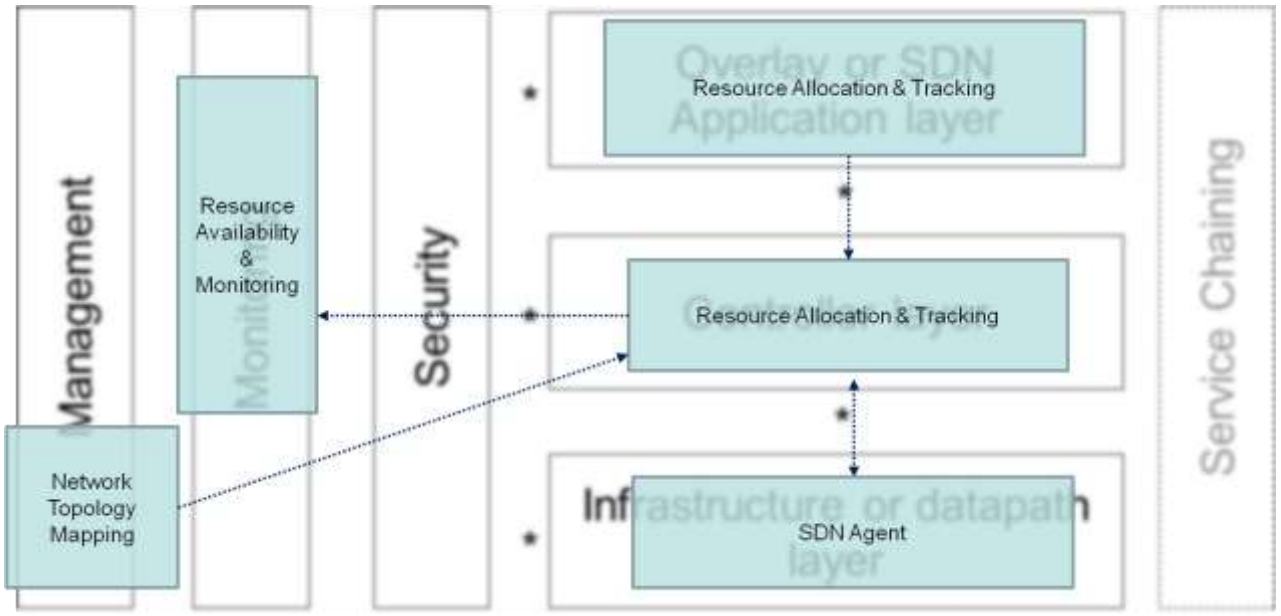

Figure 36. Cloud Infrastructure & Resource and Traffic Management

Architectural components needed are SDN controller, SDN Switches, Applications, including Resource & Traffic management algorithms, and Network Services, including network topology.

Interfaces needed are REST APIs for resource allocation, flow statistics and network topology information gathering and SDN controller flow management.

4.5.2.3 Coordinated network and cloud resource management optimization

The resource management is considered from a cloud-based SDMN operator point of view. The optimization target is as follows: The resource consumption should be minimized considering the traffic demand and the QoS and reliability requirements. Therefore, the most important tasks of the resource management are to adapt the resources to the changing traffic demand, to provide sufficient resources to compensate failures, and, in the case of network virtualization, to allocate sufficient resources to virtual networks and to dynamically adapt these resources to the current traffic demand.

In the following, a split of the resource management into a network resource and cloud resource management is assumed. Hereby network resources comprise the paths (in the transport network layer) with their assigned capacities, the transport nodes (forwarding nodes) and their interfaces as well as the mobile network components which are not yet realized as virtual network functions. The cloud resource management comprises the mapping of the virtual network functions (which are running on VMs) to the corresponding cloud datacenters as well as the respective allocation of cloud resources (in terms of processing power, memory and switching capacity).

Through the split into network and cloud resource management different grades of coordination between both management systems might be possible. For example one scenario could be the full integration of both systems by one network operator, while in another scenario only the network resource management stays within the operator domain, whereas the virtual network functions are outsourced to a public cloud.

One of the most important aspects of cloud computing for mobile networks is virtualized resource management within the cloud and it comprises the mapping of the virtual network functions, that are running on VMs, to the corresponding cloud data centers as well as the respective allocation of cloud resources to address the end-to-end networking needs for processing power, memory and switching capacity. Resource and traffic management are also closely linked to each other. The resource management decides which and how many virtual network function instances are provided at which cloud data center location and determines the optimal traffic routing within the transport network to reach these functions. Security management within the SDMN addresses the issues of resource availability, policy enforcement, negotiation and management.

All of these functions and services have to be analyzed and verified to understand the performance and quality of service (QoS) related implications and improvements. This work has no direct impact on new interfaces and

functionalities, but analyzes the benefits and costs of the application of cloud-technologies and network function virtualization.

4.6 Security management in the virtualized and SDN controlled mobile network

4.6.1 Description of the security challenge and principles

Virtual and software defined network techniques make it easier to modify and configure network functions using centralized controllers and separating the control functionalities to the cloud, making it not necessary to intervene directly on different network elements. This makes controllers a critical element in the network that needs to be secured, guaranteeing high availability at all times. SDN aims at simplifying and enhancing network control and management, while making easy to implement new applications. The potential to use SDN and SDMN to construct easier, cheaper, better security solution will be studied.

Controllers become a security concern and where they are located and who has access to them needs to be managed correctly. Communications between the controllers and network elements needs to be assured by encryption techniques (e.g., SSL) and the keys need to be managed securely. But are these techniques sufficient to assure high availability? Denial-of-Service (DoS) attacks are difficult to detect and counter. Controllers could be vulnerable to these types of attacks and guaranteeing that they are available at all times becomes difficult. Furthermore, every change and access needs to be monitored and audited for troubleshooting and forensics; and this is more complicated in virtual environments where visibility is often reduced.

Existing security solutions (e.g., SIEM, IDS, IPS, firewalls...) need to be adapted and correctly controlled since they were meant mostly for physical and not virtual systems and boundaries. The lack of visibility and controls on internal virtual networks created makes many security applications virtually (no pun intended) blind. The impact of virtualization on these technologies needs to be determined. For instance some of the subjects that need to be addressed are the following: Security applications need to be able to monitor virtual connections; Virtualization can help isolate systems but can also be used to make malicious systems that are difficult to detect; Virtualization creates boundaries that could be breached by exploiting vulnerabilities and bugs in the virtualization code (e.g., hypervisors); Centralized controllers make security management easier but become single points of failure; Whole systems are actually files that can more easily be stolen; Virtualization facilitates changes making it necessary that security applications keep up with this new dynamicity; etc.

SIEM (Security Information and Event Management) will be necessary in order to gain security and context awareness. If an incident happens, the system should be able to determine the source, recover and protect itself against it in the future. It should be verified that everything that comes out of the system is logged. Due to the fact that central managers have control over the network, it is necessary to log every change and treat it accordingly in a management solution. Log analysis and event correlation in SDN will fast become a big data issue. Tools need to be developed that can address all the forensics and compliance requirements.

4.6.2 Security management positioning and research scenarios

The consolidated view of the research activity in the security is presented in figure below. The interfaces and contact points are indexed with the numbers, in order to indicate which of the interfaces are impacted or important in the studied scenario.

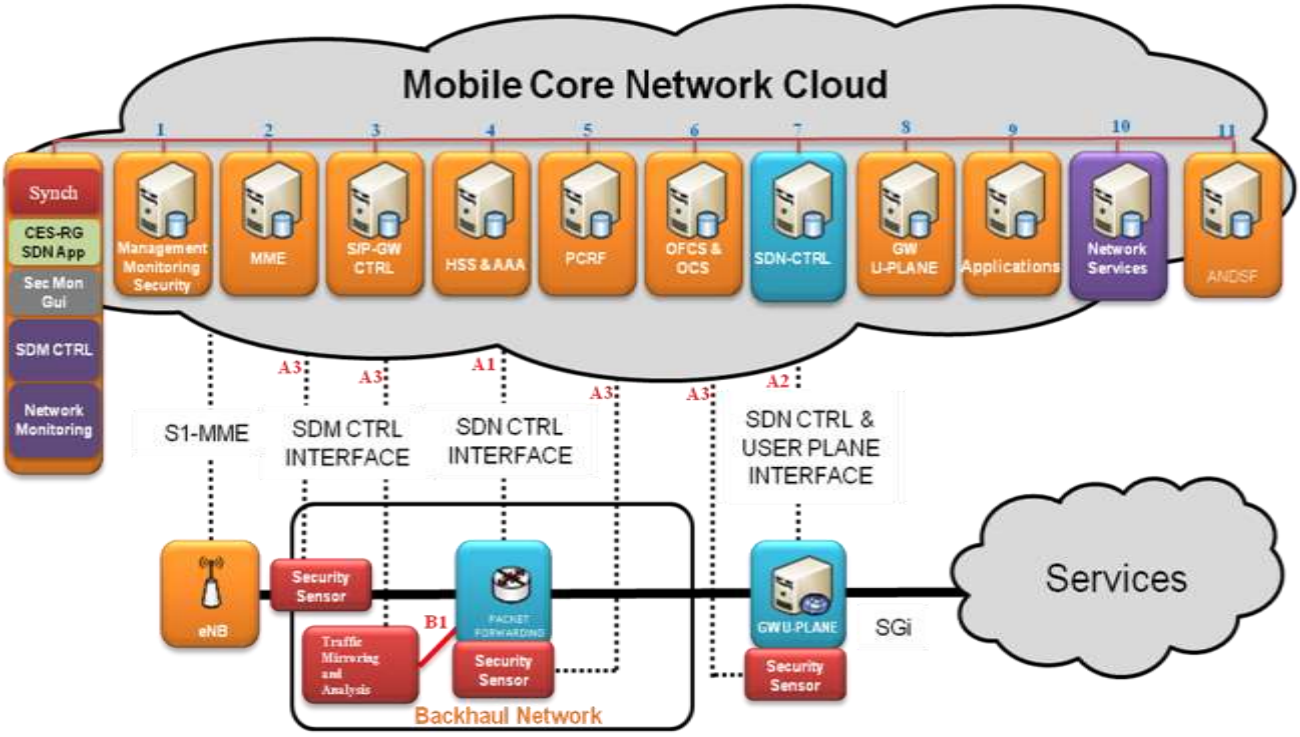

Figure 37. Security management in the virtualized and SDN controlled mobile network

New additional modules:

- Security sensor: Active or passive monitoring probe for the detection of security and behavior related information (e.g., security properties and attacks) and mitigation (e.g., filtering). It can be installed on the Network Elements (virtualized or not) or in Network TAPs; and analyze virtual or physical network interfaces.
- SDM CTRL: new module or extension of SDN CTRL to allow control of monitoring function (e.g., management of network monitoring appliances, traffic mirroring, traffic load balancing and aggregation); accepts requests from network functions and applications…). SDM CTRLs are distributed following either a peer-to-peer or hierarchical model. They interact with the Management/monitoring/security function and act as distributed analysis or decision points for the defined security policies (security SLAs).
- Network monitoring: virtualization of monitoring function (part of the traffic analysis moved to the cloud)
- Traffic Mirroring and Analysis: passive backhaul traffic monitoring required by different network functions

New additional interfaces:

- SDM CTRL INTERFACE: allows controlling the use of monitoring resources, recuperating traffic or metadata for analysis. Can be implemented as JSON API.
- Monitoring requests and status: allows applications and network functions to send requests for monitoring based information and allows monitoring functions to send status and recommendations. Can be based on a publish/subscribe model and implemented as JSON API.

Scenario 1: Synchronizing the network security and network traffic management require all the concerned entities (e.g. HSS & AAA and Management Monitoring and Security, MME, SDN controller) to cooperate.

There is a synchronizing entity (SDN controller functional unit) to be implemented in the SDN controller.

For 3GPP and non-3GPP networks (e.g. Wi-Fi) to interwork, the ANDSF functionalities are required in the cloud.

The Interfaces between functionalities used or required in this concept are 1, 7, 11, A1, A2 as shown in the consolidate figure.

Scenario 2: CES (Customer Edge Switching) [22] security focuses on interconnecting end users within the same network or across different networks securely, leveraging host-defined policies for access control in Control Plane and commercial off-the-shelf (COTS) equipment for the Data Plane. CES implements gateway functionality for the private network acting as a cooperative firewall with other CES devices. CES enables end-to-end communications between

hosts located in different private networks. The access to these private hosts is granted by policy. CES also implements security mechanisms for preventing DoS attacks and source address spoofing [23].

The target is to increase cooperation between edge nodes and operators in order to achieve global trust. As a result, the blocks of more interest on are:

- SDN Controller: Fundamental framework for running the different SW instances (SDN applications) and connecting them to each other.
- CES & RG App: The Customer Edge Switching (CES) and Realm Gateway (RG) act as a connection broker for creating UE's data connections to other hosts located in different CES nodes or public networks such as the Internet.
- PCRF App: Used for accessing subscriber databases, retrieval of user-defined policies and specialized charging functions.
- Additional Network services: Such as Application Layer Gateways SDN Apps, DNS resolvers, DHCP server, etc.
- HSS link: It can be used for verification and authorization of end-users or CES nodes.
- MME link: As an interface for communicating the registration of new UEs to CES/RG.

CES communication with MME and SDN controller: The CES allows incoming and outgoing traffic based on the information received from the MME after new user is attached and authenticated. The communication between CES and MME allows CES to set the right flow entries in the SDN data plane.

With the CES security relates to scenario of consolidating the LTE virtual network elements into single component (eMME) that integrates SDN functionality. Target is to reduce tunneling overhead in data plane and increase throughput. Enable use of the commercial-off-the-shelf switching network components, without mobile specific functionality. Dynamic allocation of resources in the cloud across data centers according to the traffic demands.

Interfaces used (required): 2, 4, 5, 7, 9, 10.

Blocks Added: CES-RG SDN App and Additional Network Services in the cloud.

Scenario 3: The three level security management approach (GUI and policy management, sensor deployment and agent deployment) to solve the security monitoring and management challenges.

Network security functions (AuthN, AuthZ, Cipher, Filtering…) shall be deployed as software modules, in order to enforce network security policies defined in the admin access and network Security Policy Administration Point tool (SPAP).

Interfaces used (required): 1, 7, A1, A2, B1

Blocks Added: Security Monitoring GUI, Traffic and Mirroring Analysis in the Backhaul network, and Sensors in Datapath and GW-U-Plane.

Scenario 4: Software Defined Monitoring (SDM) controller allows controlling SDM enabled appliances to optimize the metadata extraction to what is needed by the network functions. The SDM needs to interact with the SDN to set up network taps or traffic mirroring for the monitoring and security appliances. The Management System and SON will manage the SDM controllers and define what policies need to be applied, depending on the state of the network provided by the monitoring applications.

The PCRF instructs the Traffic Detection Function (TDF, DPI device) using Sd to look for specific application flows. The TDF will alert the PCRF using Sd when these flows are detected. Then the PCRF can instruct the PCEF to install a change rule using Gx. For this the TDF informs the SDM to set up the detection and the Monitoring application will notify the TDF.

QoS, resource and security monitoring applications will analyze the metadata provided by the monitoring and security appliances. In turn these applications will inform the network functions (e.g., management, SON, TDF) that require information on the state of the network and application flows. Eventually, they will inform the SDM controllers if they act as decision points.

Interfaces used (required): 1, 7, A3; Optional: B1, 5, 9, 10.

Blocks Added: Network Monitoring and SDM CTRL in the Cloud, Security Sensors and Traffic & Mirroring Analysis in the Backhaul network.

4.6.2.1 Mapping scenario 1. to layered SDN model and reference points

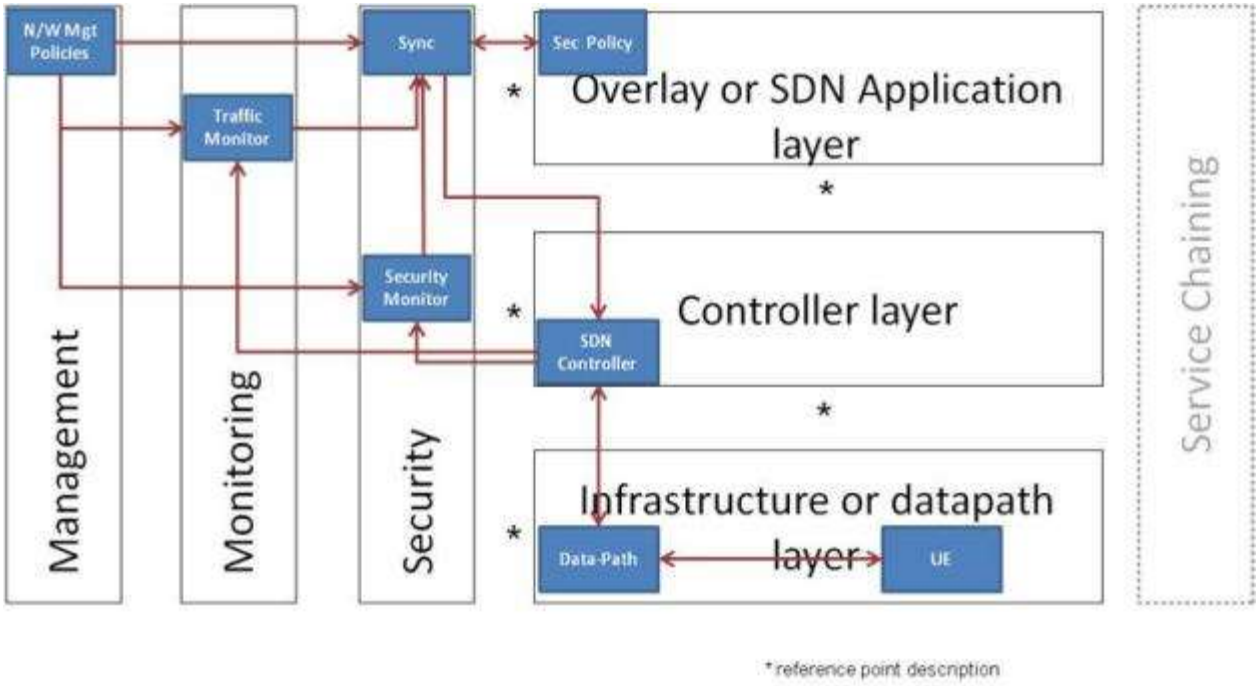

Figure 38. Synchronized Network Security and Traffic Management

- Application Layer:
  - Security Policies are implemented on top of the controller and as part of the application layer.
- Controller Layer:
  - The controller takes the decision based on the input from the policy layer through security Sync and converges the policy to action in the data-path layer. It also provides the security and traffic stats from data-path layer to security and traffic monitors respectively.
- Security Layer:
  - The Synch in the security layer takes input from the traffic monitor about the topology, mobility etc. Similarly it takes the security stats from the security layer. The synch synchronizes the network security and the network traffic. The Synch then provides the instruction to the controller layer
- Management:
  - The traffic and security are checked to be intact and according to the policies defined in the management layer.

4.6.2.2 Mapping scenario 3. to layered SDN model and reference points

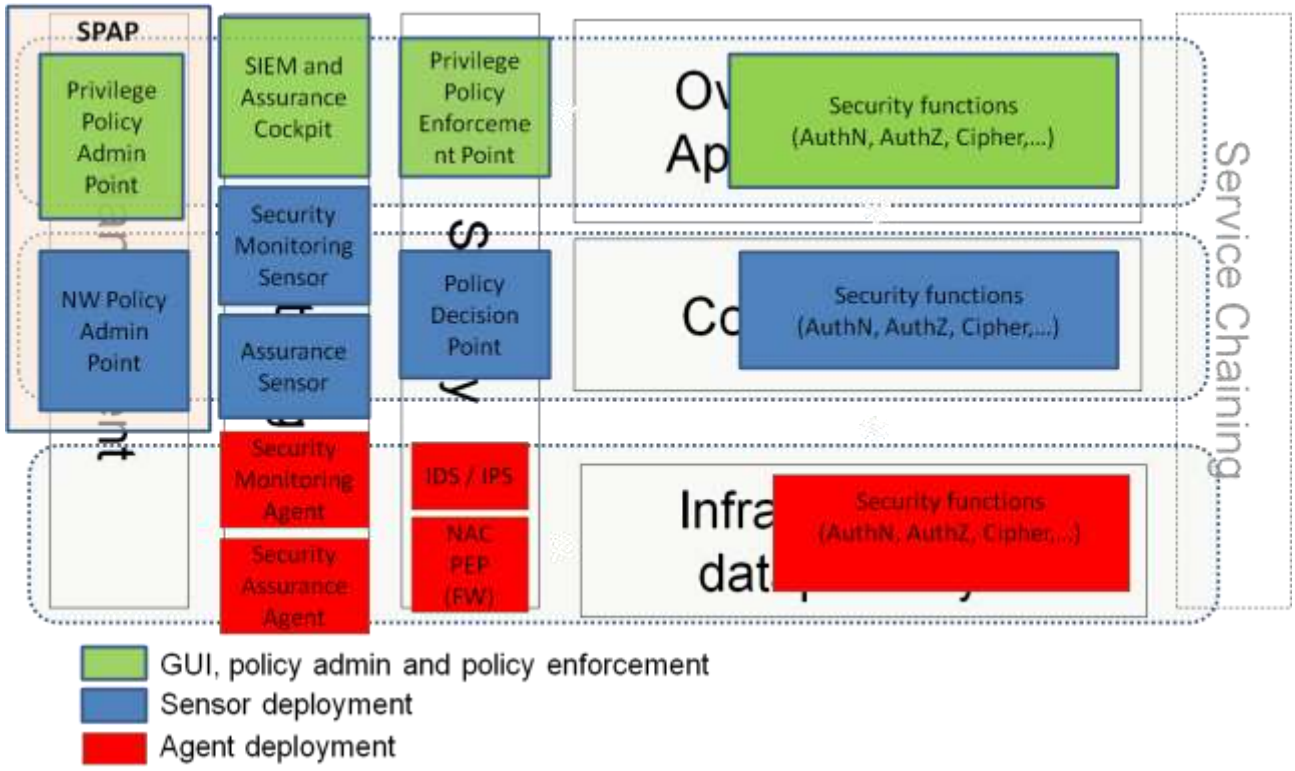

Figure 39. Three level security management architecture proposal

Management

- SPAP: Security Policy Administration Point is the tool to storage the security policies to be deployed in all three network data and control layers. Network security policies and privileged admin access control shall be distinguished.

Monitoring

1. The following three level functionalities are deployed for GUI, agents and sensors (SIEM) for collection and correlation of information in order to gain security and assurance status awareness.
2. Security monitoring and assurance sensor to track network activity as well as security related issues, gathering different sources of information, normalizing as well as consolidating them.
3. Security monitoring and assurance agents to perform specific checks and base measurements to be escalated to central processing unit (SIEM), through sensors.

Security

- This layer provides additional security policy enforcement modules to support the security functions execution deployed in all three control and data layers. The examples are:
    - Identity and access management auxiliary modules to support admin access control management
    - Network firewalling
    - Network intrusion detection systems

Main use case(s):

- Security optimization in SDN based mobile networks
- Security management
- Synchronized Network Security and Traffic Management

Related basic assumption(s):

- Assumption on managing the security. How to include security functions of physical and virtual elements and interfaces.
- Assumption on Security and Traffic Management Synchronization: Coordinate security with traffic management.
- Assumption on Delay-Security Constraints: Optimize security setup to reduce delays

Main targeted benefit(s):

Scalable network security monitoring targeted to the needs of applications and network functions. This can be referred to respectively as Application Centric Network Security and Software Defined Security Monitoring. Security in SDN networks will have to solve the issues of availability, policy enforcement, mitigation, management and enhance overall security along the different parts of the architecture.

5. Consolidated architecture view

This document includes the mapping of the multiple research contributions from different partners into SDMN architecture. This mapping results in multiple components that utilize SDN as basic technology to deploy the required functionality on areas such as EPC user plane control, transport, load balancing, security, monitoring, QoE, resource optimization, etc. Some of the researched components have a considerable impact to the mobile network architecture and some are rather optimizations to the existing functionalities and services or SDN and cloud required support functionalities (like monitoring and security).

Therefore, a consolidated architecture would include elements in different parts of the current architecture that could be grouped based on whether they are deployed in the user plane/transport network or they are cloud residing control functions as depicted in the following figure.

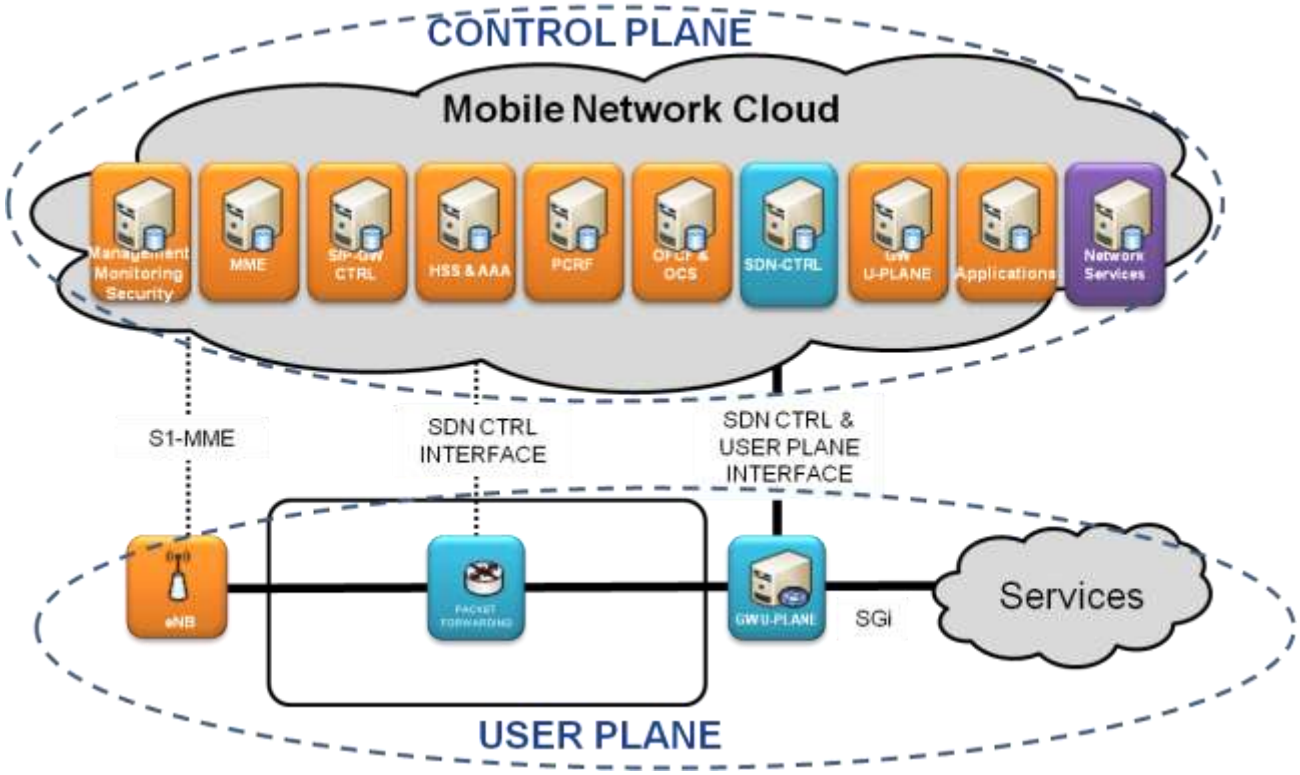

Figure 40. Generic mobile network architecture evolution view and division to control / user planes

The components deployed on the user plane consist of the components that have to be deployed in the physical packet transport network. These components are linked to the cloud control elements envisaged running on the cloud.

The components deployed as cloud control functions consists of QoE monitoring, resource management and orchestration, service slicing among others which functionality is implemented and deployed on the cloud.

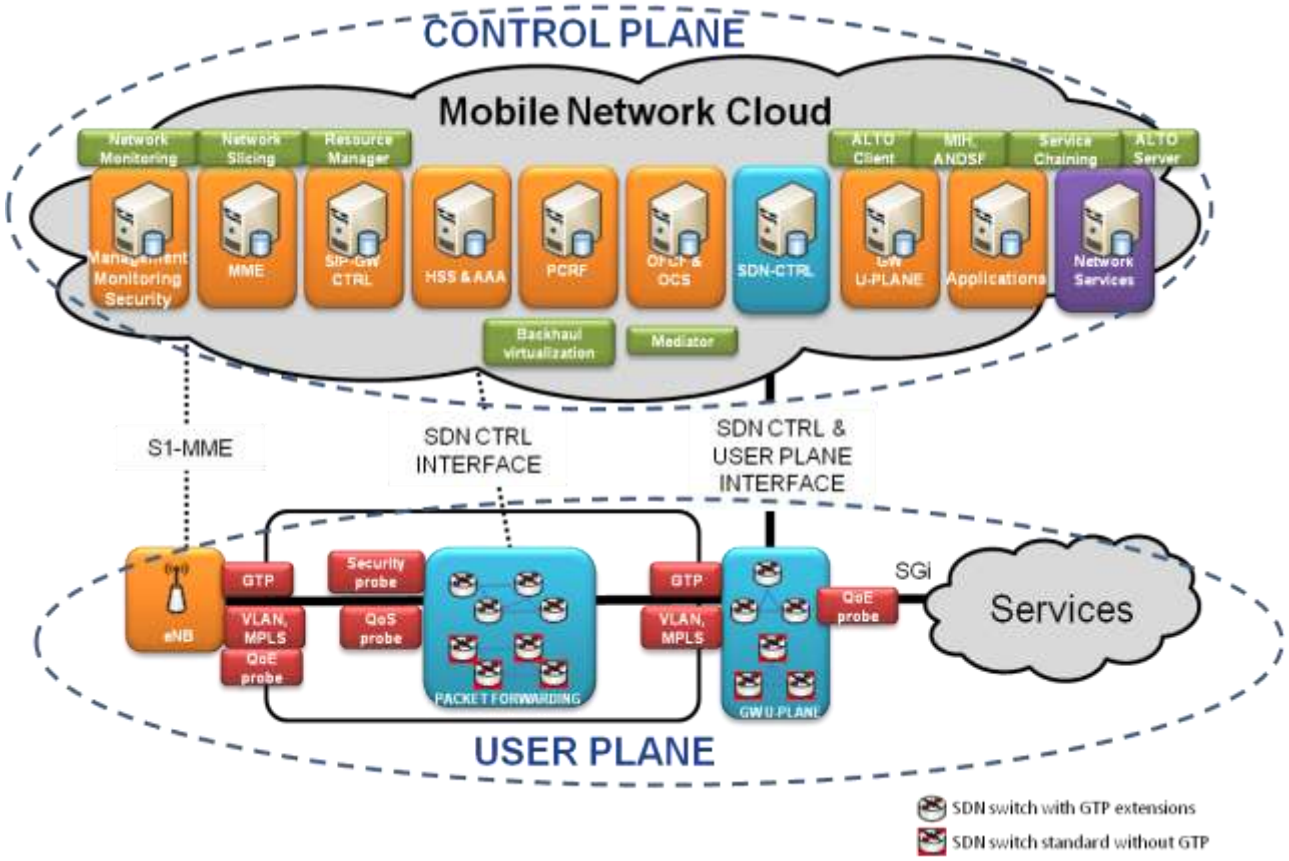

Figure 41. Consolidated SIGMONA research additions to the architecture

5.1 Consolidated architecture and basic assumptions

The proposed consolidated architecture is addressing the basic assumptions and requirements identified for future mobile networks as shown in the following table.

Basic assumptions	Consolidated architecture solution
2.1 Migration: Compatibility with the legacy systems vs. Clean-slate deployment	EPC network elements maintain most of the current 3GPP interfaces to enable migration from legacy mobile networks. At least a clear evolution path to new architecture needs to be provided for both user and control plane.
2.2 Virtualization and running network functions in cloud	At least the EPC control plane network functionalities run on the cloud as shown Fig. 40. The user plane is enhanced with SDN functionality.
2.3 Resilience: Fault tolerant SDN based mobile networks	Usage of SDN switches for the backhaul user plane with the network monitoring provides the required reliability.
2.4 QoS provision in virtualized mobile core networks in SDN-based forwarding paths	Network monitoring together with resource management and control (like ALTO) functionality will provision the needed QoS.
2.5 SDN based mobility management versus 3GPP and MIP mobility	The usage of SDN or MIP provides the needed mobility management functions.
2.6 Locator and identity assignment to UEs in SDMNs	MIP and integration of SDN with legacy 3GPP network elements.
2.7 Security and Traffic Management Synchronization	Resource manager and ALTO functionality provides

coordination	the needed traffic management to optimize the usage of user plane resources.
2.8 Optimize security setup constraints to reduce delays	Security probes and the security and network monitoring functions provide the needed security setup.
2.9 Managing the security functions of physical and virtual elements and interfaces	Security probes will enable the security at different interfaces in the network system.
2.10 Security cooperation between edge nodes and operators towards global trust	CES and other proposed security methods provide the needed framework to setup security and trust cooperation between edge nodes.
2.11 Regulation for more competition vs. Let the market forces decide	Regulation activity will be analyzed for the proposed consolidated architecture.
2.12 Privacy and trust regulation driving force	CES and other proposed security methods provide the needed security ad trust cooperation.
2.13 Network monitoring adapted to network virtualization	Network monitoring provides the continuous monitoring of virtualized EPC functions as well as information about congestion in user plane.
2.14 Service provisioning and optimization orchestrator entities	SDN based applications will enable the required management and orchestration of network elements both user plane transport elements but also control elements. The orchestration can be built on top of the SDN north bound interface.
2.15 Availability of resources	Resource management and ALTO provide the needed allocation of resources mainly for efficient usage of user plane transport resources.
2.16 Cost reduction impact of LTE network virtualization	The deployment of EPC network functions on the cloud benefits from virtualization and commodity servers.

Table 1. Consolidated architecture solutions for the basic assumptions

5.2 Architecture options in business deployment point of view

The different architecture options have been analyzed and prioritized by a survey to partners, in order to understand the most essential scenarios for techno-economical modeling. Not all of these are in the direct research scope of any partners in this project, but might be interesting in architecture evolution evaluation sense. This analysis includes not only the new architecture scenario comparisons to the base line, but also network service focus, and network ownership considerations.

This section describes the different scenarios identified from techno-economic point of view to be validated and confirmed whether those scenarios could be deployed with the proposed consolidated architecture. Following is the list of those scenarios to be considered:

1. The mobile manufactures split the network elements defined in 3GPP standardization. This first scenario considers the split of those elements into control and forwarding part but keeping proprietary interface in that separation.
2. The interface in the split is based on the open standards to allow different vendors to control and forwarding part (i.e. SDN mobile network)
3. The control part is implemented on Virtual Machines (VM) in cloud environment so it follow traditional 3GPP network elements with virtualization and an open split based on SDN
4. The control elements that follow 3GPP standards can be deployed individually or combined resembling virtualized functions following NFV architecture.

The above listed models can be implemented with the proposed consolidated architecture since it facilitates the split of control and user plane based on SDN technologies. Moreover, the usage of SDN can be done with open standards such as OpenFlow, Forces, Netmap, etc. that fulfill the needs of some operators for vendor swapping but SDN can be

supported with proprietary interfaces complementing open standards.These may provide the required differentiation for vendor specific features. Furthermore, the SDMN architecture allows the virtualization of the control elements and allows vendors to implement new functionality and features that will provide competitive advantage. Therefore, the proposed SDMN consolidated architecture fulfils all the models identified from economical point of view.

6. Conclusions

The proposed Software Defined Mobile Network (SDMN) architecture options help transform the current rigid and disparate mobile networks into scalable and dynamic ecosystems. SDN is considered the enabling technology for the SDMN making it the future paradigm for obtaining more flexible networks that can dynamically adapt to the needs of operators, content providers and service providers. Network horizontal scalability linked to cloud computing abilities promises a drastic change and simplification to the operators' networks. There is no doubt that new business models could emerge and motivate changes, but technical challenges are still heavy.

The proposed architecture, where the mobile network user plane is based on SDN technology, facilitates the migration by maintaining a part of the legacy network functions. The usage of virtualization and further split of control and data forwarding allows more efficient usage of network resources and improved security features running on the cloud. The cloud based SDMN architecture allows the core network to adapt to the service requirements dynamically. Different end user applications require different properties, like latency, capacity, optimum mobile tunnel termination location etc. These various requirements can be served better with an evolved architecture that supports an evolution path with respect to legacy systems. The centralized control plane in the cloud enables optimal usage of resources and use of the emerging cloud platform native services. The SDN controlled user plane (mobile network and transport) provides potential for optimizations of the topologically distributed forwarding elements. Moreover, extending the SDN principles to the radio network can further enhance the optimizations and would allow obtaining the mobile network architecture with end-to-end flexibility.

The traffic management proposals include macroscopic- and microscopic-level solutions. Macroscopic-level solutions target efficient routing, endpoint selection, service chaining and offloading. Microscopic-level solutions are focusing on QoS/QoE provision. An important aspect of these solutions is that they will enable higher dynamicity, flexibility for the operator for traffic management and policy control, than in legacy mobile network architectures due to the programmable network concept. For service and content providers, SDN techniques facilitate the provisioning of network services in a deterministic, dynamic, and scalable manner that is vendor neutral and based on the open standards. The software programmability enables agile and automated network configuration and traffic management. Plug-and-play ability is becoming ever more important for incremental network growth, as in the case of micro-cell base stations that are being deployed all around on walls and lamp posts.

In the resource management area, focus is on network resource availability awareness, which is the foundation for any resource management process. New coordinated network and cloud resource management algorithms are needed, dealing with the adaptation of network resources to the changing traffic demand, provision of sufficient resources to compensate failures, and, in the case of network virtualization, allocation of sufficient resources to virtual networks and to dynamically adapt these resources to the current traffic demand. Cloud resource management algorithms require the definition of new models, which describe the cloud resource demands of selected virtual network functions and the main focus areas are scalability, load balancing, multi-tenant platform support and QoS aspects in the cloud infrastructure.

Network monitoring architecture and functions need to be adapted to the requirements and constraints of SDMN. The network monitoring should be aware of the virtualized networks and provide statistics for the routing applications sitting on top of the SDMN controller. Applications can both manage the SDMN controller for optimal routing and the switches through the Management, Monitoring and Security interfaces, resulting in application controlled security and quality. The cloud-based SDMN architecture allows a more flexible QoS/QoE monitoring compared to the traditional architecture, as individual flows can be redirected to measurement probes that are realized in SW functions on the VMs in the cloud. Thus the performance of the monitoring system can be easily adapted according to the current requirements without the need to upgrade the hardware elements.

SDN enables highly reactive security monitoring, security analysis and response systems to facilitate network forensics, security policy alteration and security service insertion through the centralized control plane and programmable network nodes. However, both programmability and central control introduce new security vulnerabilities. An important risk is that, by facilitating software control, security breaches could affect legitimate traffic. Hence it is necessary to synchronize network security with the network traffic and for that we propose an architecture that allows this synchronization.

3GPP specifications recommend multiple IP tunneling options and IP mobility management schemes for handling IP mobility of users. We can divide the proposed mobility management schemes into post-distributed mobility management and SDN-based mobility management categories. The former covers the evolution of existing (distributed) mobility management solutions for the NFV-SDN architecture, which should in general be prepared for coordination with resource orchestrator, handling of SDN control plane. The latter category involves solutions, which break with the standardized tunneling concepts and apply SDN-technologies, such as OpenFlow, to coordinate the forwarding plane. Both categories may be supported by common control plane elements, such as ANDSF or MIH information service; furthermore the functional elements of the protocols should be prepared for network function virtualization.

Operators will profit from solutions integrating the proposed modifications by: Allowing the virtualization of at least part of the monitoring tasks; Improving the operators' visibility to their network status from both security and performance perspectives; Allowing them to deal with new vulnerabilities introduced by SDN; and reducing CAPEX and OPEX, thanks to SDN enabled flexibility, capacity control and automation. SDN brings new methods to monitor and manage the network, which is translated into better controls for improved end user experience. However, optimization methods have to be applied for proper allocation and placement of cloud resources (e.g. for processing, storage and forwarding) to benefit from this flexibility to obtain cost savings. During the course of the SIGMONA project a detailed techno-economic analysis is being carried out, to precisely quantify the CAPEX/OPEX savings of the new Cloud-based SDMN architecture.

7. References

[1] 3GPP TS 23.401, General Packet Radio Service (GPRS) enhancements for E-UTRAN access, V13.0.0, 2014-09

[2] General information about the ETSI ISG NFV: http://www.etsi.org/technologies-clusters/technologies/nfv

[3] ETSI ISG NFV Architectural Framework, ETSI GS NFV 002 (V1.1.1, 2013-0): http://www.etsi.org/deliver/etsi_gs/NFV/001_099/002/01.01.01_60/gs_NFV002v010101p.pdf

[4] ETSI ISG NFV open area for sharing the draft documents: http://docbox.etsi.org/ISG/NFV/Open/Latest_Drafts/

[5] OpenStack Community Welcome Guide, Revision 11, released 24. Jun. 2014: http://www.openstack.org/assets/welcome-guide/OpenStackWelcomeGuide.pdf

[6] 3GPP releases: www.3gpp.org/specifications/67-releases

[7] IETF BOF (Birds of a Feather) meeting content: http://tools.ietf.org/bof/trac/wiki

[8] E. Haleplidis, Ed., K. Pentikousis, Ed., S. Denazis, J. Hadi Salim, D. Meyer, O. Koufopavlou, "SDN Layers and Architecture Terminology", http://www.ietf.org/id/draft-irtf-sdnrg-layer-terminology-04.txt

[9] Wiki page of Software Defined Networking Research Group (SDNRG), http://trac.tools.ietf.org/group/irtf/trac/wiki/sdnrg

[10] Open Networking Foundation (ONF) White paper, Software-Defined Networking: The New Norm for Networks, April 13, 2012: https://www.opennetworking.org/sdn-resources/sdn-library/whitepapers

[11] J. Heinonen, et. al: "Dynamic Tunnel Switching for SDN-Based Cellular Core Networks", All Things Cellular workshop, August 22, 2014, Chicago, USA (co-located with ACM SIGCOMM 2014)

[12] N. Varis, J. Manner, J. Heinonen: "A Layer-2 Approach for Mobility and Transport in the Mobile Backhaul", August 23-25, 2011, ITST 2011, Saint-Petersburg, Russia

[13] Z. Faigl, L. Bokor, P. M. Neves, K. Daoud, P. Herbelin: "Evaluation of two integrated signaling schemes for the Ultra Flat Architecture using SIP, IEEE 802.21, and HIP/PMIP protocols," Computer Network – The International Journal of Computer and Telecommunications Networking, vol. 55, pp. 1560–1575, 2011.

[14] Z. Faigl: "Performance Analysis of Signalling Overhead in Host Identity Protocol-based Secure Mobile Networks: Ultra Flat Architecture or End-to-End Signalling", Wireless Networks, 2014.

[15] Z. Faigl, M. Telek: "Modeling the signaling overhead in Host-Identity Protocol-based secure mobile architectures," Journal of Industrial Management and Optimization, vol. 11, no. 3, 2015, in press.

[16] L. Bokor, Z. Faigl, S. Imre: "Survey and Evaluation of Advanced Mobility Management Schemes in the Host Identity Layer", International Journal of Wireless Networks and Broadband Technologies 3: pp. 34-59. 2014.

[17] Z. Faigl, Zs. Szabó, R. Schulcz, "Application-layer traffic optimization in software-defined mobile networks: a proof-of-concept implementation", 16th International Telecommunications Network Strategy and Planning Symposium (Networks 2014), Madeira Island, Portugal, Sep. 17-19, 2014.

[18] R. Alimi (Ed.), R. Penno (Ed.), Y. Yang (Ed.), "Application-Layer Traffic Optimization (ALTO) Protocol", IETF RFC 7285, Sep. 2014.

[19] 3GPP TS 23.203, "Policy and charging control architecture (Release 12)", v12.2.0, September 2013.

[20] M. Eckert, T. Knoll: "ISAAR (Internet Service quality Assessment and Automatic Reaction) a QoE Monitoring and Enforcement Framework for Internet Services in Mobile Networks", 4th International Conference, MONAMI 2012, Hamburg, Germany, September 24-26, 2012.

[21] A. Gudipati, D. Perry, L. E. Li, S. Katti: "SoftRAN, Software Defined Radio Access Network", Hot Topics in Software Defined Networking workshop, August 16, 2013, Hong Kong (USA (co-located with ACM SIGCOMM 2013): http://www.stanford.edu/~skatti/pubs/hotsdn13-softran.pdf

[22] R. Kantola, “Implementing Trust-to-Trust with Customer Edge Switching,” in Proc. 24th IEEE International Conference on Advanced Information Networking and Applications Workshops (AINA 2010), pp. 1092-1099, Perth, Australia, 20-23 Apr. 2010.

[23] J. Llorente, R. Kantola, N. Beijar and P. Leppäaho, “Implementing NAT Traversal with Private Realm Gateway”, in Proc. IEEE International Conference on Communications (ICC 2013), pp. 2174-2179, Budapest, Hungary, 9-13 June 2013.