




# *Small cEllS coordinAtion for Multi-tenancy and Edge services*

# Grant Agreement No.671596

Topic: H2020-2014-ICT-14

*Advanced 5G Network Infrastructure for the Future Internet*

Research and Innovation Action

## Deliverable D2.2

# Overall System Architecture and Interfaces

| | |
|---|---|
| Document Number: | H2020-5GPPP-GA No.671596/*WP2/D2.2/30.03.2016* |
| Contractual Date of Delivery: | *31.03.2016* |
| Editor: | Dr. Ioannis Giannoulakis, Dr. Tassos Kourtis - NCSRD |
| Work-package: | *WP2* |
| Distribution / Type: | Public (PU) / Report (R) |
| Version: | *1.0* |
| Total Number of Pages: | 71 |
| File: | *SESAME_Deliverable 2.2_v1.0_Final* |

## Abstract

The SESAME EU-funded project targets innovations around three central elements in the wider 5G context: (i) The placement of network intelligence and applications in the network edge through Network Functions Virtualization (NFV) and Edge Cloud Computing; (ii) the substantial evolution of the Small Cell (SC) concept, already mainstream in 4G but expected to deliver its full potential in the challenging high dense 5G scenarios, *and*; (iii) the consolidation of multi-tenancy in communications infrastructures, allowing several operators/service providers to engage in new sharing models of both access capacity and edge computing capabilities.

The present Deliverable *D2.2* is the second specification document of the SESAME project, which presents the current outcomes of *Task 2.2,* with regard to the design and specification of the SESAME overall system architecture. The aim has been to produce an architectural proposal which fulfils most key Small Cell- and NFV-*related* requirements, implements all use cases defined in the prior Deliverable *D2.1* and, *at the same time*, allows implementation in a relatively short time-frame with reasonable technical complexity.

The first step is the survey of currently proposed architectures, including industry initiatives (LTE Small Cell Networks & Edge Cloud Concepts), actual research projects (e.g., Mobile Cloud Networking, T-NOVA, UNIFY), comparison to other 5G initiatives, as well as the current specifications and trends in ETSI NFV ISG.

Considering the use cases and requirements laid out in *D2.1* as well as the state-of-the-art (SOTA) in similar initiatives, including standardisation trends, a high-level overall architecture is proposed, which encompasses all the component entities of the full SESAME system. The architecture is structured by "merging" the two main SESAME worlds, namely the Small Cells featuring the LTE network and the NFV environment.

Simultaneously, the main concepts of SESAME dedicated to the implementation of Cloud-Enabled Small Cells (CESCs) are also presented and discussed, so as to deliver edge cloud computing in a multi-service ecosystem. The SESAME concept is analysed at the component/sub-system and system level. At the component/sub-system level, we detail our plan to deploy multi-operator enabled small cells, enhanced with a virtualised execution platform, while at the system level, we present the envisaged architecture to manage and control the cloud-enabled small cell infrastructure.

Compliance with the current Small Cell Forum, 5G and the ETSI NFV ISG vision is sought at all stages of system design, especially concerning the terminology and the main architectural blocks, while several extensions are also proposed.

# Version History

| Version | Date | Comments, Changes, Status | Authors, contributors, reviewers |
|---|---|---|---|
| 0.1 | 26.02.2016 | Initial draft by NCSRD | I. Giannoulakis |
| 0.1a | 03.03.2016 | Initial contributions by NCSRD | I. Giannoulakis |
| 0.2 | 16.03.2016 | Contributions by i2CAT | C. Ruiz, P. Khodashenas |
| 0.2a | 16.03.2016 | Contributions by EHU | J. O. Fajardo, F. Liberal |
| 0.2b | 16.03.2016 | Contributions by ITL | A. Albanese, C. Meani |
| 0.2c | 16.03.2016 | Contributions by ZHAW | I. Trajkovska, A. Edmonds |
| 0.2d | 16.03.2016 | Contributions by UPC | J. Perez-Romero |
| 0.2e | 16.03.2016 | Contributions by FLE | M. Wilson |
| 0.2f | 17.03.2016 | Contributions by UoB | V. Vasilakis |
| 0.2g | 17.03.2016 | Contributions by ATN | D. Munaretto |
| 0.2h | 17.03.2016 | Contributions by UNIS | K. Moessner |
| 0.2i | 17.03.2016 | Contributions by ORION | E. Kafetzakis |
| 0.2j | 17.03.2016 | Contributions by OTE | E. Sfakianakis, I. Papafili, M. Belesioti, I. Chochliouros |
| 0.2k | 17.03.2016 | Contributions by VOSYS | P. Bliznakov |
| 0.3 | 18.03.2016 | Merged and circulated for review | I. Giannoulakis |
| 0.4 | 21.03.2016 | Review by ATOS | J. García |
| 0.4 | 21.03.2016 | Review by UPC | J. Perez-Romero |
| 0.5 | 22.03.2016 | Review by NCSRD and submission to the coordinator | I. Giannoulakis |
| 0.6 | 23.03.2016 | Contributions by OTE | M. Belesioti, I. Chochliouros |
| 0.7 | 23.03.2016 | Contributions by SMNET | A. Dardamanis |
| 0.8 | 24.03.2016 | Pre-Final version of the Deliverable, fully reviewed by OTE | M. Belesioti |
| 1.0 | 29.03.2016 | Final version of the Deliverable, fully reviewed by OTE, submitted to the Commission | I. Chochliouros |

# Contributors

| First Name | Last Name | Partner | Email |
|---|---|---|---|
| Ioannis | Giannoulakis | NCRSD | giannoul@iit.demokritos.gr |
| Anastasios | Kourtis | NCSRD | kourtis@iit.demokritos.gr |
| Daniele | Munaretto | ATN | daniele.munaretto@athonet.com |
| Javier | García | ATOS | javier.garcial@atos.net |
| Jose Oscar | Fajardo | EHU | joseoscar.fajardo@ehu.es |
| Fidel | Liberal | EHU | fidel.liberal@ehu.eus |
| Mick | Wilson | FLE | Mick.Wilson@uk.fujitsu.com |
| Pouria | Sayyad Khodashenas | i2CAT | pouria.khodashenas@i2cat.net |
| Cristina | Ruiz | i2CAT | cristina.ruiz@i2cat.net |
| Jordi | Ferrer Riera | i2CAT | jordi.ferrer@i2cat.net |
| August | Betzler | i2CAT | august.betzler@i2cat.net |
| Daniel | Camps Mur | i2CAT | daniel.camps@i2cat.net |
| Eduard | Escalona | i2CAT | eduard.escalona@i2cat.net |
| Antonino | Albanese | ITL | antonino.albanese@italtel.com |
| Claudio | Meani | ITL | claudio.meani@italtel.com |
| Emmanouil | Kafetzakis | ORION | mkafetz@orioninnovations.gr |
| Ioannis | Chochliouros | OTE | ichochliouros@oteresearch.gr |
| Evangelos | Sfakianakis | OTE | esfak@oteresearch.gr |
| Ioanna | Papafili | OTE | iopapafi@oteresearch.gr |
| Maria | Belesioti | OTE | mbelesioti@oteresearch.gr |
| Vassilios | Vassilakis | UoB | V.Vasilakis@brighton.ac.uk |
| Jordi | Perez-Romero | UPC | jorperez@tsc.upc.edu |
| Oriol | Sallent | UPC | sallent@tsc.upc.edu |
| Ferran | Casadevall | UPC | ferranc@tsc.upc.edu |
| Pavel | Bliznakov | VOSYS | p.bliznakov@virtualopensystems.com |
| Irena | Trajkovska | ZHAW | traj@zhaw.ch |
| Vincenzo | Pii | ZHAW | piiv@zhaw.ch |
| Marcello | Coppola | STM | marcello.coppola@st.com |
| Athanassios | Dardamanis | SMNET | ADardamanis@smartnet.gr |

# Glossary

| Acronym | Explanation |
|---|---|
| 3GPP | 3rd Generation Partnership Project |
| 4G | 4th Generation (of mobile communications) |
| 5G | 5th Generation (of mobile communications) |
| 5G-PPP | 5th Generation-Public Private Partnership |
| ANR | Automatic Neighbour Relationships |
| API | Application Programming Interface |
| APP | Application |
| APPS | Applications |
| ARM | Advanced RISC Machine |
| BGP | Border Gateway Protocol |
| BS | Base Station |
| BSS | Business Support System |
| BYOD | Bring Your Own Device |
| CA | Controller Adapter |
| CC | Cloud Computing |
| CC | Cloud Controller |
| CCO | Capacity Optimization |
| CDL | Choreography Description Language |
| CESC | Cloud Enabled Small Cell |
| CESCM | CESC Manager |
| CLI | Command-Line interface |
| CN | Core Network |
| CoMP | Coordinated Multi-Point |
| COTS | Commercial off-the-shelf |
| CP | Control Plane |
| CPE | Customer Premises Equipment |
| CPU | Central Processing Unit |
| cSON | centralized Self-Organizing Network |
| CWMP | CPE WAN Management Protocol |
| DC | Data Centre |
| DHCP | Dynamic Host Configuration Protocol |
| DL | Downloading |
| DM | Device Manager |
| DNS | Domain Name System |
| DoS | Denial-of-Service |
| DoW | Description of Work |
| DP | Data Plane |
| DSA | Dynamic Spectrum Access |
| dSON | Distributed Self-Organizing Network |
| DSP | Digital Signal Processor |
| E2E | end-to-end |
| eDSA | Extended Dynamic Spectrum Access |
| EM | Element Manager |
| EMS | Element Management System |
| EN | European Norm |
| eNB | Evolved Node B |
| EPC | Evolved Packet Core |

| | |
|---|---|
| ETSI | European Telecommunications Standards Institute |
| EU | European Union |
| E-UTRA | Evolved Universal Terrestrial Radio Access |
| E-UTRAN | Evolved Universal Terrestrial Radio Access Network |
| FCAPS | Fault, Configuration, Accounting, Performance and Service |
| FP | Framework Programme |
| FP7 | 7th Framework Programme |
| FPGA | Field Programmable Gate Array |
| GA | Grant Agreement |
| GPU | Graphics Processing Units |
| GRE | Generic Routing Encapsulation |
| GS | Group Specification |
| GSM | Evolved Universal Terrestrial Radio Access Network |
| GSMA | Global System for Mobile Communications |
| GTP | GPRS Tunnelling Protocol |
| GW | Gateway |
| GWCN | Gateway Core Network |
| HARQ | Hybrid Automatic Repeat reQuest |
| HeNB | Home eNB |
| HetNet | Heterogeneous Network |
| HOT | Heat Orchestration Template |
| HSS | Home Subscriber System |
| HW | Hardware |
| IaaS | Infrastructure as a Service |
| ICIC | Inter-Cell Interference Coordination |
| ICT | Information and Communication Technology |
| ID, Id, id | Identifier |
| IEEE | Institute of Electrical and Electronics Engineers |
| IFIP | International Federation for Information Processing |
| IoT | Internet of Things |
| IP | Integrated Project |
| IP | Internet Protocol |
| IPR | Intellectual Property Right |
| ISG | Industry Specification Group |
| IT | Information Technology |
| ITG | Infrastructure Template Graph |
| IVM | Infrastructure Virtualisation and Management |
| KPI | Key Performance Indicator |
| KQI | Key Quality Indicator |
| KVM | Kernel-*based* Virtual Machine |
| LAA | License assisted Access |
| LAN | Local Area Network |
| Light DC | Light Data Centre |
| LIPA | Local IP Access |
| LTE | Long Term Evolution |
| MAC | Medium Access Control |
| MANO | Management and Orchestration |
| MCN | Mobile Cloud Networking |
| MEC | Mobile-Edge Computing |
| MF | Monitoring Function |

| | |
|---|---|
| MLB | Mobility Load Balancing |
| MIMO | Multiple-Input Multiple-Output |
| MME | Mobility Management Entity |
| MNO | Mobile Network Operator |
| MOCN | Multi-Operator Core Network |
| MP | Management Protocol |
| MPLS | Multiprotocol Label Switching |
| MRO | Mobility Robustness Optimisation |
| µS | micro server |
| MVNO | Mobile Virtual Network Operator |
| NAS | Non Access Stratum |
| NBI | Northbound Interface |
| NC | Network Control |
| NE | Network Element |
| NEO | Network Operation |
| NetSoft | Network Softwarization |
| NF | Network Function |
| NF-FG | Network Functions Forwarding Graph |
| NF-IB | Network Function Information Base |
| NFS | Network Functions System |
| NFV | Network Functions Virtualization |
| NFVI | NFV Infrastructure |
| NFVO | NFV Orchestrator |
| NSH | Network Service Header |
| NO | Network Operator |
| NM | Network Manager |
| NMS | Network Management System |
| NS | Network Service |
| NSH | Network Service Header |
| NSP | Network Service Provider |
| OAM | Operation Administration and Management |
| OASIS | Organization for the Advancement of Structured Information Standards |
| OCCI | Open Cloud Computing Interface |
| ODL | Open Daylight |
| OF | Open Flow |
| OL | Orchestration Layer |
| OVS | Open Virtual Switch |
| OSS | Operation Support System |
| P2P, p2p | Peer-to-Peer |
| PAP | Policy Administration Point |
| PCI | Peripheral Component Interconnect |
| PDCP | Packet Data Convergence Protocol |
| PDP | Policy Decision Point |
| PGW | Packet Data Network Gateway |
| PEP | Policy Enforcement Point |
| PHY | Physical layer |
| PIP | Policy Information Point |
| PLMN | Public Land Mobile Network |
| PM | Performance Management |
| PNF | Physical Network Function |

| | |
|---|---|
| PoC | Proof of Concept |
| PoP | Point of presence |
| PPP | Public Private Partnership |
| Qemu, QEMU | Quick Emulator |
| QoE | Quality of Experience |
| QoS | Quality of Service |
| RAN | Radio Access Network |
| RB | Radio Bearers |
| REM | Radio Environment Measurement |
| RF | Radio Frequency |
| RIA | Research and Innovation Action |
| RISC | Reduced Instruction Set Computing |
| RLC | Radio Link Control |
| RM | Resource Management |
| RNC | Radio Network Control |
| RO | Resource Orchestrator |
| RRC | Radio Resource Control |
| RRM | Radio Resource Management |
| R-TP | Radio Transmission Point |
| SC | Small Cell |
| SCaaS | Small Cells as a Service |
| SCF | Small Cell Forum |
| SCN | Small Cell Network |
| SCNO | Small Cell Network Operator |
| SDN | Software Defined Network |
| SDO | Standards Development Organisation |
| SFC | Service Function Chain |
| SGW | Server Gateway |
| SGW | Serving Gateway |
| SI | Service Interface |
| SIC | Service Instance Component |
| SIPTO | Selected Internet IP Traffic Offload |
| SIPTO@LN | SIPTO at the local network |
| SLA | Service Level Agreement |
| SM | Service Manager |
| SO | Service Orchestrator |
| SoC | System on Chip |
| SO-I | SO implementation |
| SOM | SO Management |
| SON | Self-Organizing Network |
| SOTA | State-of-the-Art |
| SP | Service Provider |
| STG | Service Template Graph |
| SW | Software |
| TM | Telecommunications Management |
| TMF | Telecommunications Management Forum |
| TNM | Transport Network Management |
| TOSCA | Topology and Orchestration Specification for Cloud Applications |
| TP | Transmission Point |
| TR | Technical Report |

| | |
|---|---|
| TRAN | Terrestrial Radio Access Network |
| TS | Technical Specification |
| TS | Technical Standard |
| TV | Television |
| UDP | User Datagram Protocol |
| UE | User Equipment |
| UI | User Interface |
| UL | Uploading |
| vDPI | virtual Deep Packet Inspection |
| vEPC | virtual EPC |
| VIM | Virtualised Infrastructure Management |
| VLAN | Virtual Local Area Network |
| VM | Virtual Machine |
| VM2VM | Virtual Machine-to-Virtual Machine |
| VNF | Virtual Network Function |
| VNFM | VNF Manager |
| VS | Virtual Switch |
| VSCNO | Virtual Small Cell Network Operators |
| W3C | World Wide Web Consortium |
| WAN | Wide Area Network |
| WG | Working Group |
| WP | Work Package |
| WS | Web Services |
| WS-BPEL | Web Services - Business Process Execution Language |
| WS-CDL | Web Services - Choreography Description Language |

# Table of Contents

**ABSTRACT** ..... **2**
**VERSION HISTORY** ..... **3**
**CONTRIBUTORS** ..... **4**
**GLOSSARY** ..... **5**
**TABLE OF CONTENTS** ..... **10**
**LIST OF FIGURES** ..... **12**
**LIST OF TABLES** ..... **13**
**1. INTRODUCTION** ..... **14**
1.1. DELIVERABLE OUTLINE ..... 14
1.2. DEFINITIONS OF TERMS AND SESAME CONCEPTS ..... 15
**2. SURVEY ON LTE SMALL CELL NETWORKS & EDGE CLOUD CONCEPTS** ..... **17**
2.1. SMALL CELL FORUM ..... 17
*2.1.1. Small cell virtualization: Functional split* ..... *17*
*2.1.2. Impact of virtualization on the management architecture* ..... *19*
*2.1.3. Multi-operator small cells* ..... *21*
*2.1.4. Self-Organizing Networks* ..... *22*
2.2. OVERVIEW OF ETSI MEC ..... 25
2.3. OVERVIEW OF ETSI NFV ISG AND ETSI MANO ..... 26
**3. SURVEY ON ARCHITECTURES OF PROJECTS RELEVANT TO SESAME** ..... **28**
3.1. H2020 5G-PPP PROJECTS RELEVANT TO SESAME ..... 28
3.2. FP7 T-NOVA ARCHITECTURE ..... 30
3.3. FP7 MCN ARCHITECTURE ..... 32
3.4. FP7 UNIFY ARCHITECTURE ..... 35
3.5. 5G DESIGN PRINCIPLES IN 3GPP SA2 ..... 38
**4. SESAME OVERALL ARCHITECTURE** ..... **40**
4.1. KEY FEATURES OF THE SESAME SYSTEM AND ARCHITECTURE PRINCIPLES ..... 40
4.2. HIGH-LEVEL DESCRIPTION OF SESAME MAIN ARCHITECTURAL ENTITIES ..... 42
*4.2.1. CESC* ..... *42*
*4.2.2. Overview of the "Light Data Centre" concept proposed by SESAME - Envisaged system implementation* ..... *44*
*4.2.3. CESC clustering* ..... *45*
*4.2.4. Infrastructure Virtualization* ..... *46*
*4.2.5. CESC Manager* ..... *46*
4.3. COMPLIANCE AND MAPPING TO 3GPP, ETSI NFV AND SCF ..... 48
*4.3.1. Relationships with SCF* ..... *48*
*4.3.2. Compliance and Mapping to ETSI NFV ISG* ..... *48*
*4.3.3. Compliance and Mapping to 3GPP* ..... *51*
4.4. NFV ORCHESTRATION ..... 53
*4.4.1. Interfaces* ..... *54*
*4.4.2. Service function chain lifecycle management* ..... *54*
*4.4.3. Service function chain monitoring and scaling* ..... *55*
4.5. SESAME ENVIRONMENT FOR INTRODUCING MOBILE EDGE SERVICES ..... 56
**5. SERVICE LIFECYCLE AND SEQUENCE OF INTERACTIONS** ..... **60**
5.1. SERVICE LIFECYCLE ..... 60

5.2. SERVICE COMPOSITION ..... 60
5.3. DEFINITION OF "SESAME NETWORK SERVICE", COMPOSITION AND OFFERING THROUGH THE CESCM DASHBOARD ..... 61
5.4. MANAGEMENT OF QOE ASPECTS ..... 65
5.5. SESAME SECURITY ASPECTS ..... 66
**6. ARCHITECTURE EXTENSIONS ..... 68**
**7. CONCLUSIONS ..... 69**
**8. REFERENCES ..... 70**

# List of Figures

Figure 1: Conceptual view of SESAME components ..... 15
Figure 2: Scope of SESAME components (physical view) ..... 16
Figure 3: Virtualized small cell architecture of SCF ..... 17
Figure 4: Considered splits between small cell virtualized and physical parts in SCF ..... 19
Figure 5: Baseline management system for small cells ..... 20
Figure 6: Management system for small cells with virtualization ..... 21
Figure 7: Hybrid SON architecture ..... 22
Figure 8: SON API architecture in case of cSON at NM ..... 23
Figure 9: SON API architecture in case of cSON at EM ..... 23
Figure 10: SON architecture in a virtualized environment: (a) in case of cSON at NM; (b) in case of cSON at EM ..... 24
Figure 11: MEC Architecture ..... 25
Figure 12: ETSI High-level NFV framework ..... 26
Figure 13: ETSI NFV reference architectural framework ..... 27
Figure 14: High-level view of overall T-NOVA System Architecture ..... 31
Figure 15: Overview of UNIFY architecture ..... 36
Figure 16: 3GPP next generation RAN architecture (RAN TSG) ..... 39
Figure 17: Network slicing in 3GPP next generation architecture (SA2 WG) ..... 39
Figure 18: NFV MANO (left) and Multitenant Small Cell network (right) architectural frameworks ..... 40
Figure 19: SESAME overall architecture ..... 41
Figure 20: Light DC Physical Architecture ..... 45
Figure 21: Mapping between ETSI ISG NFV functional entities and SESAME WPs ..... 50
Figure 22: LTE architecture ..... 50
Figure 23: 3GPP RAN nodes with Function allocations ..... 51
Figure 24: 3GPP LTE based eNB showing protocol stacks ..... 52
Figure 25: RAN network sharing based on MOCN [14] ..... 52
Figure 26: The mobile network management architecture mapping relationship between 3GPP and NFV-MANO architectural framework ..... 53
Figure 27: Fully centralized vs. partially centralized RAN functional architecture ..... 56
Figure 28: Physically distributed Light DC running cluster-level Service VNFs ..... 57
Figure 29: Data path in typical LTE connection and in SESAME approach ..... 57
Figure 30: SESAME solution for mobile edge traffic forwarding within the Light DC of the CESC Cluster ..... 58
Figure 31: Means of delivering a composition ..... 61
Figure 32: Service composition workflow ..... 64

# List of Tables

---

# 1. Introduction

## 1.1. Deliverable outline

As the number of mobile devices and data traffic soars, installing small cells is an effective way to achieve greater performance and capacity to both indoor and outdoor places, with significant impact upon all relevant market cases. The target of the SESAME project (GA No.671596) is to design and develop a novel 5G platform based on small cells, featuring multi-tenancy between network operators and also attach to them edge cloud capabilities to be offered to both the network operators and the mobile users. Thus, the key innovations proposed by SESAME focus on the novel concepts of a multi-operator (multi-tenancy) enabling framework and also on providing an edge-*based*, virtualised execution environment. The present Deliverable *D2.2* is the second specification document of the SESAME project, which presents the current outcomes of *Task 2.2,* with regard to the design and specification of the SESAME overall system architecture.

In order to provide a clear conceptual view of the broader SESAME system, this document considers: (i) The use cases and the requirements as they have been analysed in the prior relevant Deliverable *D2.1*, as well as; (ii) the state-of-the-art in the SESAME-*relevant* technological fields, e.g., Small Cells, NFV, Edge cloud architectures, including the most prominent standardisation trends, in order to "derive" a set of design principles and to culminate to a high-level architecture which*, in turn*, includes all the components of the SESAME concept.

Compliance with the current Small Cell Forum (SCF) and ETSI NFV vision has been targeted to the most possible degree at all stages of system design, especially concerning the terminology and the main architectural blocks; furthermore, several extensions have been proposed, especially at the parts that deal with multitenancy of operators and with virtualisation of functional components of the small cells.

SESAME aims at providing a fresh 5G mobile network architecture so as to support the ambitious goal of small cell virtualization, multitenancy and edge cloud services. Therefore, this deliverable covers a big range of elements and functionalities that cover the set of the corresponding targeted innovation areas. In this context, *Sections 2 - 3* provide an overview of the State-of-the-Art among the relevant fields, focusing further on the specific technology enablers under the specific scope of SESAME.

Considering the underlying requirements and *Use Cases* as they have been described in Deliverable *D2.1*, as well as the "key-technology" enablers that support the SESAME vision, *Section 4* provides the overall high-level architecture of SESAME. It has been elaborated to "match at the best possible extent" the important concepts of Network Function Virtualization (NFV), and multitenancy for Virtual Small Cell Network Operators (VSCNO). Also, it includes the rationale behind main design choices, as well as a high-level description of the main subsystems and interfaces. As regarding a further description of the essential modules and functionalities of SESAME (i.e., Cloud Enabled Small Cell (CESC) specifications, infrastructure virtualization, orchestration and management), two more deliverables, *D2.3* and *D2.4*, will provide further details on these aspects.

*Section 5* describes the service lifecycle and the sequence of interactions. These enablers are critical to the SESAME ecosystem and they are detailed for the system and service architectural perspectives.

*Section 6* presents architectural variations and configurations, according to additional requirements of the stakeholders.

Finally, *Section 7* concludes the deliverable.

## 1.2. Definitions of Terms and SESAME concepts

At this point, it is useful to provide definitions of terms and processes which will be used later in this document so that to describe the SESAME main concepts. Besides, for the definition of actors and their interactions the reader can also refer to *D2.1*.

- **Small Cell (SC):** Does not change in the context of SESAME.
- **Execution infrastructure, micro server (µS):** Specific hardware that is placed probably inside the Small Cell and provides processing power (also maybe some memory and storage capabilities).
- **CESC (Cloud Enabled Small Cell):** The Small Cell device which includes a micro server in hardware form.
- **Cluster of CESCs:** A group of CESCs that are collocated, exchange information and are properly coordinated. As a trivial case, one CESC can be called CESC cluster.
- **Light Data Centre (Light DC):** The hardware entity composed by the micro servers of the CESCs forming a cluster.
- **CESC Manager (CESCM)**: The architectural component in charge of managing and orchestrating the cloud environment of the Light DC, as well as management of small cell functions. It can manage, at the same time, multiple clusters, a cluster or a single CESC.
- **VIM:** Manager of the HW and networking resources (lifecycle, provision, placement, operation) constituting of a cluster of micro-servers, namely the Light DC, and the networking nodes and links (virtual and physical).

Figure 1, *below,* provides an overview of SESAME physical system perspective.

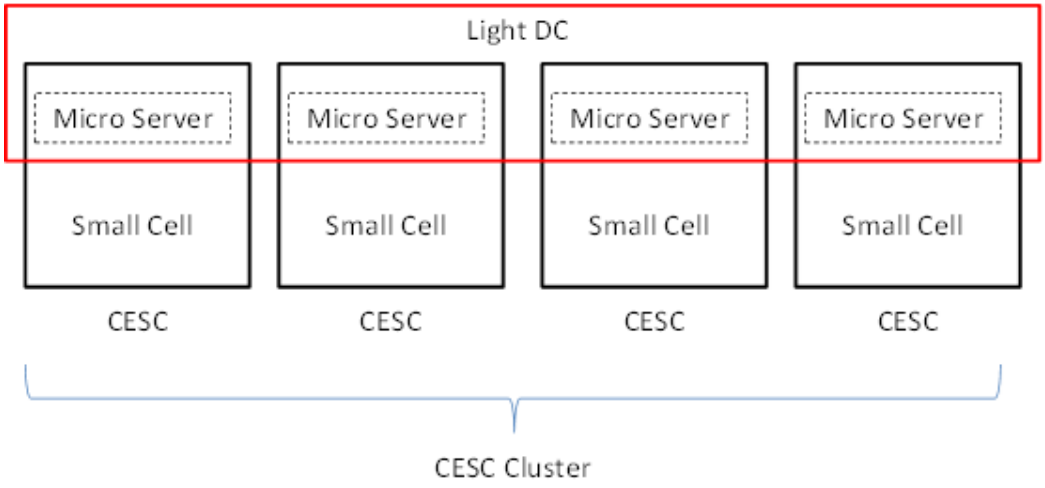



**Figure 1: Conceptual view of SESAME components**

In reference to the SESAME approach, the Small Cell concept is evolved so that not only be able to provide multi-operator radio access capacity with virtualised Small Cells that can be integrated within the operator (tenant) infrastructures, but also to be able to deliver a virtualised execution environment for delivering Cloud services at the network's edge. In order to achieve this, *however*, the Small Cell needs to provide mobile-edge computing capabilities which, *in turn*, will allow the virtual or mobile operators to increase the capacity of their own 4G/5G RAN infrastructures or to extend the range of their provided services, while maintaining the required agility to be able to offer these extensions, *on demand*.

Apart from being able to "abstract" these resources, *however*, some considerations need to be made on "how to separate or combine the network and the computing resources", and also, "which small cell functions should be physical network functions" and "which ones should be virtual".

The SESAME approach to this is performed by enhancing the Small Cells with Micro Servers that are able to provide virtualized computing and networking resources and by being able to "form" clusters, *thus*

*creating a Light DC at the edge*. This Light DC is further complemented by additional components that reside either close to the edge or to the backbone such as the VIM or the CESCM, in order to provide the proper reference points -or scope- for the whole network, *as depicted in the following diagram (given in the following Figure 2).*

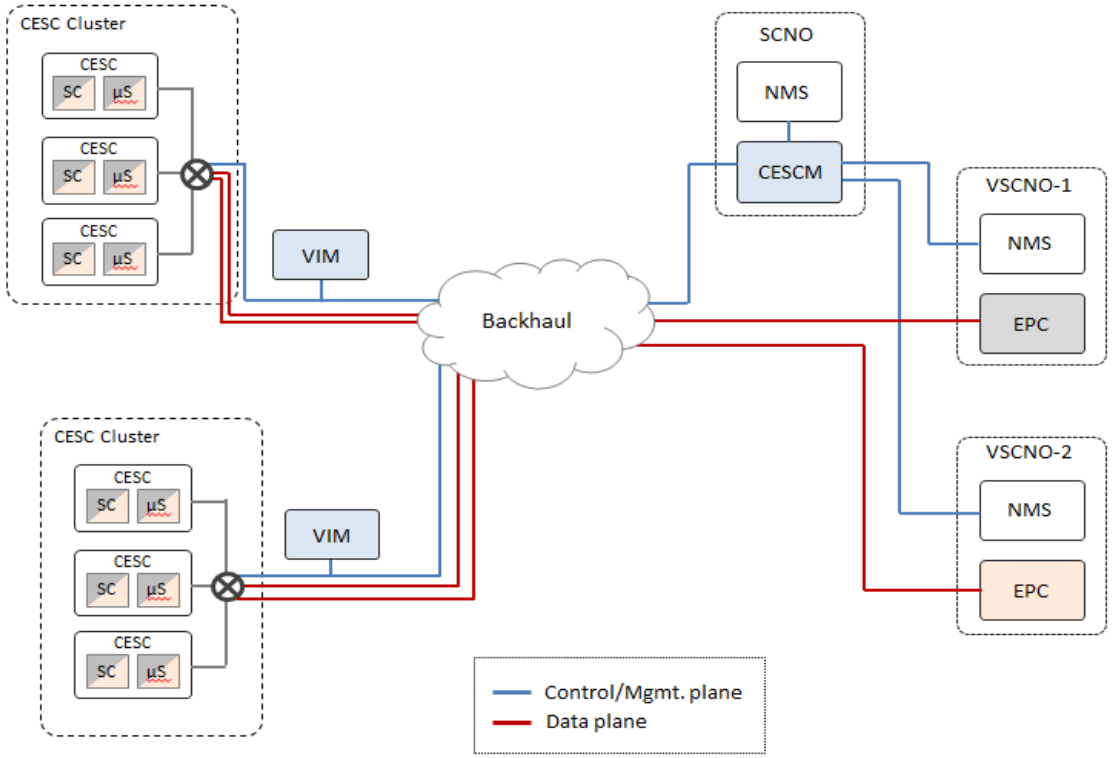



**Figure 2: Scope of SESAME components (physical view)**

# 2. Survey on LTE Small Cell Networks & Edge Cloud Concepts

## 2.1. Small Cell Forum

The Small Cell Forum (SCF) is an organization that intends to "drive" the wide-scale adoption of small cells and accelerate the delivery of integrated HetNets (Heterogeneous Networks). For that purpose, it works to ensure the adoption of industry-wide standards, a positive regulatory environment, common architectures and interoperability [55]. The following sections present a brief survey of the main aspects analysed by SCF that are related with the SESAME architecture.

### 2.1.1.Small cell virtualization: Functional split

SCF has carried out different studies analysing the introduction of virtualization technologies in small cell networks [49], [51], [52]. The small cell virtualization approach considered by SCF is illustrated in Figure 3 [51]. A small cell is split into two components: (i) A remote small cell, where functions are non-virtualized, i.e. they are Physical Network Functions (PNFs) that are implemented via a tightly coupled software and hardware, *and*; (ii) a central small cell, where functions are virtualized, i.e. they are Virtual Network Functions (VNFs) that are implemented by abstracting the hardware from the software, so that they are executed on a pool of shared computation, storage and networking resources. A central small cell will serve multiple remote small cells. The central and remote small cells are physically connected through the fronthaul link.

The main benefits of this small cell centralization and virtualization approach identified in [49] include: (i) Improved coordination of the radio functions between multiple remote small cells (e.g. coordinated scheduling, inter-cell interference coordination, etc.); (ii) enhanced scalability of small cell deployments with simplified management of the many physical small cells thanks to the centralization of functionalities at the central small cell; (iii) reduced cost from being able to lower peak-to-mean ratios for compute resources; (iv) accelerated upgrade lifecycles enabling new features to be deployed on a centralized virtual platform, *and*; (v) flexibility that enables optimum work load placement according to the availability of compute and transport resources.

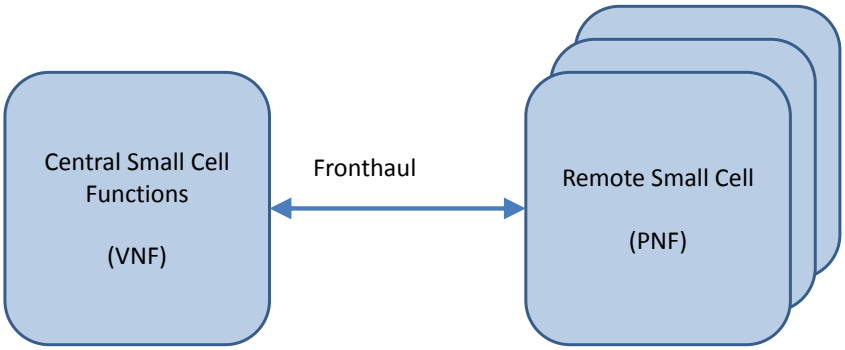



**Figure 3: Virtualized small cell architecture of SCF**

According to this virtualization approach, SCF has studied in [51] the functional split between VNF and PNFs, i.e. "which small cell functions should reside at the remote small cell" and "which ones at the central small cell". Considering different layers of the user/control plane protocol stacks of the radio interface, namely Radio Resource Control (RRC), Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC), Medium Access Control (MAC) and Physical layer (PHY) layers, the analysis follows a top-

down approach, where gradually more layers are virtualized by moving them from the remote small cell to the central small cell, thus resulting in the possible functional splits shown in Figure 4.
For each functional split the analysis determines the fronthaul latency and bandwidth requirements. It is worth mentioning that the analysis of splits done in [51] and reflected in Figure 4 does not explicitly separate between the user and control plane protocol stacks (e.g., RRC belongs to control plane while the rest of layers belong to both control and user planes). However, when analysing a given split, the fronthaul requirements are determined based on both user and control planes. In the following, a summary of the considered splits is presented:

**Centralized services:** This consists in virtualizing the small cell functions and applications that are above layer 3, such as content caching, Operation Administration and Management (OAM), Dynamic Host Configuration Protocol (DHCP) / Domain Name System (DNS) services, SON, etc. This split provides the common virtualization benefits of scalability and flexibility in how services are deployed, together with efficiency in power consumption and ease of integration of services from multiple vendors.

**Split RRC-PDCP:** In this split, the RRC protocol is virtualized at the central small cell, while the rest of protocols (PDCP and below) "stay" in the remote cell as PNFs. Therefore, the virtualized small cell includes the same applications and functions of the previous "centralized services" split and the RRC. This functional split allows the virtualization of certain functions that are supported by RRC, such as centralized connection mobility control, measurement reporting and handover trigger control, performance measurement management and inter-cell RRM. Estimated fronthaul data and control bandwidths for this functional split are 187.5 Mb/s for DL and 62.5 Mb/s for UL. This functional split is considered to be feasible with all fronthaul latencies considered in the analysis (i.e. from 250 μs up to 30ms).

**Split PDCP-RLC:** This case assumes that the virtualization split is done between PDCP and RLC, so that RRC and PDCP layers are virtualized at the central small cell, while the remote small cell keeps the RLC layer and below. The estimated fronthaul bandwidth requirements are similar to the split RRC-PDCP case, and this functional split is also feasible with all the considered fronthaul latencies.

**Split RLC-MAC:** In this case, the RLC layer functions are virtualized at the central small cell, while the MAC and Physical layer functions reside at the remote small cell. This provides some storage and processor utilization benefits, but it introduces complexity into the small cell implementation, because the downlink RLC is tightly coupled to the MAC and scheduler which remain at the physical side. Then, a flow control method is needed to decouple RLC and MAC to support this use case. The fronthaul bandwidth requirements are similar to those of the previous cases, and it requires backhaul latencies of maximum 6ms.

**Split MAC:** In this case, the upper part of the MAC layer which includes the MAC scheduler is virtualized on the central small cell, while the lower part of the MAC layer, which includes the Hybrid Automatic Repeat reQuest (HARQ) processing, remains at the remote small cell. This split enables central multi-cell scheduling. The bandwidth and latency requirements are similar to those of the previous split.

**Split MAC-PHY:** In this case all the MAC layer functionalities are virtualized at the central small cell, while the remote small cell only carries out physical layer functions. The virtualization of the HARQ results in tighter latency constraints on the fronthaul, requiring one-way latencies below 250 μs, which can be increased up to 2ms in case that HARQ interleaving technique is used. Bandwidth requirements are similar to those of previous splits.

**Split PHY:** This involves the virtualization of some physical layer 1 (L1) functions in the central small cell. Then, the central small cell will implement the upper part of the PHY layer while the remote small cell will keep the lower part. In this respect, there are in turn different options for doing this split, depending on the amount of baseband functions that are virtualized. They range from the case in which no L1 processing is done at the remote small cell, which simply becomes a remote radio head, to the case where most of the baseband processing is done at the remote small cell and only channel coding functions are kept at the central cell. This use case provides added benefits of resource sharing and load balancing for the DSPs, FPGAs, and hardware accelerators, but it involves increased requirements for the fronthaul. For

example, fronthaul bandwidth requirements reach up to 2.45 Gb/s for the case in which all the L1 processing is virtualized, and fronthaul latency should be below 250 μs.

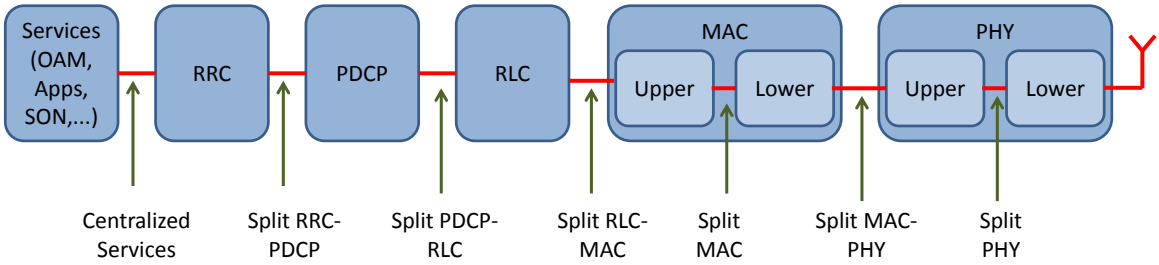



**Figure 4: Considered splits between small cell virtualized and physical parts in SCF**

In [52] the implications of centralization and virtualization are analysed from the perspective of their impact on certain techniques used to improve coverage and capacity. In particular, the focus is placed on four techniques, namely carrier aggregation, cross-carrier scheduling, high order MIMO and Coordinated Multi-Point (CoMP). The use of virtualization/centralization allows a more efficient implementation of these techniques in small cell networks, thanks to the capability of coordinating them between multiple small cells. However, this coordination capability is dependent on the considered functional split (e.g. in the case of carrier aggregation and cross-carrier scheduling the split has to be at the MAC or below, while in the case of high order MIMO it has to be at the PHY layer).

### 2.1.2. Impact of virtualization on the management architecture

The impact of small cell virtualization on the management architecture is analysed in [53], using ETSI-NFV and 3GPP architectures that support combined PNF and VNF systems. In particular, Figure 5 depicts the baseline management system for non-virtualized small cells and Figure 6 presents one of the options identified in [53] for managing small cells that include PNFs and VNFs.

The architecture shown in Figure 5 is aligned with the 3GPP management architecture of [11], and relies on the Small Cell Element Management System (EMS), in charge of the management of the small cells, and the Small Cell Network Management System (NMS), in charge of the management of the network. NMS and EMS are connected through the Itf-N interface[1]. Besides, SCF defines that the EMS performs the management of the small cells by means of the *TR-069 protocol* and the *TR-196v2 data model*, both specified by the Broadband Forum [58]. Different 3GPP Performance Management (PM) measurements are collected by the small cells and reported to the EMS.

The architecture shown in Figure 6 for the management of small cells composed of PNFs and VNFs is aligned with the ETSI-MANO[2] framework of [27] and its adaptation in 3GPP addressed in [7]. In the approach illustrated in Figure 6 the EMS is split in two components, the PNF EMS and the VNF EMS, that are in charge of managing the PNF and the VNF, *respectively*. This split is done under the consideration that the requirements that led to the definition of a TR-069-*based* EMS, like that of Figure 5 to manage a PNF small cell may not be directly applicable to the centralized VNF. Besides, Figure 6 also shows the need

[1] The "Itf-N" or "interface N" is built up by a number of IRPs and a related Name Convention, which realise the functional capabilities over this interface. The basic structure of the IRPs is defined in 3GPP TS 32.101 and 3GPP TS 32.102. Actually it is a Multi-vendor interface between Network Managers (NMs) and Element Managers (EMs)/ Network Elements (NEs). The Interface N is specified by multiple Integration Reference Points (IRPs).

[2] NFV MANO is a working group (WG) of the European Telecommunications Standards Institute Industry Specification Group (ETSI ISG NFV). It is the ETSI-*defined* framework for the management and orchestration of all resources in the cloud data center. This includes computing, networking, storage, and virtual machine (VM) resources. The main focus of NFV MANO is to allow flexible on-boarding and sidestep the chaos that can be associated with rapid spin up of network components.

to connect the PNF EMS and VNF EMS for coordination purposes, which can be done through the 3GPP defined Itf-P2P interface[3].
Another option identified in [53] for carrying out this coordination is through the NMS and the Itf-N interface.

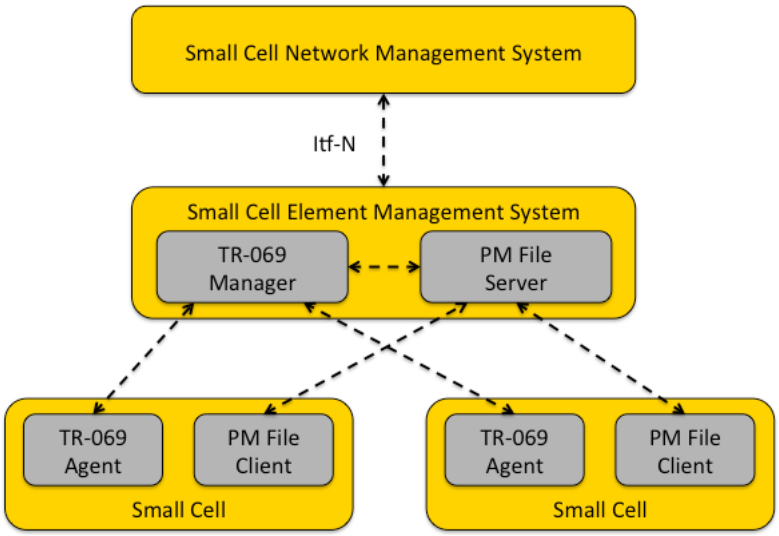



**Figure 5: Baseline management system for small cells**

In addition to the virtualization of the small cells that has been discussed until now, the study presented in [50] analyses also the application of virtualization techniques to other core network functions, such as small cell gateways, enterprise concentrators and enterprise gateways.

The document also considers the virtualization of the small cell management functions of Figure 5 (i.e. the TR-069 manager and the PM file server).

[3] Standardized "peer-to-peer interfaces" ("Itf-P2P") are conceived between sub-management functions to allow consistent management of the borders between the management domains. This includes cross-border operations, auditing and synchronization of common Network Resource Model elements and definition of the format for data exchange across these borders. Further details on the Itf-P2P interface are available in 3GPP TR 32.806.

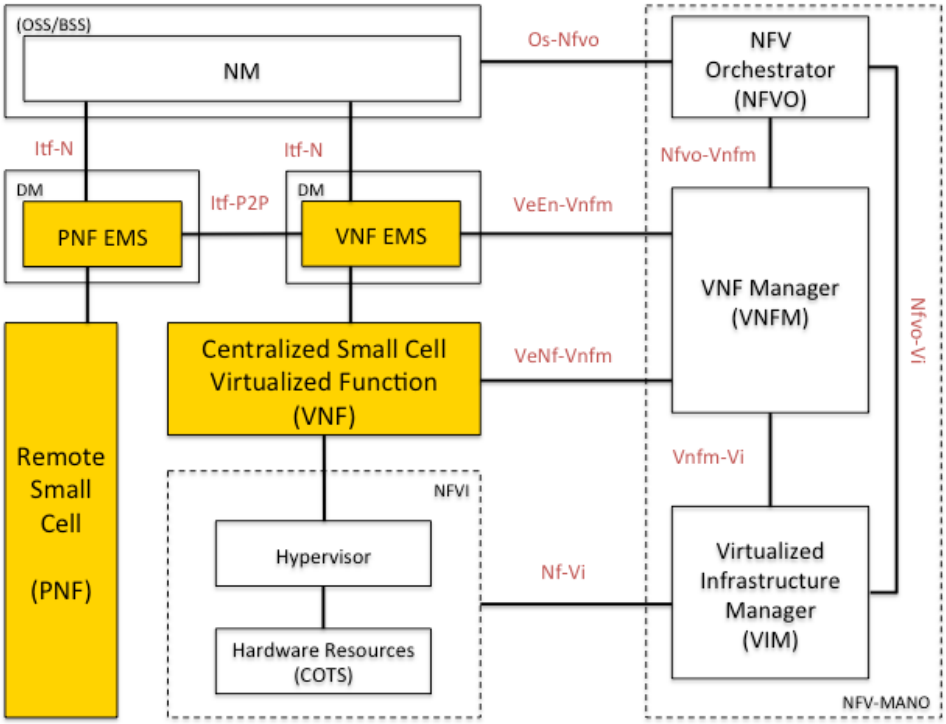



**Figure 6: Management system for small cells with virtualization**

### 2.1.3.Multi-operator small cells

SCF has also analysed the multi-operator small cells in which deployed small cells are shared between multiple operators. In the recent document [43] an analysis is presented on the market drivers for accelerating the adoption of multi-operator support, as this capability is seen as "critical" for several reasons, such as improving the economics and deployment process of dense indoor or urban networks, enhancing the revenue opportunities for Mobile Network Operators (MNOs) investing in small cells, supporting third-party neutral host models, and enabling services which rely on the participation of all operators to support the widest range of customers (e.g. mobile commerce and Internet of Things-IoT).
The study in [43] identifies different approaches to kick-start the multi-operator model, such as: (i) neutral host models in which the small cell network is deployed and run by an independent third party and supports a wide variety of MNOs and Mobile Virtual Network Operators (MVNO); (ii) it is important that the resource management model of the small cells is made fully MOCN-*aware* to enable greater operator control over its resources; (iii) new approaches to shared and flexible spectrum, e.g. through the provision of dedicated spectrum for the neutral host, through the use of dynamic spectrum access or technologies like LTE- License Assisted Access (LAA) that enable the use of unlicensed spectrum, etc.; (iv) inclusion of virtualization technologies that will facilitate *"network-as-a-service"* models, in which an MNO or neutral host can support flexible allocation of network and storage resources to multiple service providers, facilitating the allocation of bandwidth on demand dynamically adapting to changing needs of users and applications.

In [44], multi-operator small cell networks are addressed from the regulatory perspective, with main focus on the spectrum regulation to enable spectrum sharing between multiple operators, analysing different licensing models (e.g. license exempt, Licensed Shared Access, TV White Space, etc.).

From a more technical perspective, different alternative approaches to delivering shared infrastructure in small cells were discussed in a prior document [46]. The focus is mainly placed on the network sharing capabilities specified by 3GPP, such as Multi-Operator Core Network (MOCN), and on the management of multi-operator small cells (e.g. for broadcasting multiple PLMN-ids or for identifying neighbour cells in shared networks).

Based on the current information included in the abovementioned SCF documents, current status on multi-operator small cell networks has mainly covered market considerations and surveys about the current regulatory situation in different countries. Nevertheless, the documents have not addressed yet the technical implications of multi-operator small cells in areas like virtualization, functional splits, management, SON functions, etc., areas where SESAME is expected to contribute.

### 2.1.4.Self-Organizing Networks

Different Self-Organizing Network (SON) use cases for small cells are identified by SCF in [45] and [47] for enterprise and urban scenarios, *respectively*. They are classified into four categories, as follows: (i) Configuration (e.g. configuring a small cell for initial operation); (ii) planning, deployment and operation (e.g. addition of a new small cell, decommissioning of a small cell, reactivation of a small cell, new neighbour discovery, etc.); (iii) optimization (e.g. PCI conflict detection and resolution, transmit power optimization, load balancing, mobility robustness optimization, etc.), *and;* (iv) maintenance (e.g. registered agents push/pull SON metrics). In addition, some specific use cases where SON is supported by some form of human intervention are also discussed (e.g.: small cell positions are optimized with help from a technician, SON function requests or initiates a technician walk test to optimize some metric, etc.).

From an architectural perspective, a hybrid architecture for SON is adopted in [48], as depicted in Figure 7. SON functionalities are split between the small cells (*distributed SON - dSON*) and a central SON server (*centralized SON - cSON*) that can be seen as part of the Network Manager (NM) and/or the Element Manager (EM). The cSON provides guidelines and parameter ranges to dSON functions, based on information retrieved from the small cells (e.g. performance counters, alarms, etc.). The local dSON entities adjust local parameter settings within the provided guidelines. This is done autonomously at each eNB or interacting over X2 with neighbour eNBs.

Two general interactions approaches between cSON and dSON are envisaged in [48], depending on the particular function. In the “autonomous operation”, the cSON indicates parameter ranges valid for the dSON, which takes the parameter value decisions autonomously. Instead, with the “controlled operation” the cSON server takes a decision and signals the final parameter value to the dSON, which just assists the cSON with performance measurements and alarms.

A SON Application Programming Interface (API) is defined in [48] to allow exchanges of information between dSON and cSON. The corresponding architecture depends on whether the cSON is located at the NM or the EM. In case that cSON is at the NM, the SON API has two components (see Figure 8), one mapped onto the type 2 interface (e.g. Itf-N) between the NM and the EM, and the other onto the type 1 interface (e.g. TR-069) between the EM and the dSON at the small cells, and the EM includes a SON API agent responsible for translation between the type 1 and type 2 components. Instead, in case that cSON resides at the EM (see Figure 9), the SON API is only mapped onto the type 1 interface between the EM and the small cells.

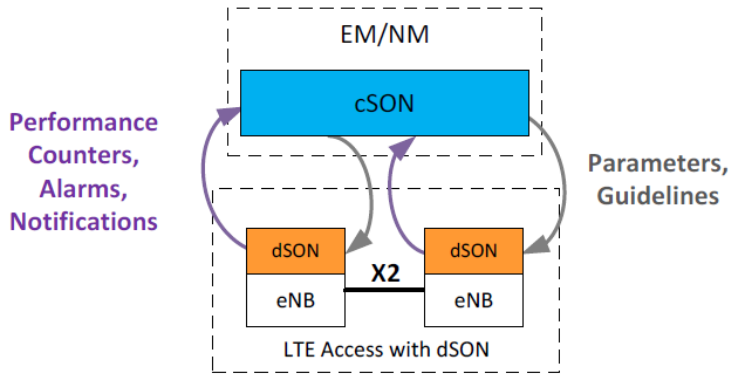



**Figure 7: Hybrid SON architecture**

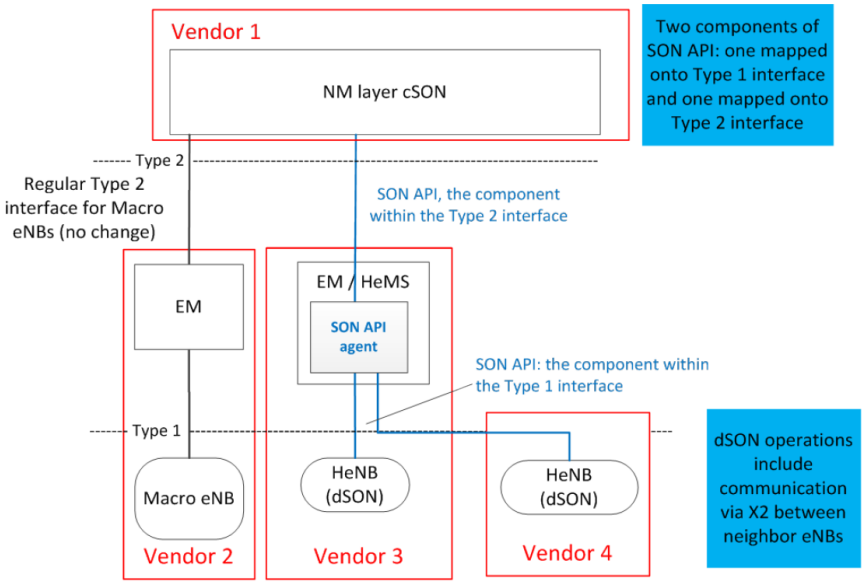



**Figure 8: SON API architecture in case of cSON at NM**

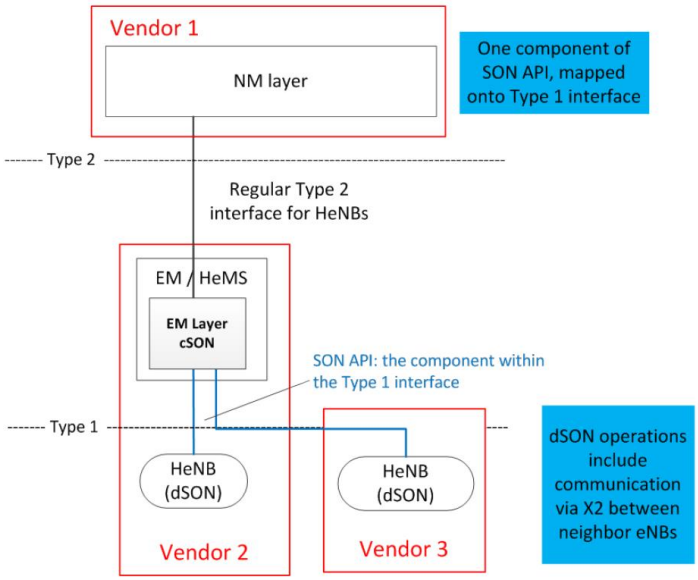



**Figure 9: SON API architecture in case of cSON at EM**

The SON API defines different procedures such as dSON parameter retrieve, PM configuration/reporting, REM (Radio Environment Measurement) configuration/reporting and event configuration/notification.

Figure 10 depicts the SON architecture when virtualization technologies are considered [53] as discussed previously in *Section 2.1.2*. In particular, a SON entity can be co-located with the small cell VNF, so that the dSON functionality can be split between the PNF and the VNF. In that case, Figure 10(a) depicts the SON API for the case where the cSON resides at the NM and Figure 10(b) considers the case where the cSON resides at the EM.

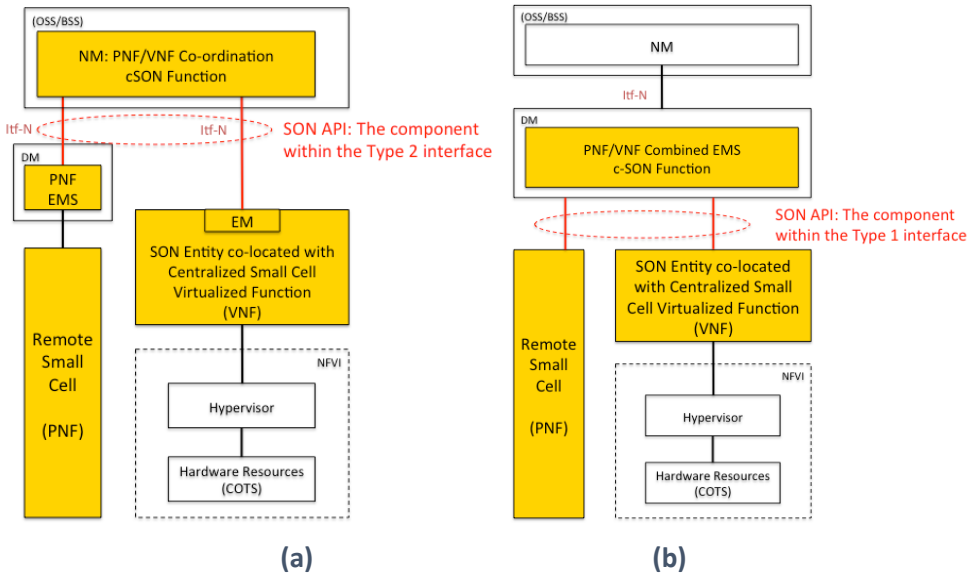



**Figure 10: SON architecture in a virtualized environment: (a) in case of cSON at NM; (b) in case of cSON at EM**

## 2.2. Overview of ETSI MEC

The idea of a computing platform placed at the edge of the mobile network is not new and is carried on inside ETSI by the Mobile-Edge Computing (MEC) Industry Specification Group (ISG) whose activity started on *December 2014*. It follows the current trends towards cloud-*based* architectures operating in an IT environment but with the peculiarity of being located at the edge of the mobile network, within the Radio Access Network and in close proximity to mobile subscribers. MEC platform wants to take advantage from the existing NFV infrastructure - that provides a virtualization platform to network functions enhancing it with new computing/storage resources and creating a virtualization environment for a wide range of applications running at the mobile network edge. Distinctive features of the MEC architecture are low latency, proximity, location awareness, high bandwidth, real-time insight into radio network information. This facilitates accelerated delivery of content, services and applications from the edge of mobile networks, closer to end-users. The mobile subscriber's experience can be significantly improved through more efficient network and service operations, enhanced service quality, minimized data transit costs and reduced network congestion.

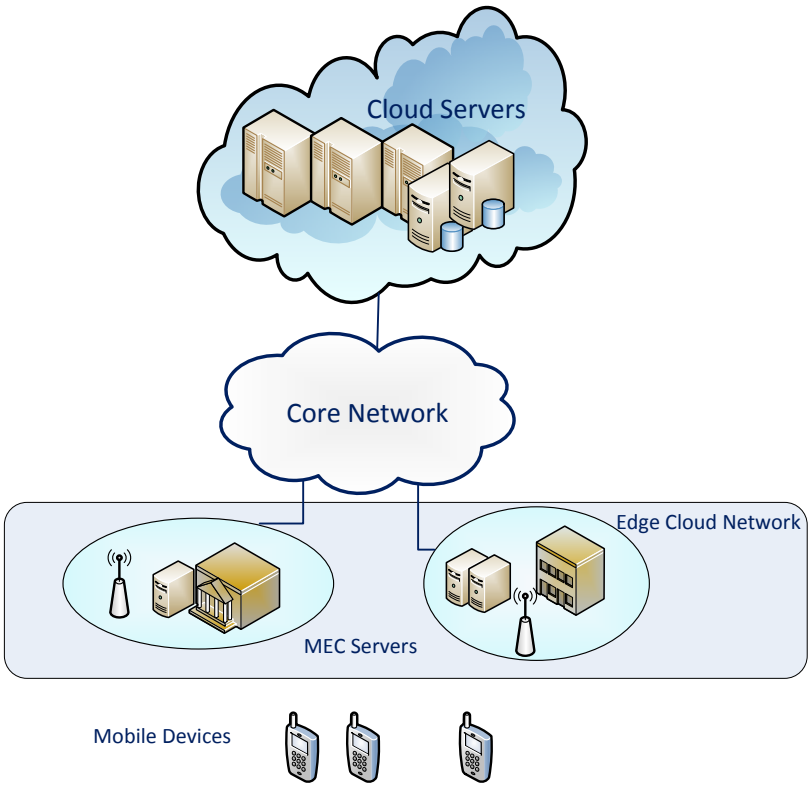



**Figure 11: MEC Architecture**

The MEC servers provide computing, storage and bandwidth capacity that is shared by multiple virtual machines installed on top of them; being owned and managed by the infrastructure provider, they are directly attached to the base stations (BSs). Traditionally, all data traffic originating at the data centres is forwarded to the mobile core network. The traffic is routed through the core network to a base station which delivers the content to the mobile devices. In the mobile edge computing scenario, MEC servers take over some or even all of the tasks originally performed in a data centre. Being located at the mobile edge, this eliminated the need of routing these data through the core network, leading to low communication latency.

## 2.3. Overview of ETSI NFV ISG and ETSI MANO

Today's IT and Networks sectors have been producing a collaborative work using their expertise in the "field" and their resources towards a joint effort. The target is "to come to an agreement" on common architectures and standardized functionalities and approaches that address the identified technical challenges in these fields. As a result, a network operator-*led* Industry Specification Group (ISG) with open membership was setup in the *last quarter* of *2012* under the umbrella of ETSI to work through the technical challenges of Network Functions Virtualisation (NFV).

It should be mentioned that ETSI ISG NFV does not provide exact standards, but rather it produces documents that contain guidelines. The ETSI ISG NFV delivers its findings in the form of Group Specifications not in the form of European Norms (EN) or Technical Standards (TS). The outputs are openly published and shared with relevant standards bodies, industry Fora and Consortia to encourage a wider collaborative effort. If misalignments are detected, the ETSI ISG NFV will collaborate with other SDOs in order to meet the requirements.

The NFV ISG also provides an environment for industry to collaborate on Proof of Concept (PoC) platforms in order to demonstrate solutions, which address the technical challenges for NFV implementation and to encourage growth of an open ecosystem.

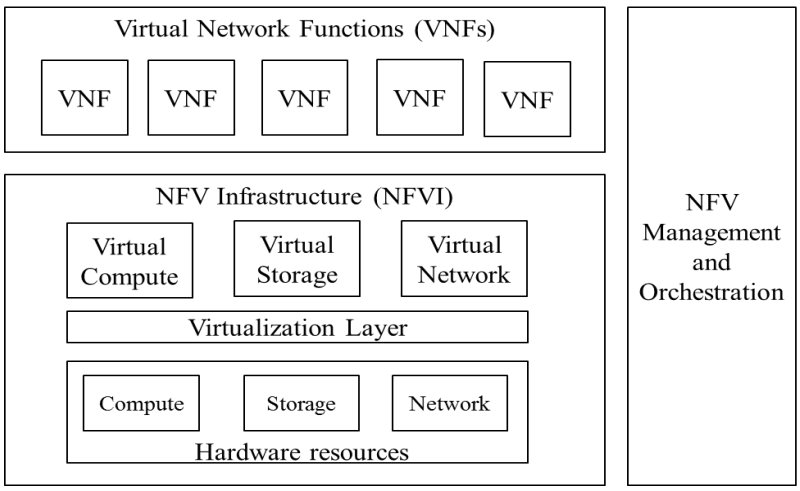



**Figure 12: ETSI High-level NFV framework**

The NFV concept envisages the implementation of Network Functions (NFs) as software-only entities that run over the NFV Infrastructure (NFVI). Figure 12 published in *October 2013* by the ETSI ISG NFV in its document on global architecture, illustrates the high-level NFV framework, where three main working domains can be identified:

- Virtual Network Function (VNF), as the software implementation of a network function which is capable of running over the NFVI.
- NFV Infrastructure (NFVI), which includes the diversity of physical resources and how these can be virtualised. NFVI supports the execution of the VNFs.
- NFV Management and Orchestration (NFV MANO), which covers the orchestration and lifecycle management of physical and/or software resources that support the infrastructure virtualisation, and the lifecycle management of VNFs. NFV MANO focuses on all virtualisation-specific management tasks necessary in the NFV framework.

The NFV architectural framework handles the expected changes that will probably occur in an operator's network due to the network function virtualisation process. Figure 13 shows this global architecture, depicting the functional blocks and reference points in the NFV framework.

The architectural framework shown in Figure 13 focuses on the functionalities necessary for the virtualisation and the consequent operation of an operator's network. It does not specify which network functions should be virtualised, as that is solely a decision of the owner of the network.

Three ETSI NFV ISG Working Groups have been established following the identification of the above-mentioned domains (i.e. NFVI and MANO).

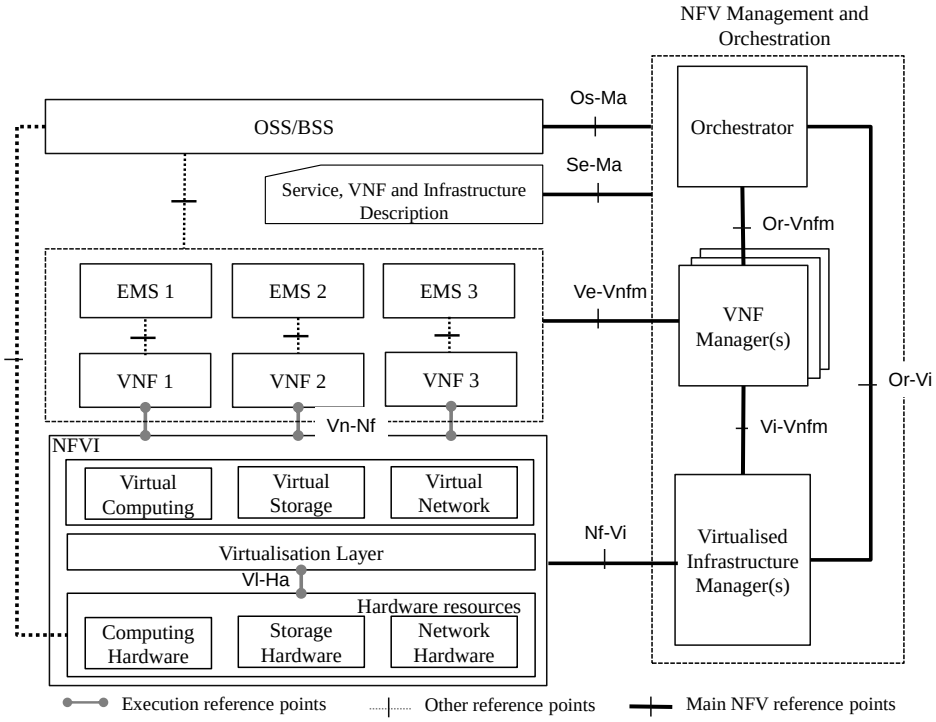



**Figure 13: ETSI NFV reference architectural framework**

# 3. Survey on Architectures of projects relevant to SESAME

## 3.1. H2020 5G-PPP Projects relevant to SESAME

A range of 5G-PPP projects are working on access network architecture related issues and the concepts and approaches they are following are highly relevant for the small cells as a service concept defined and instantiated by the SESAME project. While the approach to exploit SDN and NFV concepts has been seen as a driving factor mainly for core network features, the virtualisation of the radio access has not been investigated sufficiently so far. The key concepts of SESAME are to define a multi-operator (multi-tenancy) enabling framework that is capable of providing an edge-based, virtualised execution environment to provide "Small Cells as a Service" ("SCaaS").

The framework will rely on a number of underlying features, functions and capabilities that are provided by the complementary work undertaken in some of the other 5G-PPP projects, this includes in particular:

- ***5G-EX***, the ***5G Exchange*** [15] project defines an open platform enabling cross-domain orchestration of services over these domains. The aim is to enable collaboration between operators, on a service level, including 5G infrastructure services. *5G-EX* relies on the integration or unification of infrastructure from several operators by introducing unification via NFV/SDN compatible multi-domain orchestration. The aim is to facilitate rapid and automated service provisioning.
  The service management approach of *5G-EX* may be able to exploit different resource and other management features the SESAME framework offers.

- ***5G NORMA*** [16] develops a mobile network architecture that will be adaptable while being resource efficient. The main target is to handle fluctuations in traffic demand by implementing "multi-service and context-aware adaptation of network functions". The approach is expected to increase energy-efficiency and it considers a "mobile network multi-tenancy" approach for inherent dynamic sharing and distribution of network resources between operators.
  While the focus of *5G NORMA* is on energy and resource efficiency, it complements the flexibility that the SESAME platform provides in terms of *Small Cells as a Service*.

- ***COHERENT*** [17] defines a control framework with two main aspects that facilitate tighter inter-network coordination. The first one is to "abstract" the low layer network states and behaviours of underlying mobile networks for network-*wide* control and resource coordination. The second is the definition of common interfaces and software-development kits to enable programmability in controlling and coordinating heterogeneous mobile networks.
  This programmable control may be used in conjunction with some of the management functions and features defined in SESAME.

- ***SPEED-5G*** [19] tackles the capacity problem through the deployment of small cells and ultra-densification. The basis of the approach is to manage dynamic control across wireless network resources, which typically leads to unbalanced spectrum loads and a perceived capacity bottleneck. eDSA (extended DSA), resource management with three degrees of freedom: densification, rationalized traffic allocation over heterogeneous wireless technologies, and better load balancing across available spectrum is seen as the main technique investigated.
  The algorithms and mechanisms investigated in Speed5G provide many of the features required by the PNFs defined in SESAME.

- **M*etis II*** [18] defines a spectrum management architecture which also enables efficient integration of new and legacy air interfaces. Across the different air interfaces, an agile Resource Management (RM)

framework for efficient operation of the integrated 5G air interfaces complements the architecture. A common control and user plane framework for efficient service support complements the approach. The SESAME approach can again exploit the spectrum management architectural features as well as the resources of the integrated air interfaces, as long as a suitable adaptation layer is provided.

## 3.2. FP7 T-NOVA Architecture

The ***T-NOVA*** *project*[4] aims to design and implement a management and orchestration framework for the automated provision, configuration, monitoring and optimisation of "*Network Functions as-a-Service*" over virtualised Network and IT infrastructures.

The *T-NOVA* architecture can be hierarchically organised into four architectural layers:

- The *NFV Infrastructure (NFVI)* layer includes the physical and virtual nodes (commodity servers, VMs, storage systems, switches, routers, etc.) on which the services are deployed.
- The *NFVI Infrastructure (NFVI) Management* layer comprises the infrastructure management entities (VIM, TNM). The NFVI and management layers are conceptually grouped under the name *Infrastructure Virtualisation and Management (IVM).*
- The *Orchestration* layer is based on the *T-NOVA* Orchestrator and also includes the NF Store.
- Finally, the *Marketplace* layer contains all the customer-facing modules which facilitate multi-actor involvement and implement business-related functionalities.

We now present an overview of the aforementioned layers and their components.

***F1. IT infrastructure virtualisation, supporting resource elasticity***

State-of-the-art data-centre IT virtualisation technologies, including modern cloud platforms can fulfil most of the functionalities required for *T-NOVA*. The typical IT virtualisation structure can be followed, based on *compute*, *storage*, DC *network* and *hypervisor* domains.
On top of these, a unified management framework conforming to contemporary cloud management paradigms can be foreseen, where the compute/storage/DC network and hypervisor domains are jointly managed achieving the automated provision, management and optimisation of IT IaaS services.

***F2. Network infrastructure virtualisation, supporting resource elasticity***

Intra-data-centre network assets are controlled by the Virtualised Infrastructure Management (VIM). Contemporary Software Defined Network (SDN) technologies significantly facilitate software-*driven* and vendor-*agnostic* network management and are thus commonly employed in modern data centres, so it is realistic to assume that the intra-DC network is SDN-*enabled*.
For the wide-area network (hereafter *Transport Network*), a separate management entity needs to exist, managing end-to-end connectivity between DCs/NFVI-PoPs and managing traffic steering to establish the end-to-end network service. This management entity is identified as *Transport Network Management (TNM).*

***F3. Automated service provisioning, monitoring, scaling and optimisation***

The establishment of an end-to-end *T-NOVA* service normally requires the knowledge of the available resources, the planning of the service and the interaction with the VIMs and the TNMs managing the infrastructure assets which will be involved. All these operations need to be carried out by a higher-level management entity, *one per Service Provider*, which jointly orchestrates the underlying infrastructure. This is the role of the *T-NOVA Orchestrator*, which should incorporate functional modules dealing with *Resources Orchestration, Network Service Orchestration* as well as *VNF Management.*
In order to fulfil its role, the Orchestrator also maintains internal catalogues containing information about available and established Network Services, available VNFs and deployed instances of both, as well as infrastructure resources.

[4] More information about the T-NOVA (FP7 ICT) project can be found at: http://www.t-nova.eu/.

At the same time, a Repository for hosting the VNF images and associated metadata should be foreseen (*NF Store)*, from where the VNF images are retrieved for deployment as instances into the infrastructure.

***F4. Service advertisement and brokerage***

A dedicated, customer-facing module is foreseen, dedicated to *Brokerage* functionalities. Apart from pricing and trading policies, the Brokerage function should have access to Service and VNF catalogues as well as to the NF Store.

***F5. SLA, billing and accounting support***

In the same context, dedicated modules are also foreseen for: (i) *Accounting;* (ii) *SLA Management* and; (iii) *Billing.* These components collect resource usage data, monitor the status of the established SLAs and bill the customer accordingly. All service-*related* information can be presented through a unified *Dashboard*, which can also facilitate interactions with service brokerage procedures.

By assembling all the aforementioned components, a high-level view of the *T-NOVA* system architecture is derived, as shown in Figure 14.

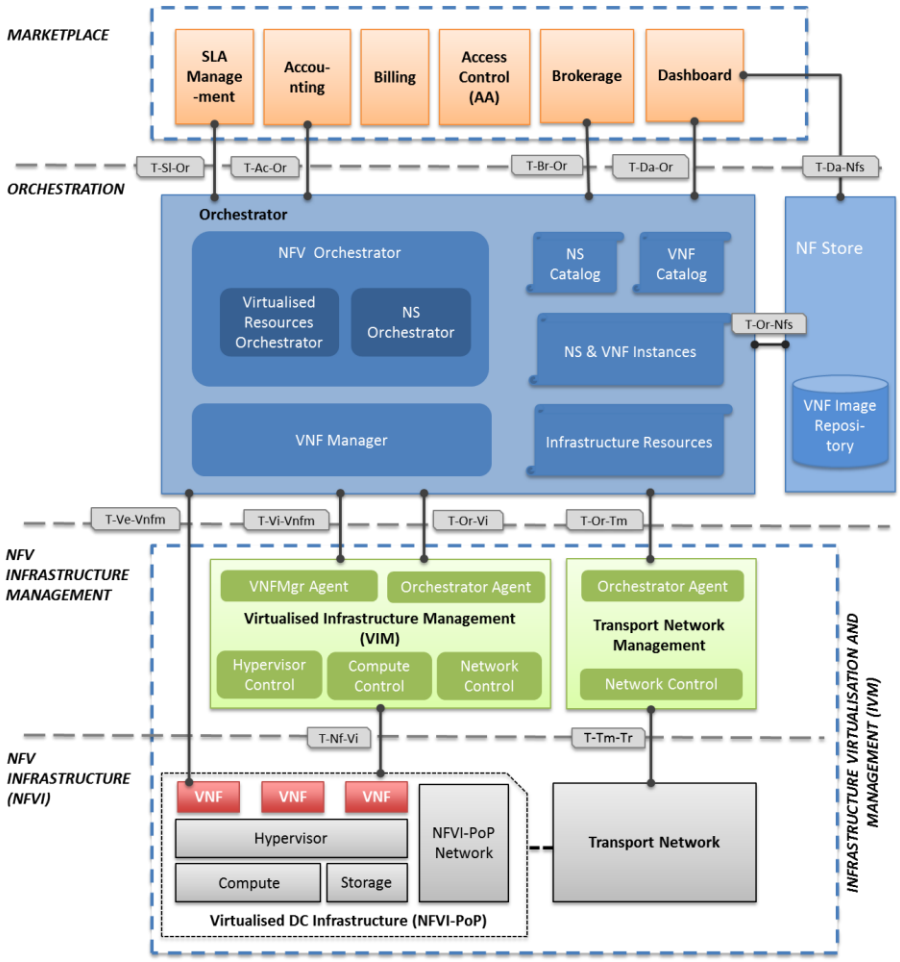



**Figure 14: High-level view of overall T-NOVA System Architecture**

## 3.3. FP7 MCN Architecture

In today's situation, there are key challenges for telecom industries. The two key questions of the telecom and mobile industries are namely:

1. How can they optimise their CAPEX/OPEX [29], [39], offering the same service at a lower cost or with greater profit, and;
2. How can they incentivise their existing subscriber base to use new and innovative services by efficiently leveraging the vast amounts of infrastructure at their disposal and in doing so create new revenue streams?

These questions are addressed by the ***Mobile Cloud Networking (MCN) project***[5].
MCN is further focused upon two key ideas as a means to address these challenges.

1. The first idea is to exploit cloud computing as infrastructure for future mobile network deployment and operation. At the core, MCN is about the movement from systems that are self-hosted, self-managed, self-maintained, on-premise designed and operated, to cloud-native-*based* software design that respects existing standards and management and operation of those designs instantiated as services or indeed, network function services.
2. The second idea is more visionary and forward looking. It illustrates "how the project envisions a future MCN service provider that fully endorses and adopts cloud computing services". These providers will leverage these services by building value-added services to commercially exploit as new innovative services, both traditional telecom services as well as new composed end-to-end (E2E) services.

MCN is driven by cloudification and delivery of cloud-native applications and services with an aim of remaining compatible with ETSI NFV architectures.
Nonetheless the considered sectors of MCN are not just IT but also mobile telecom. This is one of the **core propositions of MCN**, to use the cloud and to extend cloud services from the IT world into the mobile telecom world.

MCN -and its architecture- encompasses and integrates (converges) the three major stakeholder domains within this project. These three domains are:

1. **Radio Access Network (RAN):** Subscriber's wireless access is serviced. The challenge here within MCN is to virtualise the base station.
2. **Mobile Core:** Subscriber's access through to internet services is given. The challenge here is the servitisation and cloudification of the Evolved Packet Core (EPC) architecture.
3. **Data Centre:** Further services such as compute and storage (block storage, object storage, database etc.) services are offered. The challenge here is how these can be used efficiently within mobile and telecom environments.

The MCN architecture is based on SOA that provides a common service-*based* model, defining key entities following a common service lifecycle to enable the cloud-*based* delivery of services. It encourages the use and combination of new and innovative services by efficiently leveraging the vast amounts of cloud infrastructure and services available.
The approach in MCN is not dependent on whether the deployment is a public or private one, both can be used separately or in combination. MCN considers applications that are pre-existing, however in [36], strong arguments are made for the adaptation of services/functions so they run in a cloud environment in such a way to leverage the cloud itself. These recommendations lead to the cloudification of a service/function. MCN adopted a service-*oriented* architecture, however only sought to be designed on

[5] More information about the MCN (FP7 IP) project can be found at: *https://www.mobile-cloud-networking.eu/*

the basis of the SOA principles. These principles are: autonomy, formal-contract-*based* programmability, loose coupling, composability[6], reusability, statelessness and discoverability.

To deliver these principles required the design of architectural entities and a service lifecycle. The MCN service lifecycle take inspiration from TM Forum's eTOM[7] [57] and maps to it. The MCN life-cycle has been divided into two distinct phases: (1) the **business phase** and; (2) the **technical phase**.

The **business phase** contains all activities related to the conceptualization of the service plus the agreements of contracts between partners. This phase is largely a human- and manual-based process:

- **Design:** This is the phase where the service is conceptualized, the services that cannot be supplied by the organization are sourced from other organizations, and requirements upon the external services to be combined are collected and studied.
- **Agreement:** Here items such as pricing, Service Level Agreement (SLA), Access, etc., are agreed between two or more organizations. The agreements are generally bilateral business ones.

The **technical phase** is guided and governed by the business phase decisions and agreements between providers. In the technical phase the following sequence of phases are carried out:

- **Design:** Design of the architecture, implementation, deployment, provisioning and operation solutions. Supports Service Owners to "design" their service.
- **Implementation:** it is relevant to the designed architecture, functions, interfaces, controllers, APIs, etc.
- **Deployment:** It is the creation of the required resources. It refers to the supply of anything such that the service can be used, but does not provide access to the service. For example, placing a VM image on an IaaS and creating an instance from it.
- **Provisioning:** Provisioning of the service environment (e.g. VNFs, interfaces, network, etc.). Activation of the service such that the EU can actually use it.
- **Operation and Runtime Management:** Activities such as scaling, reconfiguration of SICs happen here.
- **Disposal:** Release and destruction of SICs and the SI itself and therefore all related resources.

Completely adopting and implementing these service lifecycles are the key MCN Architectural entities that are delivered through the MCN Service Management Framework. These entities are as follows:

- **Service Manager (SM):** Provides an external interface that presents one or more service types of be instantiated by a tenant. It is responsible for managing a collection of service orchestrators. The SM's programmatic interface (northbound interface, NBI[8]) is designed so it can provide either a CLI and/or a UI. Through the NBI, the SM gives the tenant or another upstream SO capabilities to create, list, detail, update and delete tenant service instance(s). Its Service Catalogue contains a list of the available service types offered by the provider. Its Service Repository is the component that provides the functionality to access the Service Catalogue. The SO Management (SOM) component has the task of receiving requests from the NBI and overseeing, initially, the deployment and provisioning of the service instance. Once the instantiation of a service is complete, the SOM component can

[6] The characteristic of composition reflects the notion of an “end-to-end” MCN service - i.e. which composes all the necessary services to deliver service from the user equipment (UE) all the way through to the target value-delivering service.

[7] The Business Process Framework (eTOM) is a standard maintained by the TM Forum, an association for service providers and their suppliers in the telecommunications and entertainment industries. More related information can also be found, *inter-alia*, at: *https://en.wikipedia.org/wiki/Business_Process_Framework_(eTOM)*.

[8] In computer networking and computer architecture, a *northbound interface* (NBI) of a component is an interface that conceptualizes the lower level details (e.g., data or functions) used by, or in, the component. A northbound interface is used to interface with higher level layers using the *southbound interface* of the higher level component(s). In architectural overviews, the northbound interface is normally drawn at the top of the component it is defined in, hence the name northbound interface. A southbound interface decomposes concepts in the technical details, mostly specific to a single component of the architecture. Southbound interfaces are drawn at the bottom of an architectural overview.

oversee tasks related to the execution of the SO and its later disposal. SOs are tracked in its SO Registry component.

- **Service Orchestrator (SO):** Oversees the end-to-end orchestration and composition of a SI (Service **Instance). Once created and running, it manages the SI and its SICs (Service Instance Components)** and is isolated per-tenant. Generally, only one SO is instantiated per each SI, however for reliability and quorum more can be provided. The instantiated SO is always associated with the SM that created it. It is a domain specific component as it has all the specific orchestration logic encoded within it. In particular, it is responsible for SIC instantiation and configuration, triggering of scaling and migration of SICs according to metrics collected within or by the runtime component of the CC. When the SO is created by the SM, it is just a bundle of code resources, known as the SO bundle. What is contained in the SO bundle is domain specific, however these items are express as either the SO implementation (SO-I), the SO's **Service Template Graph (STG)** or the **Infrastructure Template Graph (ITG)**. The SO-I this is the actual code implementation for the creation of a SI and implements methods mapped to the MCN lifecycle. The STG contains all the required supporting and atomic service needed by the SO. The ITG defines how resources should be composed to be able to host Service Instance Components. For example, the MCN Analytics service requires two virtual machines: one to handle compute execution and one to handle the storage backend, both of which are connected through a network.
- **Cloud Controller (CC):** Abstracts from specific technologies that are used in the technical reference implementation. It is through the CC how the service bundle is created. This component plays a key role in the support of several infrastructures (e.g. CloudSigma[9], OpenStack[10]) and platform technologies (e.g. Google App Engine, Azure, OpenShift). The CC's OCCI-*based* API's specification has been submitted to the OCCI[11] group and will figure as part of the upcoming 1.2 release of the standard.

The categorisation of services available in MCN is given in the following:

- **Atomic Service:** This is an indivisible service that executes a particular singular business or technical function. An Atomic Service is not subject to further decomposition.
- **Composed Service:** It aggregates/combines services together with orchestration logic. Both Atomic and Composed Services can be used to create further composed services. MCN supports two types of composed services.
- **Support Services:** These provide targeted, specific functionality for use by any other support or MCN service. These can be thought of being the platform services of MCN which other services use to carry out specific functions.
- **MCN Services:** These are the key services that demonstrate the use of the MCN architecture. They contain their own domain specific orchestration logic and may use both support and atomic services.

It should be noted that the capabilities of MCN, being SOA-*based*, allows for recursive composition of services. Another key note of the MCN architecture is that it is compatible with the ETSI-NFV architecture.

[9] For more related information, see: *https://www.cloudsigma.com/*.

[10] *OpenStack* is a free and open-source software platform for cloud computing, mostly deployed as an infrastructure-as-a-service (IaaS). For more related information, see, *for example*: *https://en.wikipedia.org/wiki/OpenStack*.

[11] The *Open Cloud Computing Interface* (*OCCI)* is a set of specifications delivered through the Open Grid Forum for cloud computing service providers. OCCI has a set of implementations that act as proofs of concept. More information can be found, *inter-alia*, at: *https://en.wikipedia.org/wiki/Open_Cloud_Computing_Interface*.

## 3.4. FP7 UNIFY architecture

The UNIFY project ("*Unifying Cloud and Carrier Networks*" - GA No.619609)[12] aims to extend the virtualization and automation enabled in cloud environments across the whole telecom carrier field, and thus to "unify" the networking and cloud infrastructure.
Telecom providers struggle with low service flexibility, increasing complexity and related costs. Although cloud computing and networking have been two very active fields of research, there were little integration between the vast networking assets and data centres of telecom providers. UNIFY addresses this by considering the entire network, from home networks up to data centre, as a "*unified production environment"*. The unified approach will open up the potential of virtualization, programmability/ automation to span across the whole infrastructure of the provider and guarantee an unprecedented level of agility for network operations and for deploying new, secure and quality of experience aware services, with seamless instantiation across the entire infrastructure. This will lead to a novel service-oriented carrier-grade platform for the Future Internet and brings virtualized services in the most efficient, secure and quality-aware way to the end-users.

To serve the aforementioned purposes, the UNIFY architecture defines the main architectural components and reference points which are relevant to its concept and which enable appropriate framework to support orchestration, SP-DevOps[13] and high performance data plane.
In this section, we briefly overview the UNIFY architecture and its constituent elements and highlight its relevance to the scope of SESAME.

UNIFY focuses on enablers of such unified production environment and will develop an automated, dynamic service creation platform, leveraging a fine-granular service chaining architecture. A service abstraction model and a proper service creation language will enable the dynamic and automatic placement of networking, computing and storage components across the infrastructure. A global orchestrator, which novel optimization algorithms, ensures the optimal placement of the elementary service components across the entire infrastructure. New management technologies, based on experience from data centre and tightly integrated into the service orchestration architecture, have been so proposed, developed and cope with the dynamicity of new services.

The so-called overarching view of the UNIFY architecture comprises of three major layers: the Service Layer, the Orchestration Layer and the Infrastructure Layer. The architecture also includes management components, a Network Functions System (NFS) and reference points between the major components [60] (see Figure 15).

[12] More details about the UNIFY project are given in: *https://www.fp7-unify.eu/*.

[13] The UNIFY SP-DevOps brings together a set of scalable monitoring approaches along with automated verification and activation functions for dynamic service chains, tools and workflows that empower the development and operations teams to significantly increase their efficiency will be developed. The SP-DevOps will reduce the number of manual diagnosing steps in carrier-grade software defined networks and allow the operators to determine validity of functional implementations within one agile development cycle. SP-DevOps will also design methods for rapid identification of the locations of fault or performance degradations in a service chain built from virtualized software defined service blocks. [*DevOps* (a clipped compound of "development" and "operations") is a culture, movement or practice that emphasizes the collaboration and communication of both software (SW) developers and other IT professionals while automating the process of software delivery and infrastructure changes].

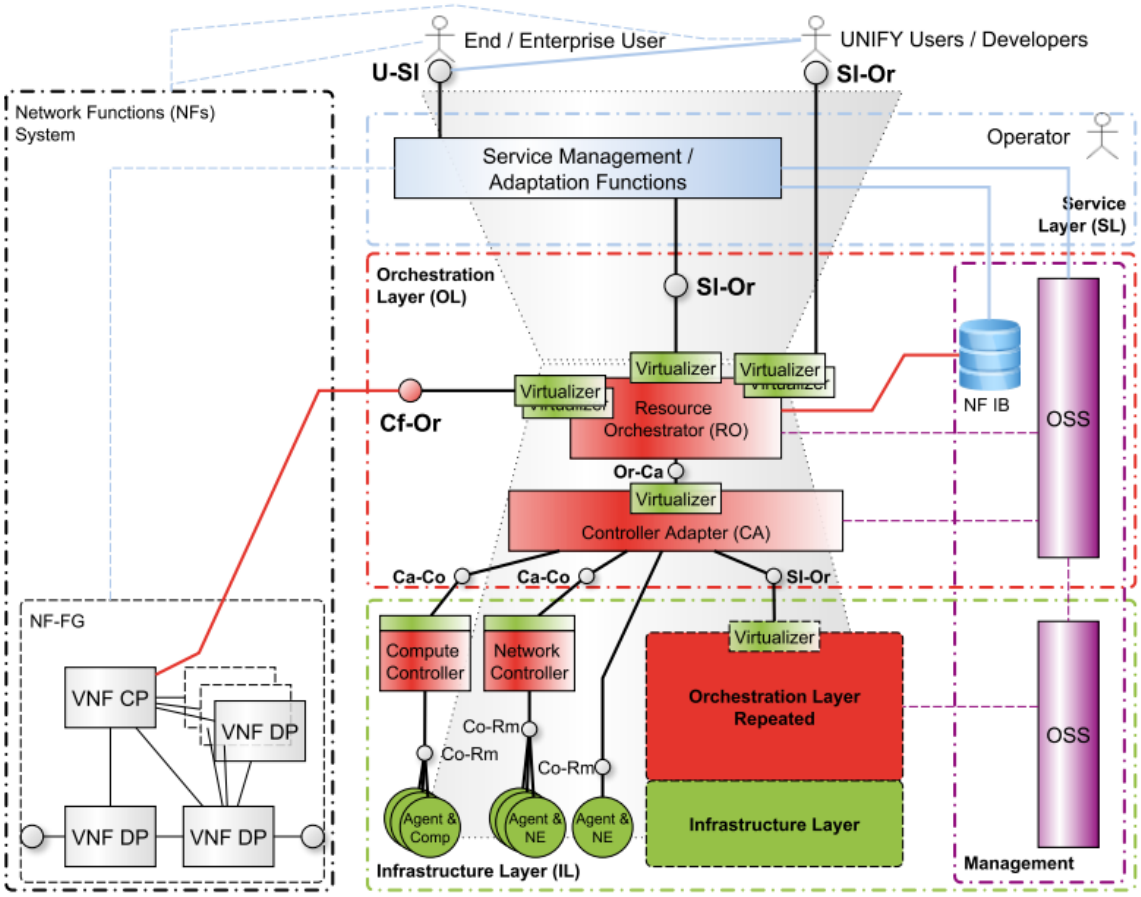



**Figure 15: Overview of UNIFY architecture**

The Service Layer (light blue) comprises "traditional" and virtualization-related management and business functions concerned with service lifecycle. Traditional functions include *inter-alia* Element Management Systems (EMSs), Operation Support Systems (OSSs) and Business Support Systems (BSSs) related to service and business management associated with Network Service Providers (NSPs) or simply Service Providers (SPs). On the other hand, virtualization-*related* management functions include lifecycle management for virtualized network services, lifecycle management for Virtualized Network Functions (VNFs) and orchestration of the latter over the resources presented by the lower layer. The management functions of the Service Layer are infrastructure agnostic and are responsible to handle the management of offered services.

The Orchestration Layer (red) comprises two major functional components, namely the Resource Orchestrator (RO) and the Controller Adapter (CA). RO performs policy enforcement and resources orchestration on virtualisers, and between them and underlying resources. CA is a multi-technology, multi-vendor/multi-domain controller that maintains the global view of available resources and supported capabilities and provides it to RO. RO and CA may be managed by a management system such as an OSS.
Finally, the Infrastructure Layer (light green) naturally comprises physical resources, local resource agents and/or controllers handling a single domain, such as a transport network or a Data Centre (DC). Again, controllers, agents and physical resources are managed by a corresponding management system such as an OSS.

Orthogonal to Orchestration and Infrastructure layers, Management (purple) includes infrastructure management functions and a Network Function Information Base (NF-IB), while orthogonal to all three layers, Network Functions System (NFS) includes instantiated NFs, including data, control and management plane components and the corresponding forwarding overlays.
Moreover, within the UNIFY architecture, so-called Reference Points are defined. Reference points practically are interfaces that allow information exchanges between "producer" and "consumer"

functional blocks of information. For instance, in Figure 15, U-Sl is the reference point between users and the Service Layer management and adaptation functions, while Sl-Or is the reference point between Adaptation Functions in the Service Layer or external consumers and the RO, etc.; similarly, all other reference points are defined.
Main components of the UNIFY architecture are Virtualisers, the Network Functions Forwarding Graph (NF-FG), Service Management and Adaptation Functions, the Universal Node, the Controller Adapter (CA), the Network Function Information Base (NF-IB), the Resource Orchestrator (RO), the Policy Enforcement, and the Monitoring.

CA and RO have been already described. Virtualisers are responsible for the allocation of abstract resources and capabilities to particular consumers and for policy enforcements, should be vendor, technology and domain agnostic and should contain the resources at compute, storage and network abstraction and capabilities. NF-FG is an abstract information model, which is central to the UNIFY framework, and it is used primarily at the Sl-Or, Cf-Or and Or-Ca reference points. NF-FG defines a selected mapping of NFs and their forwarding overlay definition into the virtualized resources presented by the underlying virtualiser. Service Management and Adaptation Functions map Key Quality Indicators (KQIs) associated to the overall performance of a service/product, which is meaningful to customers, to a set of Key Performance Indicators (KPIs).

The Universal Node is a collection of forwarding elements and embedded software resources, albeit with high performance data plane execution environment, under a joint RO for both software and network resources. The NF-IB is a database or catalogue containing resource models for NF abstractions at networking, compute and storage resource level. Policy Enforcement is focused on the Sl-Or and Cf-Or reference points and comprises the following elements: Policy Administration Point (PAP), Policy Enforcement Point (PEP), Policy Decision Point (PDP) and Policy Information Point (PIP). Finally, Monitoring is one of the most essential building blocks of the UNIFY framework in terms of maintaining performance of a service and rapid reaction to infrastructure failures. In this regard, Monitoring Functions (MFs) collect a wide range of information that indicates performance and availability of system components, e.g., VNFs.

The high-level SESAME architecture, as defined in the SESAME DoW, comprises blocks, components and interfaces, which can be mapped to layers, components and references points in the UNIFY architecture; nevertheless extension or re-formulation of such UNIFY elements would be necessary to address the SESAME objectives and scope. For instance, the NFV Orchestrator of SESAME can be mapped to the processes of the Service Layer and the interaction between RO and CA of UNIFY, while the NFV Manager could be mapped to a Virtualiser of the Service Layer interacting through the Cf-Or reference point with NFS and NF-FG component.
Moreover, abstraction models and management of resources, and the involved components and functionalities of the UNIFY Infrastructure Layer could accordingly be extended to address the Small Cells requirements.

## 3.5. 5G design principles in 3GPP SA2

Although the work related to future 5G systems in 3GPP is still in its embryonic phase, the first outcomes of the involved workgroups allow foreseeing the main working assumptions for the future mobile broadband networks.

The following aspects are considered relevant to the main concepts of SESAME.
In the scope of 3GPP SA1[14], the Feasibility *Study on New Services and Markets Technology Enablers (SMARTER)*[15] identifies the list of potential requirements for future mobile networks [3].
Among the different categories of use cases, the Network Operation (NEO) building block includes a series of potential system requirements related to system flexibility, scalability, mobility support, efficient content delivery, self-backhauling and interworking with 4G systems [2] that are clearly in the scope of SESAME.

The 3GPP SA2[16] is developing the *Study on Architecture for next Generation System* [4], aimed at providing the main architectural requirements, assumptions and principles.
One of the "key issues" in the definition of the next generation mobile networks is the concept of network slicing and the horizontal and vertical limits.

Providing network slicing in the Core Network is in a more mature state nowadays, and it is already under specification in the 3GPP SA5[17] work group under the working item OAM14-MAMO_VNF[18]. In this sense, the relationships with ETSI NFV and ETSI MANO standards are of great relevance for the virtualization of the network functions.
Concerning the virtualization of the RAN, the 3GPP RAN TSG has initiated the *Study on Scenarios and Requirements for Next Generation Access Technologies* [8]. Different functional splits of the RAN are proposed as potential solutions for the Radio Transmission Points (R-TP), all of them can coexist under a unified management system. Additionally, the proposed architecture splits the user plane and control plane.

Taking into account the different potential RAN functional splits, the extension of the virtual network slices up to the UE is yet unclear.
In this sense, [4] provides an analysis of the network slicing issue without including the RAN. In that case, the RAN may remain as a common network segment that includes a new element for "slice identification and selection" (similar to the NAS Node Selection Function[19]).

The architectural solution provided in SESAME *D2.2* needs to cope with the different potential functional splits of the RAN, introducing management elements for the physical and virtualized network functions.
It is in the scope of SESAME *D2.3* to identify the most appropriate functional splits, while SESAME *D2.4* details the management interfaces and their specific roles.

---

[14] For more information, see: *http://www.3gpp.org/specifications-groups/sa-plenary/sa1-services*.
[15] For more corresponding information, see, for example: *http://www.slideshare.net/yihsuehtsai/new-services-and-markets-technology-enablers-smarter-lte-relese-13-and-road-to-5g*.
[16] For more information, see: *http://www.3gpp.org/specifications-groups/sa-plenary/sa2-architecture*.
[17] For more information, see: *http://www.3gpp.org/specifications-groups/sa-plenary/sa5-telecom-management*.
[18] For more information, see: *http://www.3gpp.org/DynaReport/WiCr--680036.htm*.
[19] More relevant information is given in ETSI TS 136 410.
*http://www.etsi.org/deliver/etsi_ts/136400_136499/136410/09.00.00_60/ts_136410v090000p.pdf*.

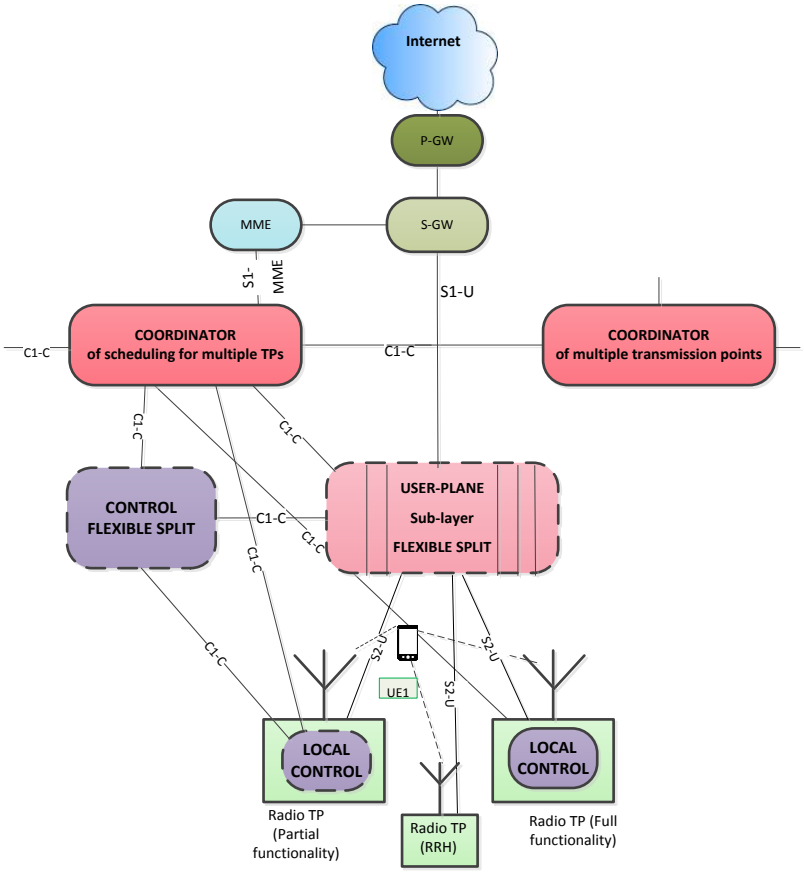



**Figure 16: 3GPP next generation RAN architecture (RAN TSG)**

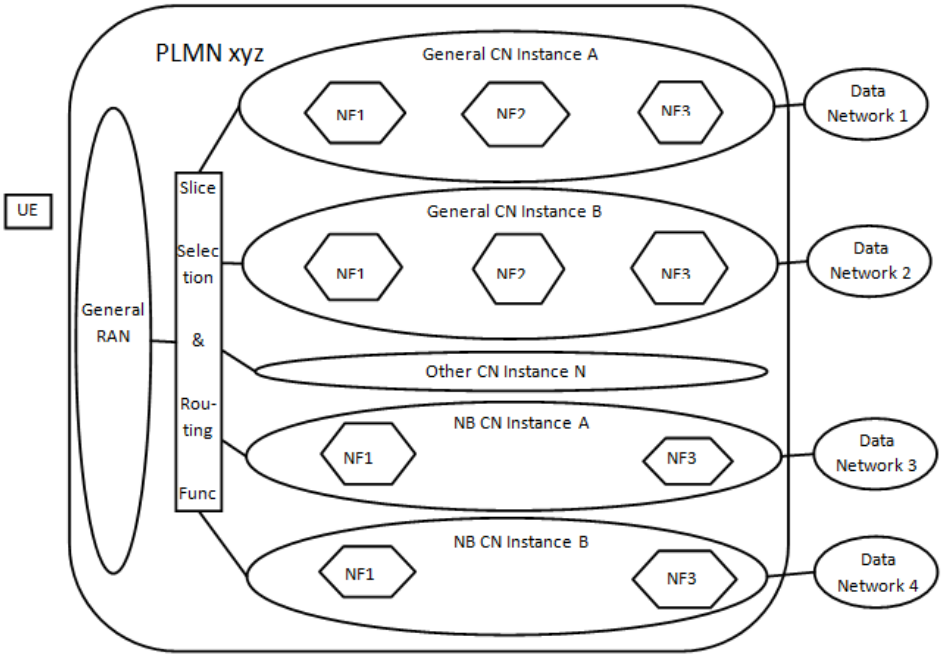



**Figure 17: Network slicing in 3GPP next generation architecture (SA2 WG)**

# 4. SESAME Overall Architecture

## 4.1. Key Features of the SESAME System and Architecture Principles

The key innovations proposed in the SESAME architecture focus on the novel concepts of virtualising Small Cell networks by leveraging the paradigms of a multi-operator (multi-tenancy) enabling framework coupled with an edge-*based*, virtualised execution environment.

SESAME falls in the scope of these two principles and promotes the adoption of Small Cell multitenancy, i.e., multiple network operators will be able to use the SESAME platform, each one using his own network "slice". Moreover, the idea is to endorse the deployment of Small Cells with some virtualized functions, with each Small Cell containing also a micro-server through appropriate fronthaul technology. A micro-server is based on a non-x86 architecture[20] using 64-bit ARMv8 technology[21]. Together with the SC, they form the Cloud- Enabled Small Cell (CESC) and a number of CESCs form the "CESC cluster" capable to provide access to a geographical area with one or more operators.

At this point, we provide a brief description of the two main technological fields that constitute the main fields of innovation of SESAME. This kind of "decomposition" has been the starting point for building at the next following stage an accurate framework for SESAME architecture. To that end, the NFV technology is going to be used as an enabler that will offer a virtualisation platform and meet the requirements of SESAME, namely NFV-driven small cell functions and NFV-based network services. The left-hand side of Figure 18 presents the MANO framework for the NFV part.

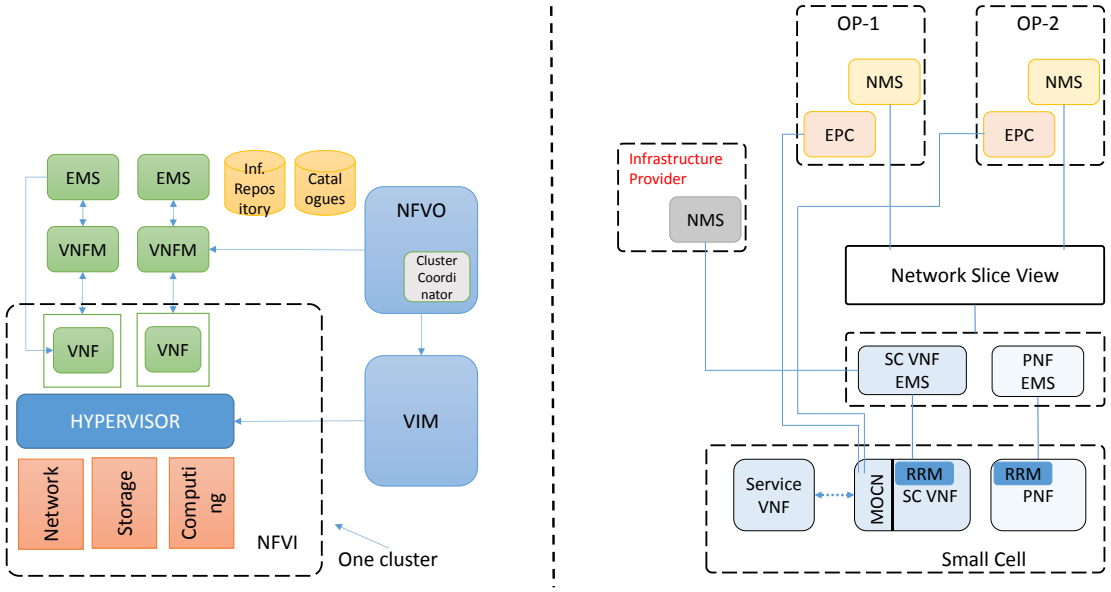



**Figure 18: NFV MANO (left) and Multitenant Small Cell network (right) architectural frameworks**

[20] The *x86* is a family of backward compatible instruction set architectures based on the Intel 8086 CPU and its Intel 8088 variant. More related information can be found, for example, at: *https://en.wikipedia.org/wiki/X86.*

[21] For more related information, see:*http://www.arm.com/products/processors/armv8-architecture.php*.

On the other hand, a Small Cell network capable to support more than one network operator is envisaged (Figure 18 – right-hand side). 3GPP specifications have already added some support for Radio Access Network (RAN) sharing [9]. Although two main architectures are identified, namely Multi-Operator Core Network (MOCN), where the shared RAN is directly connected to each of the multiple operator's core networks, and Gateway Core Network (GWCN), where a shared core network is deployed so that the interconnection of the multiple operator's core networks is done at core network level, the MOCN case has been identified as the exclusive enabler for multitenancy features in SESAME platform. The infrastructure consists of a number of Small Cells and the corresponding SC network functions such as gateways and management systems. The architecture is based on the current 3GPP framework for network management in RAN sharing scenarios [13], [11], [12]. Assuming LTE technology as the contextual framework, the interconnection of the SCs of the SCaaS provider to the Evolved Packet Core (EPC) of the tenant is done through the S1 interface, delivering both data (e.g., transfer of end-users traffic) and control (e.g., activation of radio bearers) plane functions.

Based on the required functionalities as well as architectural principles as mentioned above, it is possible to derive an overall, high-level view of the SESAME system, as proposed by Figure 19.
To that end, the CESC offers computing, storage and radio resources. Through virtualization, the CESC cluster can be seen as a cloud of resources which can be sliced to enable multi-tenancy. Therefore, the CESC cluster becomes a neutral host for mobile Small Cell Network Operators (SCNO) or Virtual SCNO (VSCNO) which want to share IT and network resources at the edge of the mobile network. In addition, cloud-*based* computation resources are provided through a virtualised execution platform. This execution platform is used to support the required Virtualized Network Functions (VNFs) that implement the different features/capabilities of the Small Cells (and eventually of the core network) and the cognitive management and *Self-X* operations, as well as the computing support for the mobile edge applications of the end-users.

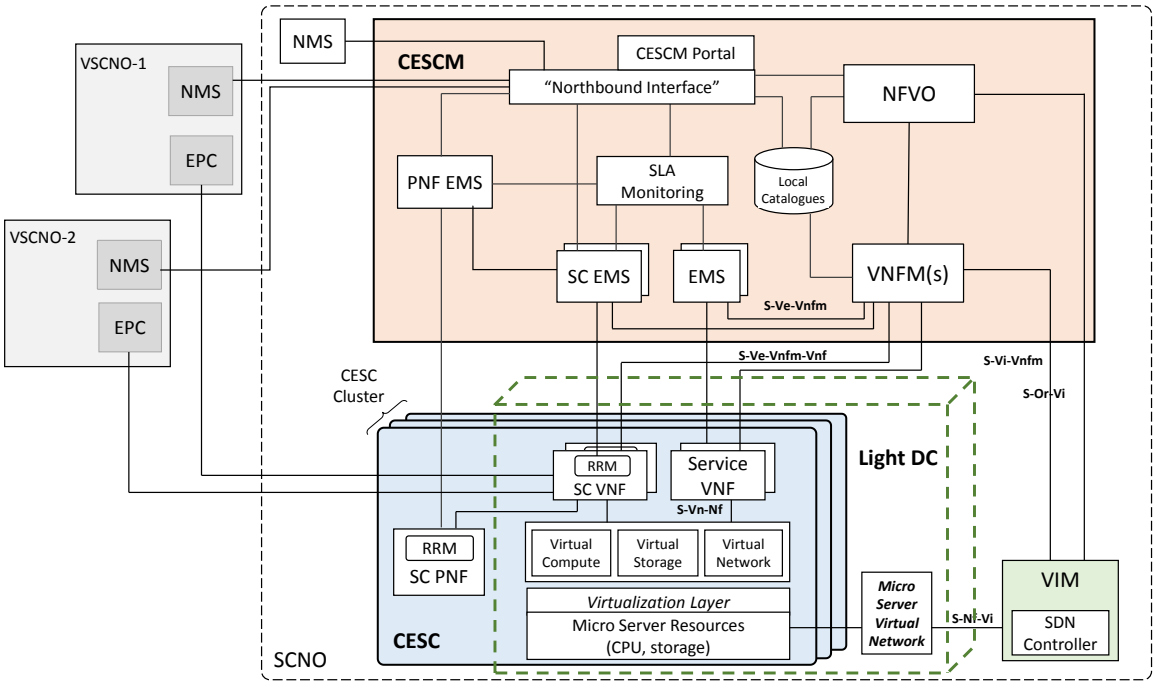



**Figure 19: SESAME overall architecture**

The CESC clustering enables the achievement of a micro scale virtualised execution infrastructure in the form of a distributed data centre, denominated Light Data Centre (Light DC), enhancing the virtualisation capabilities and process power at the network edge.
Network Services (NS) are supported by VNFs hosted in the Light DC (constituted by one or more CESC), leveraging on SDN and NFV functionalities that allow achieving an adequate level of flexibility and scalability at the cloud infrastructure edge. More specifically, VNFs are executed as Virtual Machines

(VMs) inside the Light DC, which is provided with a hypervisor (based on KVM) specifically extended to support carrier grade computing and networking performance.
Over the provided virtualised execution environment (Light DC), it is possible to chain different VNFs to meet a requested NS by a tenant (i.e. mobile network operator). Note that, in the context of SESAME, a NS is understood as a collection of VNFs that jointly supports data transmission between User Equipment (UE) and operators' Evolved Packet Core (EPC), with the possibility to involve one or several service VNFs in the data path. Therefore, each NS is deployed as a chain of SC VNFs and Service VNFs.
Finally, the CESC Manager (CESCM) is the central service management and orchestration component in the overall architecture figure. Generally speaking, it integrates all the necessary network management elements, traditionally suggested in 3GPP, and the novel recommended functional blocks of NFV MANO [37]. A single instance of CESCM is able to operate over several CESC clusters, each constituting a Light DC, through the use of a dedicated Virtual Infrastructure Managers (VIM) per cluster.
With regard to interfaces, it must be noted that Figure 19 mostly depicts reference points –which may contain one or more actual interfaces- between architectural layers. Each reference point label starts with "S-" to differentiate it from interfaces defined in ETSI NFV ISG documents (and in specific Vi-Vnfm, Or-Vi, Ve-Vnfm, Nf-Vi) – although in many cases the functionality of the reference point will be almost aligned to the ETSI definitions.

## 4.2. High-level Description of SESAME Main Architectural Entities

### 4.2.1.CESC

In the scope of SESAME, a CESC consists of a complete small cell with its standard backhaul interface, standard management connection (TR69 interface[22] for remote management) and with necessary modifications to the data model (TR196 data model [59]) to allow Multi-Operator Core Network (MOCN) radio resource sharing. The CESC is composed by a Physical small cell unit and an ARM V8 non-x86 execution platform architecture, i.e. the micro server. Physical small cell and micro server are logically connected whereby a single 3GPP compliant S1 interface over Gigabit Ethernet.
Cloud Computing and networking is realised through the sharing of computation, storage and network resources of the micro servers present in each CESC and composing the Light DC. The Light DC supports the virtualised execution environment required for implementing different features/capabilities of the Small Cells and the cognitive/self-x management operations as VNFs. The collection of CESCs forms the CESC cluster.

The CESC is meant to accommodate multi-tenancy by design, enabling the offer of the *platform as a service*, capable of providing optimised sharing of the deployed physical infrastructure among multiple VSCNOs. Different VNFs are hosted in the environment provided by the micro server for different tenant operators. VNFs encompass small cell functions and service functions, with the latter including deep packet inspection, firewall and caching, to name a few. This also provides the support for mobile edge computing applications deployed for each tenant that, operating very near to the end users, may significantly reduce the service delivery time and deliver composite services in automated manner. Moreover, the CESC is the termination point of the GTP-U tunnelling which encapsulates user IP packets from SGW which are destined to the UEs and from the UEs to the SGW.
MOCN is one of the small cell VNFs supported in the micro server. In the MOCN-*based* deployment an operator can share the RAN with other operators. In SESAME, the concept of MOCN is extended through RAN virtualisation, allowing the infrastructure provider to deliver S*mall Cell-as-a-Service (SCaaS)* through management and orchestration of CESC resources for the tenants. The multi-tenant platform is able to meet the required capacity demand within a certain area, as well as leverage reuse of spectrum. Thus, apart from improved user experience, the CESC can offer the possibility to offload traffic from congested macrocells, allowing for a variety of value-added services.

[22] The *TR-069 (Technical Report 069)* is a technical specification that defines an application layer protocol for remote management of end-user devices. It was published by the Broadband Forum and entitled *CPE WAN Management Protocol* (CWMP). More information can be found, *inter-alia*, at; *https://en.wikipedia.org/wiki/TR-069*.

The CESC exposes different views of the network resources: per-tenant small cell view, and physical small cell substrate, which is managed by the SCNO, decoupling the management of the virtual small cells from the platform itself. In the CESC, rather than providing multiple S1 (or Iu-h interface[23]) connections from the physical small cell to different operators' EPC network elements such as MME and SGW [56], such fan-out is done at the micro server. The CESC is further the termination of multiple S1 interfaces connecting the CESC to multiple MME/SGW entities as in S1-Flex[24].

The CESC includes the following components:

- ***Small Cell Physical Network Function (SC PNF):*** The SC PNF implements the radio interface and possibly the main protocol features of the LTE (H) eNB protocol stack, and, together with the SC VNFs, provides the complete functionality of the virtualised small cell. Details on the functional split between SC VNF and SC PNF are available in [25].
- ***Small cell VNF (SC VNF):*** This is part of the small cell functions which is hosted and executed in the micro server. Different small cell functions are instantiated and managed in the micro server by the management and orchestration layer in the CESC Manager. The VNFs are allocated creating VMs in the micro server. Different VNFs are instantiated depending on the specific functional split, as discussed in *Section 2.1.1.* The functional split shall define which parts of the LTE protocol stack are implemented as VNFs. As mentioned above the virtualised small cell might include MAC, RLC and PDCP at the U-plane, whereas at the C-plane RRC functions. In addition, RRM and Self-x features are also implemented as VNFs in the micro server. A SC Common VNF which is in charge of the fain-in and fan-out to/from the physical small is implemented as a VNF running in the micro server across multiple VSCNOs. As mentioned earlier, also service VNFs run in the micro server and can be chained together to perform composite network services. This prescribes defining the forwarding graph of the service chain for example relying on OVS technology.

As shown in Figure 19, both the SC PNF and SC VNF have an Element Management System (EMS) for FCAPS (Fault, Configuration, Accounting, Performance and Service) purposes. Both resides in the CESC Manager and are connected to the Network Management System (NMS).

Automated operation of frequency reuse-one CESCs in the SESAME multi-tenant system is made possible by Self-x features (e.g. Self-Planning, Self-Optimising and Self-Healing). Self-Organizing Networks (SON), also referred to as *"self-x",* features include a number of techniques for automating the operation of the network by automatically tuning different network settings and can be implemented as physical or virtual network functions instantiated and executed in the micro server. Typical examples of Self-x features include Inter-Cell Interference Coordination (ICIC) to configure the power, time and frequency resources to minimize inter-cell interference, Coverage and Capacity Optimization (CCO) to adjust RF parameters, Automatic Neighbour Relationships (ANR) to manage neighbour lists, Mobility Load Balancing (MLB) to manage traffic loads between cells and Mobility Robustness Optimisation (MRO) to optimize the operation of handover procedures. Further details about "self-x" functionalities are given in [26]. These features can be instantiated as VNFs under the control of the CESC Manager, which collects aggregated CESC statistics and acts mainly on behalf of the SCNO, since this is the infrastructure provider. Depending on the SLA agreement between SCNO and VSCNO a certain degree of decentralisation can be left also to a tenant operator, which collects statistics of its own slice of resources. Conversely, closely related RRM techniques such as scheduling could be left in the physical small cell or be virtualized depending on the functional split (see, *e.g.*, [25]).

[23] The *Iu-h interface* defines the security architecture used to provide a secure, scalable communications over the Internet. The Iu-h interface also defines an efficient, reliable method for transporting Iu-based traffic as well as a new protocol (HNBAP) for enabling highly scalable ad hoc HNB deployment. Also see the description given in: *https://en.wikipedia.org/wiki/Home_Node_B* .

[24] See: *http://wireless-century.blogspot.gr/2012/12/s1-flex-mechanism.html#!/2012/12/s1-flex-mechanism.html.*

### 4.2.2. Overview of the "Light Data Centre" concept proposed by SESAME - Envisaged system implementation

The SESAME project envisages the combination of the MEC/NFV concept with Small Cell virtualization in 5G networks, enhancing the multi-tenancy support. The main concept of SESAME is the design and implementation of a CESC, able to support edge cloud computing. It foresees the split of the small cell into physical and virtual network functions, respectively Physical Network Function (PNF) and Virtual Network Function (VNF), enabling a multi-tenancy environment to support the Multi-Operator Core Network (MOCN) requirements. The hosting of VNFs related to small cell virtualised functions, as well as service VNFs is guaranteed by the design of an advanced modular micro server whose architecture and characteristics are optimized for the MEC and Small Cell environment.

The micro-server architecture is based on use of specialized System on Chips (SoCs) that embed 64 bits multi core CPUs and HW accelerators having the necessary extension to support hardware virtualization. Typical HW accelerators are related to networking and packet processing. In case of additional computational requirements (e.g. audio and video transcoding, security, crypto engines), the micro-server is able to connect via standard PCI Express (PCIe) interfaces, cards equipped with specialized HW accelerators and or FPGA components. High capacity storage requirements (e.g. for video caching) are addressed hosting disk controllers and related disks in the micro-server.

Each micro-server is composed of a multi-core ARMv8 CPU, memory, storage, hardware accelerators (i.e. GPU, FPGA, crypto chips, etc.) and at least two network interfaces. One network interface is used for the connection to the radio-head (SC PNF), the other one assures the connectivity to the backhaul network as well as to the other micro-servers of the CESC cluster.

The software baseline of a micro-server consists of:

- ARM bootloader
- Baseline device drivers
- Linux kernel, capable of running KVM on ARMv8
- Customized version of Qemu[25] with KVM support for ARM:
  - providing access to hardware accelerators (GPU, FPGA, crypto chips, etc.) for the virtual machine
  - optionally enabling the use of huge pages for memory allocation
- Libvirt[26] with CPU pinning support
- A virtual switch providing:
  - Accelerated VM2VM communication
  - Access to the LightDC, backhaul and external networks for each virtual machine
- OpenStack Nova[27] and Neutron agents[28]

The physical aggregation of a set of CESCs (CESCs cluster) gives the possibility to share the IT resources of each micro-server belonging to the cluster and enables the implementation of a micro scale virtualised execution infrastructure in the form of a distributed data centre, denominated Light Data Centre (Light DC), enhancing the virtualization capabilities and process power at the network edge.
The HW architecture of the Light DC envisages that each micro-server will be able to communicate with all others via a dedicated LAN guaranteeing the latency and bandwidth requirements needed for sharing resources. Such a clustering is achieved using an Ethernet switch, suitably configured for properly

[25] More related information can be found, *for example*, at: https://en.wikipedia.org/wiki/QEMU.
[26] For more related information see, *for example*: https://libvirt.org/.
[27] See, *for example*: http://docs.openstack.org/developer/nova/.
[28] See, *for example*: http://docs.openstack.org/admin-guide-cloud/networking_config-agents.html.

enabling the networking between CESCs (bandwidth management, VLAN separation, etc.). It provides also the backhaul connections to the operators Evolved Packet Core (EPC) and all the links to the management system of the SESAME platform.

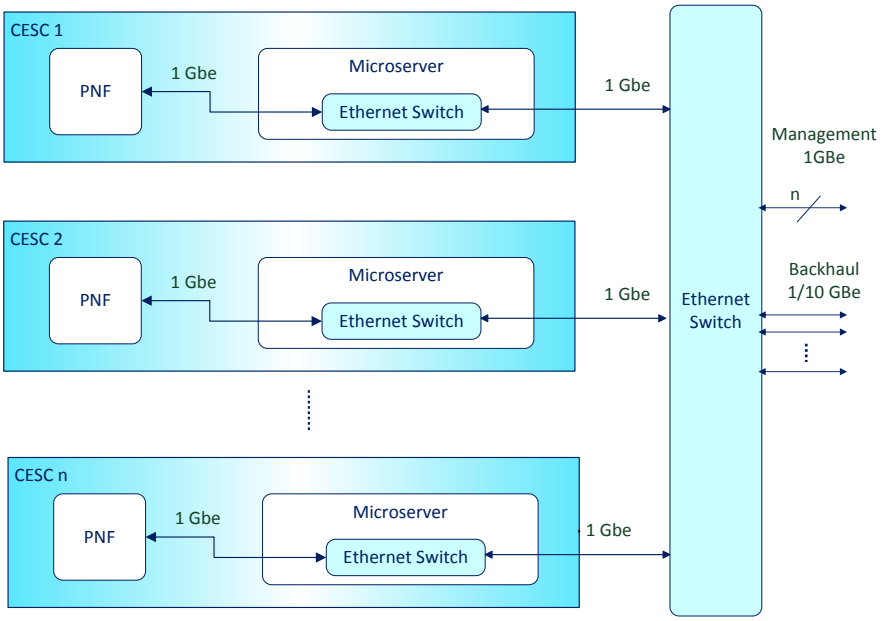



**Figure 20: Light DC Physical Architecture**

Network Services (NS) are supported by VNFs hosted in the Light DC, leveraging on technologies -like SDN and NFV- that allow achieving an adequate level of flexibility and scalability at the cloud infrastructure edge. In particular, VNFs are executed as virtual machines (VMs) inside the Light DC, which is provided with a hypervisor (a Kernel-*based* Virtual Machine (KVM)-*like* solution), specifically extended to support carrier grade computing and networking performance.
To achieve high performance at low cost and power consumption, the proposed Light DC architecture combined with the hypervisor computing extensions enable first to execute efficiently VMs on the underline ARM cores and on the other way to access HW accelerators using shared memory. Depending on the specific use case, the VMs are connected with these accelerators by using direct device assignment or any other runtime infrastructure. Furthermore, to offer high networking bandwidth and speed, each micro server runs a performance optimized user space virtual switch and leverages on specific zero copy shared memory extensions for the communication between VNFs co-hosted on the same Light DC (VM to VM).

### 4.2.3. CESC clustering

As mentioned before, one or more CESCs can form a CESC cluster, which can facilitate access to a broader geographical area with one or more operators (even virtual ones), extend the range of their provided services, while maintaining the required agility to be able to provide these extensions on demand, as shown in Figure 2.

More specifically, while each cell can provide a separate virtualised execution environment, through the application of a hypervisor, including PNFs and VNFs, the mentioned extensions can only be achieved by grouping together and creating the appropriate Service Function Chains (SFCs) under the necessary scope, which a tenant may apply through the CESCM. Thus the clustering of CESCs can provide the required flexibility and scalability of the network and computing resources as depicted in Figure 19.

The abstraction layer of the virtualized resources of the CESC cluster, however, is created and controlled by the VIM. Finally, in another perspective, the CESC cluster apart from being a cloud or pool of resources can be sliced as well to enable multi-tenancy.

### 4.2.4. Infrastructure Virtualization

VNFs in the SESAME architecture are executed on virtual machines distributed across the micro-servers of the CESC cluster. This implies computing hardware accelerators and network resources should be available for those VNFs to use.

While virtualization enabled hardware accelerators (i.e. some GPU) can be used by multiple virtual machines through direct attachment techniques, others, such as FPGA, require explicit concurrent access management. To provide this management, SESAME focuses on the development of specific FPGA virtualization extensions, allowing several virtual machines to access the underlying hardware via API remoting mechanisms.

Network access for the virtual machines is another matter of infrastructure virtualization. The software switch deployed on the micro-server allows VNFs in the same CESC cluster to communicate between them, connect to their respective VNF EMS as well as to the external world (i.e. VSCNO network) through the backhaul network. It also plays an essential role in setting up SFCs between the VNFs.
The software switch is in charge of the binding of the SC VNF to the physical network interface that connects it to the SC PNF.

### 4.2.5.CESC Manager

The Light DC concept offers a virtualisation platform to "meet" 5G requirements, namely NFV-*driven* small cell functions, NFV-*based* multi-tenancy and NFV-*based* MEC services. However, management and orchestration of this uniform virtualised environment, able to support both radio connectivity and edge services, is a quite challenging task by itself. From a pragmatic perspective, to better suggest a solution, we need to first understand the nature of VNFs and services under the pure scope of SESAME.
Two types of VNFs are identified under the scope of SESAME:

- *SC VNFs:* These are associated to the virtualised execution of different functionalities of the SC. SC VNFs may represent different layers of the E-UTRAN[29] protocol stack, while the rest of protocol entities will remain as PNF. Distributed self-x functionalities (i.e. dSON as explained in *Section 2.1.4*) can also be implemented as SC VNFs. The use of SC VNFs offers the required support for splitting the SC resources in different virtual slices provided to different tenants (multi-tenancy).
- *Service VNFs:* These are targeting the key performance indicator (KPI) of 5G, i.e., low latency through the proximity of end users and the serving nodes, service VNFs (e.g., virtual transcoding and virtual caching) are instantiated on the Light DC.

Over the provided virtualised execution environment (Light DC), it is possible to chain different VNFs to meet a requested network service (NS) by a tenant (i.e. mobile network operator). Note that, in the context of SESAME, a NS is understood as "a collection of VNFs" that jointly supports data transmission between User Equipment (UE) and operators' EPC, with the possibility to involve one or several service VNFs in the data path. Therefore, each NS is deployed as a chain of SC VNFs and Service VNFs.
It clearly highlights that, beyond the conventional orchestration and management of the cloud resources in a virtualised environment, the proposed solution entails a series of specific challenges such as the dynamic composition of the Light DC resources based on the status of CESC cluster(s), coordination of specific type of resources (radio-*related* resources, service-*related* hardware (HW) accelerators, etc.) and isolation of dedicated network slices to each tenant.

[29] More related infuriation can be found, *for example*, at: http://lteworld.org/ltefaq/what-eutran.

Having all these principles in mind, Figure 19 illustrates the envisaged high level architecture of the SESAME solution containing the main building blocks and their internal/external interconnections. The Small Cell Network Operator (SCNO) is the owner of the radio access infrastructure, and offers the sliced NS to the Virtual Small Cell Network Operator (VSCNO), which acts as tenant mobile network operator.

The CESC Manager (CESCM) is the central service management and orchestration component in the architecture. Generally speaking, it integrates all the traditional 3GPP network management elements, and the novel recommended functional blocks to realize NFV [37]. A single instance of CESCM is able to operate over several CESC clusters at different Points of Presence (PoP), each constituting a Light DC, through the use of a dedicated Virtual Infrastructure Manager (VIM) per cluster. Note that, due to the distributed nature of the Light DC, the proposed VIM requires data packet extraction (from the traditional 3GPP data path) and a forwarding rule implementation to guarantee possible communication between SC VNFs and Service VNFs (since they may reside in different CESCs). SDN principles are used to endow the system with the required scalability. In this way, the CESCM instructs the embedded SDN controller at VIM with the specific VNF forwarding rules, and the SDN controller in return applies them to support the desired connectivity within the Light DC.

For each instantiated VNF, an Element Management System (EMS), deployed in the CESCM, is responsible to carry out key management functionalities (such as fault monitoring, configuration, accounting, performance monitoring and security (FCAPS)). Meanwhile, the SCNO Network Management System (NMS) is the central management point for the whole network of the SCNO, while the PNF EMS and the SC EMS are respectively in charge of the management of the physical and virtualized network functions residing at the SC. In particular, the PNF EMS and SC EMS include different centralized self-x functionalities (i.e. cSON as explained in *Section 2.1.4*) to carry out the automated management of different radio parameters. All SCNO/VSCNO NMS communication is handled via a northbound interface provided in the CESCM.

The lifecycle management of deployed VNFs is carried out by the VNF Manager (VNFM) included in the CESCM. By leveraging on the monitoring mechanisms, the CESCM, in conjunction with the VNFM, is able to apply policies for NS-level rescaling and reconfiguration to achieve high resource utilization. It is worth to mention that monitoring mechanisms are dictated by the CESCM Service Level Agreements (SLA) monitoring unit that allows the evaluation of SLAs between different business role players, i.e. SCNO and VSCNOs. Under the scope of SESAME, SLAs are negotiated and agreed offline between the business role players and manually inserted to the CESCM SLA monitoring. Generally speaking, a generic SLA should be built upon a measurable set of Key Performance Indicators reflecting the characteristics of the provided service.

Another essential component at the heart of CESCM is the NFV Orchestrator (NFVO). Besides management and orchestration of the abovementioned functionalities, NFVO composes service chains (constituted by two or more VNFs located either in one or several CESCs) and manages the deployment of VNFs over the Light DC. This includes not only the management of a typical Network Function Virtualization Infrastructure (NFVI) (i.e. processing power, storage and networking), but also assignment of HW accelerators. Moreover, NFVO may include features to enhance the overall system performance, e.g. to improve energy efficiency a CESC resources switch on and off procedure may be introduced at the NFVO level. This element is detailed in *Section 4.4.*

The CESCM Portal, a web-based GUI for both SCNO and VSCNOs to access management and orchestration functions without any integration work, has been also been included in the architecture. The CESCM Portal supports two login procedures. A login for the VSCNO tenants aims to provide SLA monitoring information and an entry point to browse NS and VNF catalogues. Another login for the SCNO administrator to register extra resources and add new VNFs to the CESCM catalogues. In this scenario, any request for a new VNF by a tenant needs to be directly submitted and communicated with the SCNO administrator. After an agreement between both parties, the VNF will be implemented and registered to the VNF catalogue by the administrator and can be inserted on the service chains.

A more detailed description of the CESCM key features, internal modules and interfaces can be found in the respective SESAME *Deliverable D2.4* [26].

## 4.3. Compliance and Mapping to 3GPP, ETSI NFV and SCF

### 4.3.1. Relationships with SCF

SESAME and SCF are working in closely related subjects, so that different relationships can be identified. Both SCF and SESAME consider the split of the small cell functionalities between PNF and VNF. In SESAME, this is being analysed in the context of *WP3* (*Tasks 3.1* and *3.3*) and it is explained in more details in the Deliverable *D2.3* [25]. In SCF, this discussion can be found in [51].

Regarding the implementation of the VNFs, while in SESAME the VNFs are executed locally at the Light DC (i.e. at the computational resources of a cluster of small cells), the SCF considers a centralized small cell.

From the small cell management perspective, EMS in SESAME follows the split between PNF EMS and VNF EMS for the physical and virtual parts of the small cell, *respectively*, which is in line with the SCF approach [53]. However, in addition to the small cell VNFs, SESAME considers also the so-called "service VNFs" that correspond to mobile edge service instances devoted to deploy virtualized service-*level* functions at the CESCs. This requires that the EMS is split into a third component that is in charge of managing these VNFs. More details about the management architecture in SESAME are presented in the Deliverable *D2.4* [26].

SESAME architecture supports multi-operator small cells with MOCN as a key functionality. In SCF, although multi-operator small cells are identified as a relevant capability, no specific technical considerations have been done yet about multi-tenancy in areas like virtualization, functional splits, management, SON functions, SLAs, etc.

SON functionalities in both SESAME and SCF can follow dSON, cSON or hybrid approaches. Besides, dSON can include components implemented as PNF and/or VNF. The analysis of the specific SESAME SON (also referred to as self-x) functionalities is carried out in *WP3* (*Task 3.2*), while SCF addresses these aspects in [53].

### 4.3.2. Compliance and Mapping to ETSI NFV ISG

The need to produce normative specifications enabling end-to-end interworking of equipment and services initiated a joint collaborative effort between IT and Networks industries in the *last quarter* of *2012*. To "drive" the work forward, the ETSI Industry Specification Group for Network Functions Virtualization (ETSI ISG NFV) was formed in *January 2013*. The idea was to investigate the requirements and architecture for virtualization of various functionalities within telecoms networks.
ETSI ISG NFV initially brought together seven leading telecoms network operators, including: AT&T, BT, Deutsche Telekom, Orange, Telecom Italia, Telefonica, and Verizon. Not long after, the ETSI ISG NFV community grew to over 230 individual companies, including many global network operators, telecoms equipment vendors, IT vendors, and technology providers.

By nature, ETSI ISG NFV is not a Standards Development Organisation (SDO) and its findings are delivered in the format of guidelines, focusing on the requirements and architecture specifications for hardware and software infrastructure. The final goal of recommendations is to make sure virtualized functions are developed and maintained appropriately. The ETSI ISG NFV outputs are openly published and shared with relevant standards bodies, industry Fora and Consortia. It guarantees a wider collaboration possibility and paves the way for a joint effort to address any misalignment.

The ETSI ISG NFV guidelines are the main ingredients in the Proof-of-Concept (PoC) efforts lunched by industries to address the technical challenges of NFV implementation under the open source mentality.

Having said that, the NFV recommendations are reviewed briefly in the following and the SESAME compliance with them is checked.

Figure 13 presents the high-level NFV framework envisaged by ETSI ISG NFV [27], where three main working domains are identified:

- *Virtualised Network Function (VNF):* the software implementation of a network function, executable over the NFVI.
- *NFV Infrastructure (NFVI):* available physical resources (computation, storage, networking) and the way that they are virtualised.
- *NFV Management and Orchestration (NFV MANO):* Orchestration and lifecycle management of physical and/or software resources that support the infrastructure virtualisation. That includes the lifecycle management of VNFs/network services.

Figure 13 illustrates the necessary functionalities for the virtualisation and the consequent interactions between elements. Note that, the recommendation does not specify the exact network functions that should be virtualised as that is a decision of the network owner.

The above presented ETSI ISG NFV guidelines seem well-aligned with the SESAME solution, since enough attention has been paid in the design phase of SESAME overall architecture to ensure its compliance with the ETSI NFV reference architecture. Figure 21 illustrates the mapping between ETSI ISG NFV functional entities and the SESAME work packages (WPs).

- ***WP4*** takes care of the NFVI development. The general idea is to support a substantial change on the architecture of the current small cells, from being only a wireless head to cloud-*enabled* device (i.e. CESC) equipped with, *e.g.*, novel processor architectures, Graphics Processing Units (GPU), Digital Signal Processors (DSPs), and/or Field-Programmable Gate Arrays (FPGA). To enhance the virtualisation capabilities and process power at the network edge, the proposed CESC clustering under the scope of SESAME targets the formation of a distributed data centre, denominated Light Data Centre (Light DC). Also these work package support activates developing a hypervisor (a Kernel-*based* Virtual Machine (KVM)-*like* solution) specifically extended to support carrier grade computing and networking performance.
- Service VNFs and small cell VNFs as well as their associated EMSs are designed and developed within the ***WP4*** and ***WP3*** framework, *respectively*. More specific, ***WP4*** looks for the service VNFs, targeting the key performance indicator (KPI) of 5G, *e.g.*, low latency through the proximity of end users and the serving nodes. Some example of the service VNFs are virtual video transcoding and virtual caching. SC VNFs detailed in ***WP3*** are associated to the virtualised execution of different functionalities of the SC. SC VNFs may represent different layers of the E-UTRAN protocol stack, while the rest of protocol entities will remain as PNF. The use of SC VNFs offers the required support for splitting the SC resources in different virtual slices provided to different tenants (multi-tenancy).
- Service, VNF and infrastructure descriptors will be developed mainly on ***WP5*** as part of the CESCM. This activity needs a close collaboration with ***WP3*** and ***WP4*** to pass service VNF descriptors and determine the definition/parameters of the SC VNF descriptors.
- The CESC Manager (CESCM) is the central service management component in the architecture. Generally speaking, it integrates all the traditional 3GPP network management elements, and the novel recommended functional blocks to realize NFV. Having said that, ***WP5*** takes care of CESCM and VIM development. A single instance of CESCM is able to operate over several CESC clusters, each constituting a Light DC, through the use of a dedicated VIM per cluster. Note that, due to the distributed nature of the Light DC, the proposed VIM requires data packet extraction (from the traditional 3GPP data path) and a forwarding rule implementation to guarantee possible communication between SC VNFs and Service VNFs (may reside in different CESCs). SDN principles are used to endow the system with the required scalability. In this way, the CESCM instructs the embedded SDN controller at VIM with the specific VNF forwarding rules, and the SDN controller in return applies them to support the desired connectivity within the Light DC.

- ***WP6*** develops the NVF Orchestrator (NFVO) as the essential component at the heart of CESCM. NFVO is responsible for service chain composition (constituted by two or more VNFs located either in one or several CESCs) and the appropriate deployment of VNFs over the Light DC. This includes not only the management of a typical Network Function Virtualization Infrastructure (NFVI) (i.e. processing power, storage and networking), but also assignment of HW accelerators. Besides, to improve the energy efficiency of the proposed solution, NFVO may need to take care of switching on and off resources at CESC level.

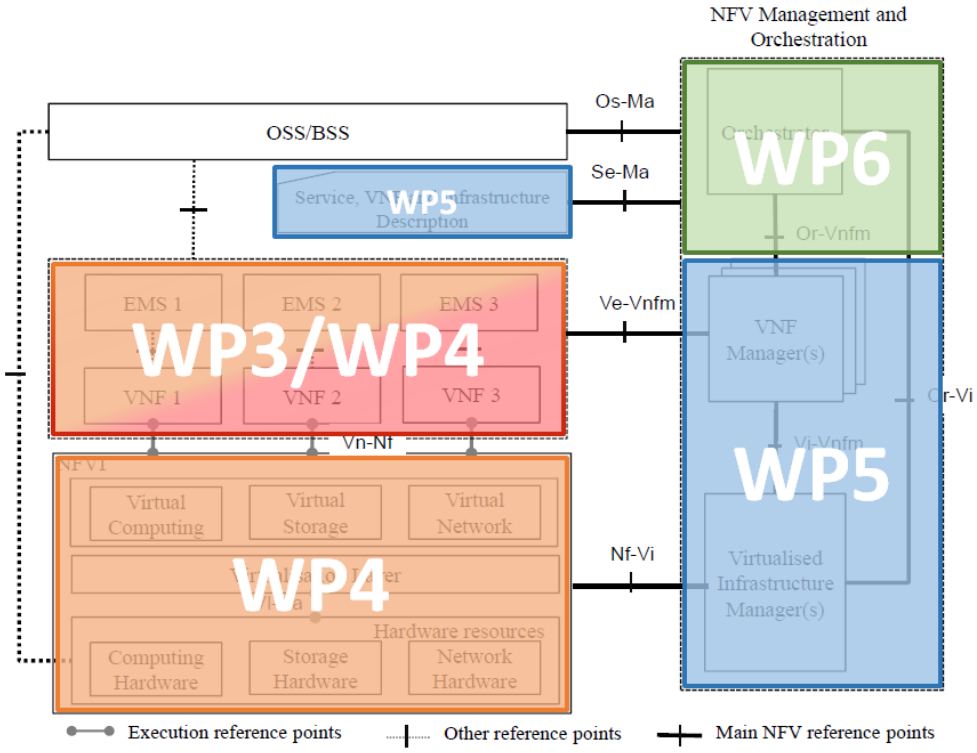



**Figure 21: Mapping between ETSI ISG NFV functional entities and SESAME WPs**

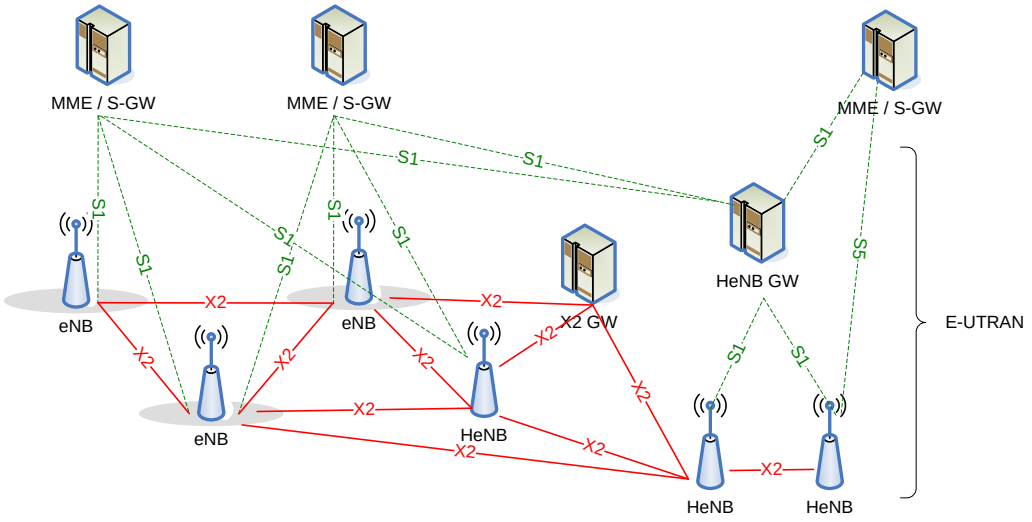



**Figure 22: LTE architecture**

### 4.3.3.Compliance and Mapping to 3GPP

The 3GPP architecture for the LTE Radio Access Network [13] has been stable for a number of years with each release of the standards building on this stable architecture and incorporating more advanced features. From a RAN perspective the base Stations (eNB) have separate connections into the network for control and user data (S1-C and S1-U) to the MME and S-GW, *respectively*. In addition, eNBs may be interconnected through an X2 interface to handle functions such as handover and interference coordination.

SESAME is particularly focused on a cluster of small cells that provide user access services either running at the edge of the network as part of the light DC or centrally either as an operator supported service or as a service running beyond the core network. In order to maintain compatibility with the existing LTE network architecture, it can be assumed that the CESC would need to support the S1 Control and User Plane protocols towards different operator networks for multi-tenancy support. However, this should not prevent the project from proposing any updates to existing standards if relevant to the ongoing research. The 3GPP architecture identifies specific functionality for the different architecture nodes which will directly be supported by 3GPP protocol architecture (Figure 23).

SESAME is targeting a Cloud-*enabled* small cell (CESC) infrastructure integrating small cell and distributed cloud technologies. The project itself is not focusing on the radio protocols themselves but recognises that virtualisation is a key technology that may impact the way in which the communications functionality integrates into the CESC infrastructure. Mapping the 3GPP communications functionality onto the CESC architecture will help identify interaction points between the communications and cloud based functions.

**3GPP RAN Protocols**

The generalised protocol architecture for the 3GPP eNodeB is described in Figure 24 below, showing the protocol stacks for the control and user planes.

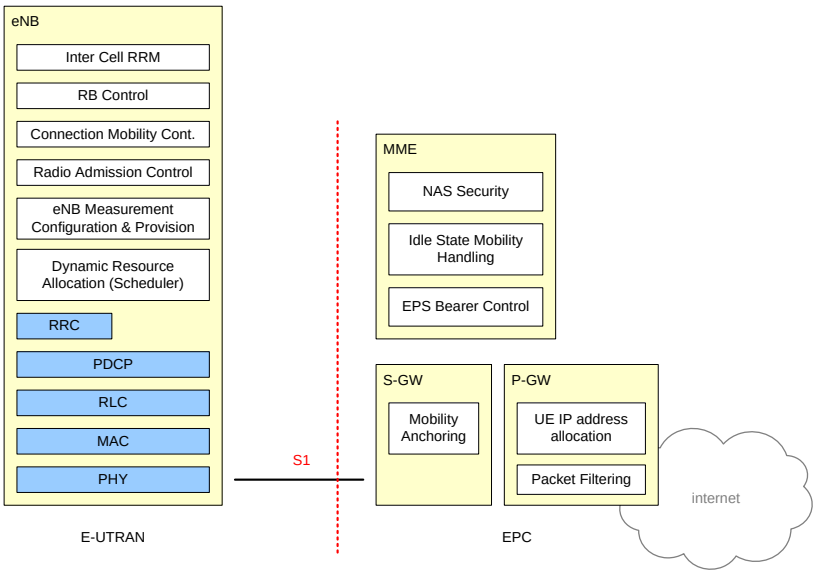



**Figure 23: 3GPP RAN nodes with Function allocations**

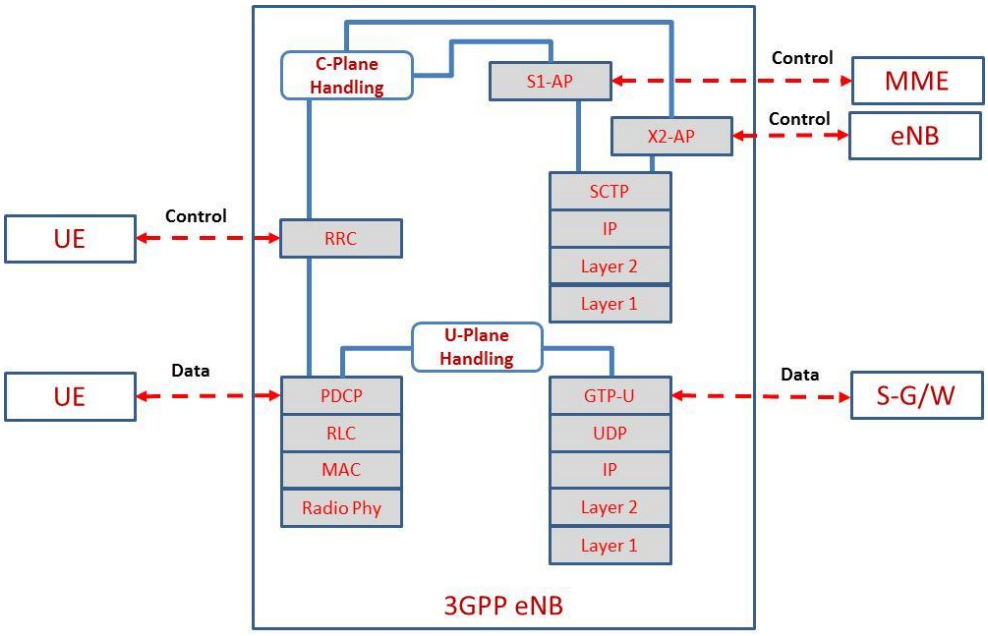



**Figure 24: 3GPP LTE based eNB showing protocol stacks**

From the perspective of the radio interface, we can assume compliancy with a 3GPP defined interface within the scope of the project. As the 5G radio interface becomes available the SESAME architecture should be able to accommodate this. From the SESAME perspective the interface towards the mobile core network has greater impact. A SESAME small cell cluster will at least in an initial deployment connect towards the core next through the S1-C and S1-U interfaces. Generally the control plane messages will be encrypted via a security gateway towards the MME. The user plane should at least during the scope of the project be based on the GTP-U tunnelling protocol from the SESAME cluster to the 3GPP Core Network. SESAME WP3 will need to consider this in the task of the Small Cell design.

**3GPP Mobile Network Sharing**

One of the challenges for the SESAME project is to support multi-tenancy in a small cell, light data centre cluster. Within the scope of 3GPP, a number of studies have been carried out focusing on the "sharing" of networks by multiple mobile operators. In this case the radio network will broadcast the list of PLMN-IDs for the mobile operators that share the RAN infrastructure. Currently this is limited to sharing by 6 mobile operators.

From the SESAME viewpoint, the most significant specification in 3GPP is that of the *Multi-Operator Core network (MOCN)* in which the radio access network is shared between the different mobile operators [14]. This is shown in the figure below.

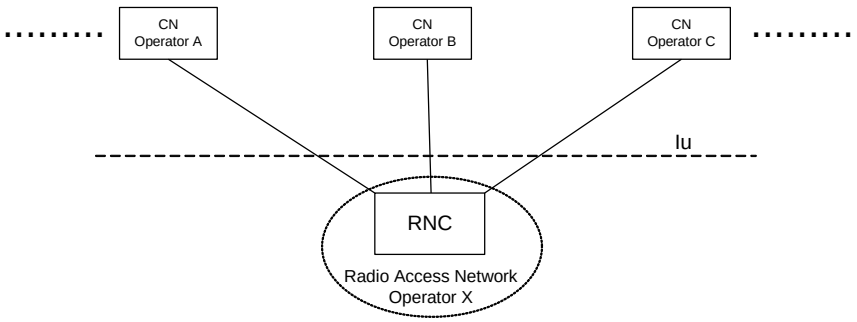



**Figure 25: RAN network sharing based on MOCN [14]**

The scope of SESAME is wider than "simply network sharing", as we are also concerned with the sharing of light data centre resources within the small cell cluster. Nevertheless sharing on the RAN network resources could use MOCN as a basis. This again will mainly impact the design of the SESAME small cell in *WP3*.

### 3GPP Network Management

The concluded 3GPP SA5 study item (i.e.: TR 32.842, as considered in [7]) has identified the following aspects regarding the management of mobile networks that include virtualized functions.

- Fault management.
- Configuration management.
- Performance management.
- Core network lifecycle management.

TS 28.500 [10], presents the management concept, architecture and requirements for mobile networks with virtualized network functions. The management requirements are organised according to the four management categories extracted from the study item. The TS 28.500 also presents a management architecture that provides a mapping between 3GPP and the ETSI NFV-MANO framework (Figure 26). The management architecture was designed for mobile networks composed of both physical and virtualized network elements.

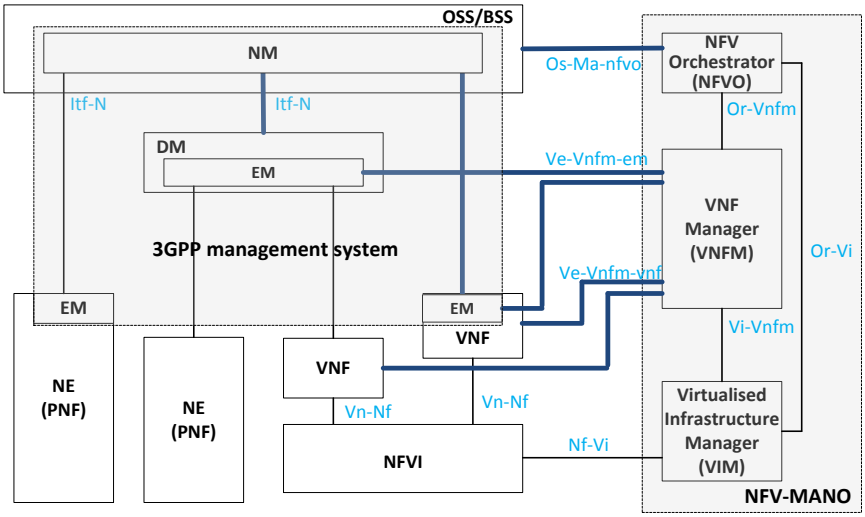



**Figure 26: The mobile network management architecture mapping relationship between 3GPP and NFV-MANO architectural framework**

Given the fact that the work items addressing each of the management categories only commenced in *November 2015*, the SESAME project should monitor its activity and assess the impact as the project progresses. In particular, *WP6* in SESAME will monitor any impact from 3GPP SA5 activity.

## 4.4. NFV Orchestration

From a general perspective, SESAME NFVO consists of the following components/functionalities:

- Interfaces to the external systems (i.e. CESCM elements);
- Service function chain lifecycle manager;
- Service function chain monitoring;
- Service function chain scaling ;

In the following a high level explanation of each one of them is presented:

### 4.4.1.Interfaces

Interfaces are crucial components that enable NFVO/CESCM components interactions. Under the SESAME context, interfaces of the NFVO are categorized as:

- *Northbound interfaces:* to support data exchange between CESCM elements (except catalogues) and NFVO.
- *Southbound interfaces:* To provide monitoring/communication line between the SESAME NFVO and virtual resources on CESC cluster (service VNFs and SC VNFs). These metrics are vital ingredient for service level performance monitoring and scaling decision making.
- *Westbound interface:* To support data exchange between local catalogues and NFVO.

Some basic features of interfaces are listed below:

- *Low latency:* To minimize the response time, especially once an error condition is detected or a threshold condition is exceeded.
- *Scalability:* To support different volume of NFVO/external systems interactions.
- *Resiliency:* To support information delivery even in the occasion of infrastructure failure or performance degradation (due to overload).

### 4.4.2.Service function chain lifecycle management

This functionality takes care of manage/provision of network services, over the Light DC. Based on ETSI NFV, network service is defined as a forwarding graph of interconnected (virtual) network functions and end points (that includes both functional and behavioural aspects of a network service).
Generally speaking NFVO needs to face two important challenges:

- Manage/coordinate the resources from different VIMs, when there are multiple VIMs in same or different PoPs.
- Manage/coordinate the creation of an end to end service that involves VNFs from different VNFMs domains.

These challenges are overcome by the following activities:

- *Resource orchestration:* It means coordination, authorization, releasing and engaging CESC cluster resources among different PoPs or within one PoP. These functionalities are handled by a direct communication with VIM(s) through the southbound interface APIs.
- *Service orchestration:* It means creation of end-to-end (E2E) service among different VNFs (that may be managed by different VNFMs – can be instantiated when/where it is necessary/applicable). Note that service orchestration is achieved via coordination with respective VNFM(s), not a direct communication with VNFs.

NFVO internally runs the VNF placement algorithms. Such an algorithm addresses the following challenges:

- Dynamic placement and elastic management of VNF over virtual nodes.
- Virtual resource provisioning and efficient local and cross-CESC site placement algorithms to meet agreements of service. That includes HW resource allocation (i.e. the general purpose micro server CPU, storage and networking resources as well as available HW accelerator – Graphics Processing Units (GPU), Digital Signal Processors (DSP), and/or Field-Programmable Gate Arrays (FPGA)).

### 4.4.3.Service function chain monitoring and scaling

Monitoring service-*related* metrics is another main responsibility of NFVO. Individual VNF monitoring data collection/processing as well as forming a service-*level* monitoring data is the main responsibility of service monitoring unit. Also, this unit includes interfaces to CESCM local catalogues to store all service-*level* monitoring information.
Note that service-*level* monitoring data may include inter-/intra- CESC cluster connection qualities. To support real time monitoring this unit needs a direct communication to the VIM(s).

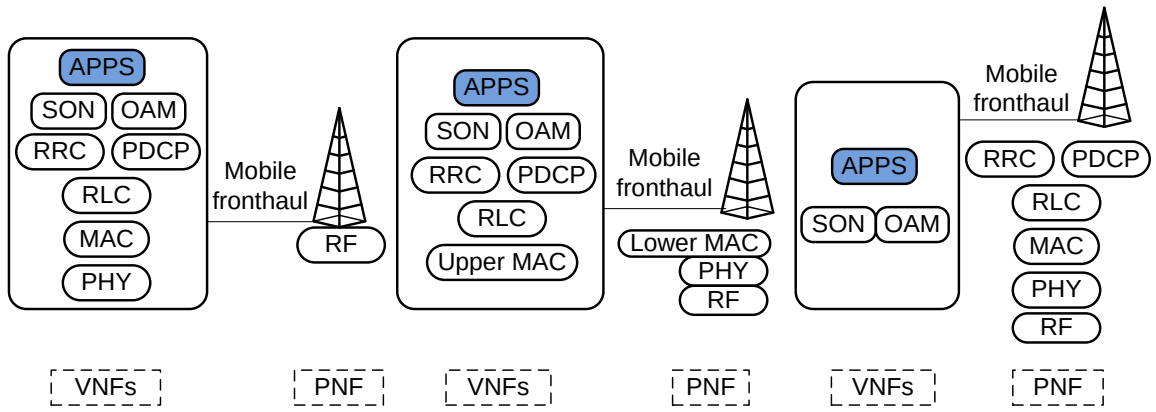
Under the SESAME context, service-level scaling up/down functionality is handled automatically whenever SLAs are violated. In this occasion, through a communication with the other CESCM elements appropriate actions can be performed based on a predefined scaling policy.

## 4.5. SESAME environment for introducing mobile edge services

The proposed "cloudification" of the Small Cells, represented in the novel CESC elements in SESAME, also allows reusing the available HW infrastructure for deploying service instances at the edge of the mobile network.
One of the emerging technologies to cope with more personalized and user-centric service provisioning is the novel Mobile Edge Computing (MEC) industry initiative [28], a promising approach to solve these types of problems from an operator-supported perspective. This initiative proposes that mobile network operators would provide an API to third-party partners, offering them access to critical features such as location awareness and network context information. This information may be exploited to deploy proximity-enabled services with close-to-zero latency characteristics, in order to optimize the management of future mobile networks.

The main novelties and innovations in SESAME are achieved by placing intelligence in the network's edge through the employment of Network Functions Virtualisation (NFV) and edge cloud computing directly to the Small Cell part. The SESAME proposed use of CESCs as a new multi-operator enabled Small Cell integrates a virtualised execution platform (i.e., a micro-server) for deploying Virtual Network Functions (VNFs) and also to execute novel applications and services inside the access network infrastructure.
The consolidation of multi-tenancy in communications infrastructures is also in scope, allowing several operators/service providers to engage in new sharing models of both access capacity and edge computing capabilities.

By enhancing Small Cells with an execution infrastructure, the inclusion of mobile edge computing capabilities emerges and this leads to increasing responsiveness from the edge of the network. This allows enriching the end users' experience and at the same time, operators can open the radio network edge to third-party partners, allowing them to rapidly deploy innovative applications and services.
In SESAME, MEC-*driven* service instances must be deployed over the cloud resources available at the RAN side, represented by the Light DC.




a) fully virtualized SC b) partially virtualized SC (1) c) partially virtualized SC (2)

**Figure 27: Fully centralized vs. partially centralized RAN functional architecture**

Figure 27 shows different possible splits of a CESC from a functional perspective. By taking the 4G LTE protocol architecture as the "reference" it is shown how different SC functions can be deployed as VNFs in the micro-server or as PNFs embedded in the SC firmware. Regardless the functional split of the CESC in PNFs and VNFs, the mentioned MEC services would be deployed as "APPS" features through VNFs running in the Light DC.
In other words, some capability of each CESC micro-server will be devoted to running the SC virtualized functions (SC VNFs in SESAME terminology), while the additional capacity will be made available for running mobile edge service instances (Service VNFs in SESAME terminology) devoted to deploy shared service-*level* functions within the CESC cluster.

In this sense, from the Service VNF standpoint, the service instances will be orchestrated at cluster level to be deployed in the distributed micro-servers of the Light DC as illustrated in Figure 28.

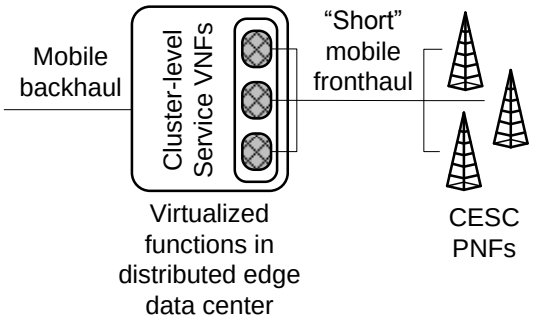



**Figure 28: Physically distributed Light DC running cluster-level Service VNFs**

The way to implement these types of mobile edge services in an environment – such as the proposed in SESAME – highly depends on the functional split of the SC. Taking into account the current 3GPP LTE architecture presented in Figure 27, the implications of the LTE data plane and the possible solutions to implement the SESAME data path are different.

Figure 29 illustrates the evolution of the eNB taking into account a functional split where the GTP tunnels are ended in the SC PNF part.

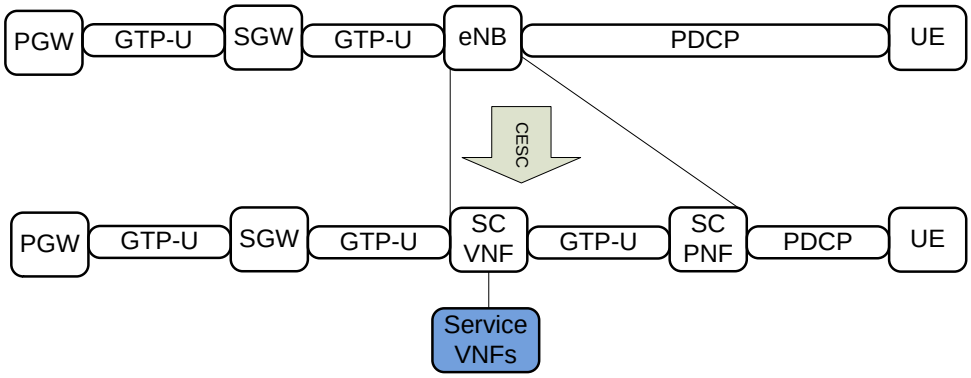



**Figure 29: Data path in typical LTE connection and in SESAME approach**

The data path resulting from the standard 3GPP LTE network architecture is depicted in Figure 29 (upper plot). All the user data traverses the Evolved Packet Core (EPC) through dedicated tunnels based on the GPRS Tunneling Protocol (GTP) protocol. These GTP-U tunnels exchange user data packets from the Packet Data Network Gateway (PGW) to the eNB, going through the Serving Gateway (SGW). The eNB is able to decapsulate the IP packets from the GTP-U tunnel and to forward them towards the corresponding User Equipment (UE).

The standard data path entails two main issues: First, all the IP data must traverse the EPC to the PGW for its forwarding to external applications. Second, taking into account the protocol split proposed in Figure 29, the VNFs are not able to access the IP packets without breaking the GTP-U tunnel.
Different solutions exist to cope with both problems. For the first problem, the 3GPP has introduced in *Release 10* the concept of local breakout. Two main solutions are currently available: Local IP Access (LIPA) and Selected Internet IP Traffic Offload (SIPTO) [5], [6], [1].

Both options allow offloading some user traffic from the PGW through the addition of a local gateway with different alternative architectures. The SIPTO at the local network (SIPTO@LN) solution is the "best candidate" for Small Cells, the local breakout is not fully integrated into the Home eNB (HeNB).
In the scope of the SESAME architecture, the applications are designed to run within the Light DC following the MEC concepts. Therefore, using local breakouts within the Small Cells would entail significant advantages both from the management perspective and from the standpoint of reduced latency in the data path. The desired operation environment (lower plot in Figure 29) would entail that the VNFs of the Small Cell (SC VNFs) are capable of handling the data path to/from the "Service VNFs" without going out of the CESC cluster.

The proposed solution for the local data path management in SESAME is illustrated in Figure 30, and involves the coordinated use of the CESCM and a SDN controller.

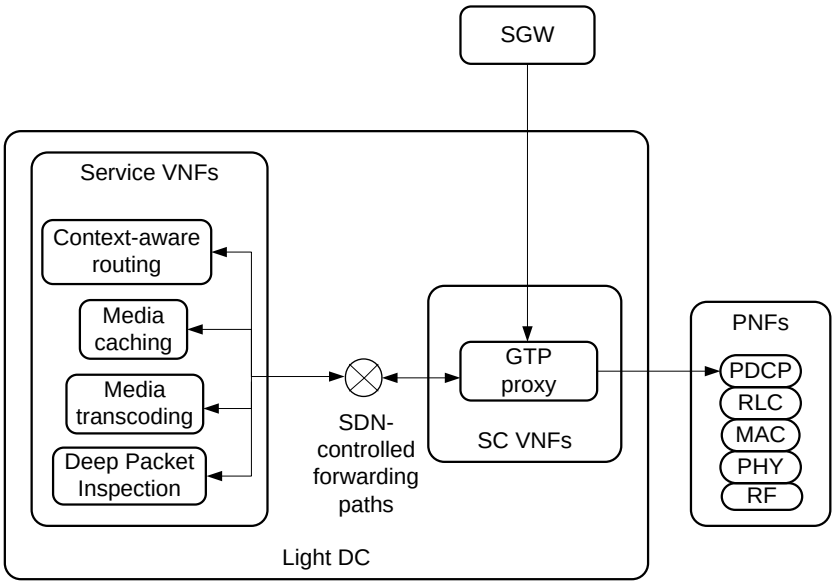



**Figure 30: SESAME solution for mobile edge traffic forwarding within the Light DC of the CESC Cluster**

The CESCM is the SESAME component that is in charge of the deployment, monitoring, configuration and administration of the mobile edge services instantiated over the hardware of the Light DC (resulting from the aggregation of the CESCs). A single CESCM can generically operate over a cluster of CESCs and, *consequently,* acquire an overall knowledge of the status of the virtual and physical resources across the different infrastructures.

The main functional components that build SESAME services are VNFs and it is therefore the role of the SESAME Orchestrator (in particular the NFVO located within the CESCM) to manage their deployment over the virtualized infrastructure offered by the VIM. This deployment of VNFs and their interconnection into composed edge services (the virtual infrastructure exposed by the VIM is made of the edge-*deployed* CESCs) is made with the awareness of different parameters and metrics that ensure an optimal allocation of compute resources and an efficient use of the radio access network in scenarios where end-users demand high performance and multiple network operators share the same infrastructure.

Building over these principles, the CESCM obtains the ability to dynamically compose edge services and apply real-time monitoring procedures that allow the enforcement of the agreed SLAs between the different entities (CESC operators, mobile network providers, end-users).

A fundamental role in the interconnection of the deployed VNFs is played by the SDN controller as a key entity to enforce the rules of traffic steering and chaining. The emergence of the NFV concept and the expansion of VNF solutions have materialized service function chaining (SFC) among virtualized functions as a legitimate use case for a cloud data centre.

The advantage of SDN is that based on the Open Flow (OF) protocol, the routing can be steered over a specific networking path by programmatically applying OF-*based* rules (flows) on the SDN controller or the virtual switch (OVS) inside the VM hosting the VNF. This approach to implement SFCs is based on flow programming rules, that leaves the packets untouched while applying actions on the OF ports of the switches, in order to gear the desired route of the packets in the chain.

# 5. Service Lifecycle and Sequence of Interactions

## 5.1. Service lifecycle

In SESAME, we have defined a set of high level workflows that show the creation, modification and termination of a service instance.
In this section we delve a little deeper into the actual lifecycle that encompasses the creation and deletion and points towards a lifecycle that will be equally applicable to VNFs (atomic services) and that of composed services.

In order for a service to be technically managed the following life cycle phases need to be accommodated:

- **Design:** Design of the architecture, implementation, deployment, provisioning and operation solutions. Supports Service Owners to "design" their service.

- **Implementation:** of the designed architecture, functions, interfaces, controllers, APIs, etc.

- **Deployment:** creation of the required resources. Supply of anything such that the service can be used, but does not provide access to the service. For example placing a VM image on an IaaS and creating an instance from it. This is part of the creation phase in the high level workflows.

- **Provisioning:** Provisioning of the service environment (e.g. VNFs, interfaces, network, etc.). Activation of the service such that the EU can actually use it. This is part of the creation phase in the high level workflows.

- **Operation and Runtime Management:** Activities such as scaling, reconfiguration of SICs happen here. This is part of the modification phase in the high level workflows.

- **Disposal:** Release and destruction of SICs and the SI itself and therefore all related resources. This is part of the termination phase in the high level workflows.

These are the phases that should take part during both the request for creation and deletion of a service or network function instance. Depending on the context, one or more of the phases may have no actions or logic associated. These phases are also applicable to a service composition that involves service function chaining. One of the key advantages that this approach has is the separation of deployment (creation) from provisioning (configuration). By doing this the creation of multiple resources and/or services can be accomplished in parallel with a “join-point” reached when all resources and/or services are created, where the resolution of configuration parameters can be carried out and then acted upon during the provisioning phase. This leads to minimised delivery times of the requested VNF or Service, especially those involving composition and function chaining.

## 5.2. Service composition

In order to aid our discussions we define:

- A Resource to be any physical or virtual component of limited availability within a computer or information management system.
- A VNF to be to be an orchestrated set of resources, generally residing in the same administrative domain and uses function chains to connected the resources and manage the network between them.

- A Service to be a delivered VNF composed with supporting BSS and OSS tools and/or services.
- A Service Composition to be the act of bringing existing services together through IT processes and software to deliver their functionality in support of value added functionality delivered by the service provider who requires the composition.

***Typically used Approaches***
There are two main means of delivering a composition (Figure 31).

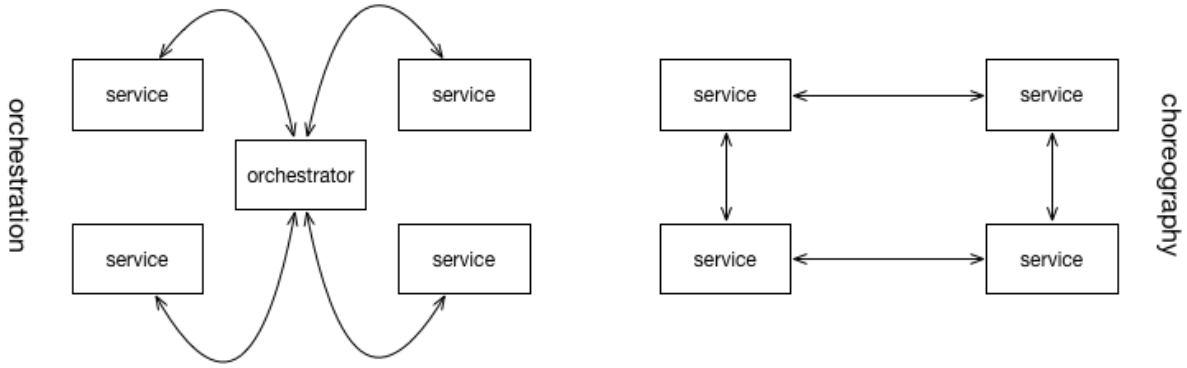



**Figure 31: Means of delivering a composition**

Those means are:

- **Orchestrated:** Orchestrated compositions take a centralised approach, where a process carries out the logic to deploy and provision the required dependencies. This is typical of systems like those based in WS-BPEL [61] and other orchestrators like Juju [33] or Heat [31].
- **Choreographed:** Choreographed compositions take a decentralised approach, where each service in the composition knows about its immediate collaborators. This knowledge is supplied to the system as initial seeding information. The deployment and provisioning of each part is carried out locally to the dependency. This is typical of systems like those based on WS-Choreography [62] or WS-CDL [63].

Whereas the general approach to composition is the orchestration in many enterprise scenarios the choreography approach has largely taken hold in academic contexts.

## 5.3. Definition of "SESAME network service", composition and offering through the CESCM Dashboard

***Service Function Chaining***
The emergence of the NFV concept and the expansion of VNF solutions have materialized service function chaining (SFC) among virtualized functions as a legitimate use case for a cloud Data Centre (DC). Yet in a premature state, applying SFC concepts in a fully virtualized environment requires changes and adaptations on the existing protocols in order to be able to apply the same concepts and achieve the desired behaviour on a network level. A typical burden in environments with fully virtualized functions running on virtual endpoints is the aggregated protocol encapsulation in the packets' headers added as they traverse those endpoints.

Currently, there are two key approaches to implement SFC solution in virtualized scenarios: packet based and flow based. The first requires manipulation of the packets, for instance by introducing some changes in the header field (packet tagging or rewrites) [54] or simply by applying protocols that introduce one more layer of abstraction on the top of the existing header fields – designed especially for this type of

service. Such dedicated protocols have been ultimately introduced by Cisco and leveraged in the Open Daylight (ODL) community to support the ODL SFC integral project. In this case, an additional header called Network Service Header (NSH), is introduced in order to enforce end-to-end traffic as an overlay connection above the service chain path. The problem of such solution is that it alters the datagrams and this can potentially cause a problem in the case where the VNF that runs on some of the virtual machine (VM) hops along the chain, requires the datagrams in their original structure. One example is a virtual function such as vDPI (virtual Deep Packet Inspection) that requires the packets in their original structure in order to enforce a correct behaviour.

The advantage of SDN is that, based on the Open Flow (OF) protocol, the routing can be steered over a specific networking path by programmatically applying OF based rules (flows) on the SDN controller or the virtual switch (OVS) inside the VM hosting the VNF. This is the second approach to implement SFC, based on flow programming rules, that leaves the packets untouched while applying actions on the OF ports of the switches, in order to gear the desired route of the packets in the chain. This routing logic is simplified compared to the first one, as it avoids unnecessary overheads on the top of the already existing ones (ex. in the scenario of inter tenant communication in OpenStack [38]) and the packets are left intact, completely agnostic of the existing chain. A solution based on this approach requires that the network environment is fully SDN capable in order to apply the chain rules along the full virtual network graph. The resulting routing flows need to be maintained to reflect alterations in the function chain (e.g. a VNF altering the packet header could invalid the end-to-end chain). Higher level chaining abstractions and programming languages are needed in order to allow service developers to programmatically declare the sequence the VNF traffic should follow, leaving to the underlying runtime system the actual implementation of such rules ([22], [23], [40]). For the chain routing to be deterministic, there has to be a field that keeps track of the chain hops. Using the VLAN ID as a workaround for this purpose could be one possible approach, since the chain routing does not follow the standard Ethernet routing.

Extending the concept of SFC for the scope of the SESAME project and in the context of the 5G-PPP ecosystem, requires deeper understanding on the NFV concepts in such a scenario and carefully elaborating the requirements to establish the desired functionality in environment that is not fully prepared to support it. Today there exist VNFs deployments as substitutes for the EPC integral blocks as well as on the radio access fronthaul side. SFC concepts based on SDN can be applied in LTE environment to enforce the packets among a logical network graph and achieve certain service functionality among the virtualized components such as BSS, HSS, RAN, etc. [41].
The role of the controller in this holistic approach has been analysed in the literature and some experimental and industry implementations already have been presented [32]-[34]. However, SFC solution based on SDN in this scenario is an unexplored area and potentially one of the use cases to be addressed in SESAME.

From a single data centre point of view, the SDN controller stands in between the components such as CESCM (VIM) and the light DC. In this case, the previously described approach for SFC applies if the SDN controller takes the charge of a single light DC deployment. The steering and rule enforcement policy are kept within the controller application logic and enforced over the network that hosts the light DC specific NFVs. If the routing happens across different light DCs, then the SFC approach may alter depending on the placement of the controller within the given architecture. This has to be in accordance with the networking protocols to be adopted in that case, for example MPLS and BGP as used today for intra data-centre routing. If SFC is established on GTP tagged traffic that enters the Light DC, then a component/function that removes the GTP header and extracts and reconstructs the IP packets is required as intermediator between the ingress/egress ports of the Light DC.

***Difference between Service Composition and Chaining***
The key difference of service composition and function chaining is at the level where the management and configuration happens, i.e. the control plane. The two carry out different tasks; however, their tasks are complementary, each needs the other. For composition the control plane manipulations are that at the level of the software that is deployed upon resources. For SFC, the control plane manipulations are those related to configuring the data plane (network) between two or more resources.

In order to compose and chain services/functions those need to be described as external dependencies. How this is to be described is a piece of work currently underway, however there are two identified candidates; TOSCA[30] and NFV-GD. With these inputs the NFVO can carry out the composition through orchestration lifecycle. Therefore, the composition of a service/function should also maintain the lifecycle as an individual atomic service/function.

A Composed Service aggregates/combines services together with orchestration logic. Both Atomic and Composed Services can be used to create further composed services.
Included in this process can be the complementary process of service function chaining that allows for the chaining of NFVs as well as services. How this can be achieved is shown in the next section.

***Composition and SFC***
Bringing composition and SFC as two ways to implement a specific service using several atomic services and network functions, is it essential to start with a top down approach when designing the use case scenario. This approach establishes valid business application model and identifies the building blocks of that application. Once defined the services and the network functions in the target scenario, a requirements analysis is needed to understand their specifications and certain limitations. Service function chaining among network functions enforces data-path rules among physical and virtual network endpoints (links and ports). Operating on layer 2 it should be agnostic of the functionality enforced by the VNFs.

As previously pointed out, depends on the specific implementation of the SFC rules that makes these requirements more or less relevant, but there is not yet a general solution that would be able to "predict and address" all types of application effects over the packets.
Taking as example a transcoding service/function that has specific requirements of the packets to be received (no VLAN/GRE traffic accepted) and changes the packet's IP and MAC addresses. This would essentially result in conflictive case for an SFC solution that manipulates virtual IP/MAC addresses. Moreover the packet has to be adjusted before sending it to the service in order to meet the specific header requirements by the given service. Whether this behaviour should be triggered by the SDN controller SFC application, other VNF in the chain, or the application itself in order to make the chain end-to-end valid is a matter to be further elaborated.

Such a scenario can be defined as a combination of chaining among specific network related VNFs and services operating on application level - a service composition. At a current stage there is not yet a strict limitation on the type of service a VNF can include. With this in mind a composite service can include several VNFs (regardless of the type of the service) with specific data plane instructions for the traffic flow rules and the order of the VNFs in the service).

[30] *Topology and Orchestration Specification for Cloud Applications (TOSCA),* is an OASIS (Organization for the Advancement of Structured Information Standards) standard language to describe a topology of cloud-*based* web services, their components, relationships, and the processes that manage them. The TOSCA standard includes specifications to describe processes that create or modify web services.For more related information see, *for example*, the description given in: *http://docs.oasis-open.org/tosca/tosca-nfv/v1.0/tosca-nfv-v1.0.html*

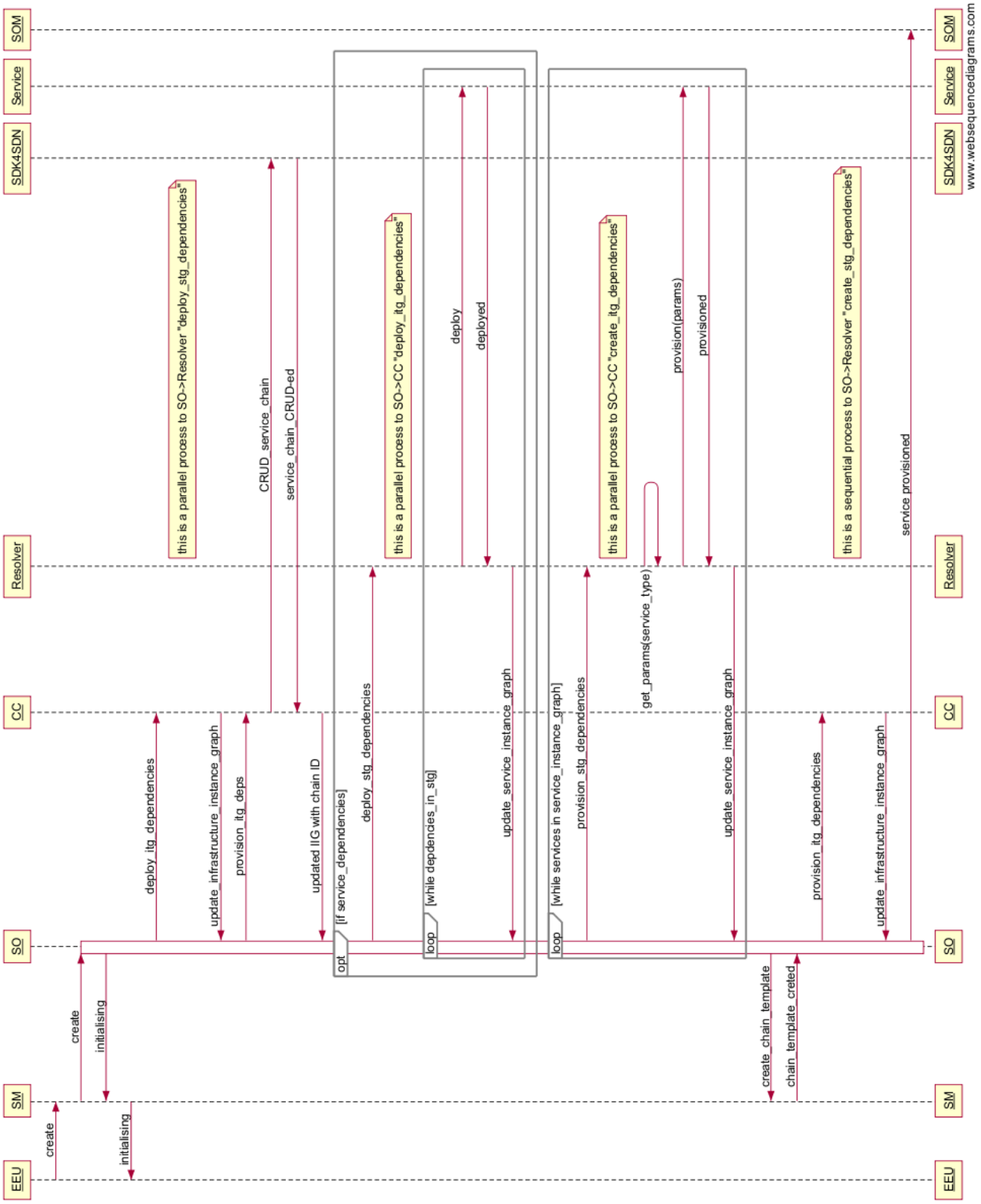



**Figure 32: Service composition workflow**

## 5.4. Management of QoE aspects

This section provides also definitions and relations for Deliverables *D2.3* [25] and *D2.4* [26]. One of the key aspects in future 5G networks will be the flexibility of providing different type of services over a shared network infrastructure, including the support of multiple tenants and different types of services per tenant.

Therefore, three main features are of utmost relevance to provide this capability to the external users:

- The availability to expose the capabilities of the network infrastructure to the external users.
- The capability by the external users to request and configure the required service levels.
- The capability by the network provider to properly propagate the heterogeneous service requirements to the underlying shared virtualized infrastructure.

In SESAME, these features are supported by different functional elements:

- The SCNO exposes the available network functions and network services through the CESCM portal.
- The VSCNO is able to select / create network services to be deployed, with specific KPIs (e.g., radio coverage, radio coverage for determined periods of time, maximum/average/minimum UL/DL bitrate/delay/packet loss ratios for the coverage area, etc.) or even with more QoE-*related* KPIs (based, *e.g.*, on the combination of temporally available bitrate and dedicated multimedia transcoding bitrates for the users of the VSCNO).These requirements will be formalized through SLAs between the SCNO and the VSCNOs, which depending on the granularity of the service may apply to aggregated or per-user control.
- Based on the available resources, the NFVO of the SCNO will map these requirements to the underlying resources. It must be noted that the network service requirements may include a combination of mobile network functions and mobile edge service functions, which need to be carefully orchestrated over all the VSCNOs.

As a result, the flexible configuration of the underlying infrastructure, including pre-defined and NFVO-*driven* scaling actions are a key functionality of the SESAME system.

Another relevant feature is the possibility to adapt the configuration of the Service VNFs to the actual performance of the mobile network at cluster level. In this way, the possible effects of potential temporal radio performance degradations are attenuated at QoE-level.

Taking into account the convergent worlds of small cell networks and virtualized network functions, SESAME will continuously apply a two-level management of the SLAs:

- From the perspective of traditional mobile network management, SESAME will introduce enhanced self-x functions that will facilitate the dynamic configuration of the underlying cluster of small cells. At QoE level, this may lead to improved balancing of service requests between different CESCs, to self-reconfiguration of certain radio parameters that better cope with the current traffic patterns, etc.
- From the perspective of traditional mobile network management, SESAME will be able to maximize the use of the underlying virtualized resources in a holistic way, taking into account both the network and service level requirements. In this sense, the VNF placement and scaling algorithms at the NFVO will be able to activate / deactivate the use of certain small cell resources, aiming at coping with the SLAs while minimizing the resource usage and energy consumption.

The terms of SESAME SLAs and interactions between SESAME actors are further defined in SESAME *D2.3*, while the role of the SLA Monitoring element and the different FCAPS operations are in the scope of SESAME *D2.4*.

## 5.5. SESAME Security Aspects

Network and system security is a very crucial issue because the SESAME system is expected to support both customer enterprises and end users, who cannot tolerate financial losses or data privacy violations, and therefore they seek the highest possible security guarantees.

Hence, the considered SESAME scenarios and functional components must be evaluated to identify and analyse security issues from the early stages of system design and software development, as well as to model and analyse threats and vulnerabilities in existing software and protocols that will be used in the SESAME system. The aim is to prevent a wide range of attacks, such as control hijacking, reverse engineering, malware injection, eavesdropping, *just to name a few*. At the same time, the innovative SESAME concepts and technologies could provide invaluable opportunities of developing novel solutions for attack prevention, management and recovery.

First of all, the physical security of CESC infrastructure and hardware integrity has to be ensured. Hence, appropriate security controls (such as [42]) should be deployed by the CESC infrastructure owner to prevent hardware tampering. Also, it is important to consider attacks that are initiated from the cloud side. This is particularly relevant to scenarios where multiple enterprises using private clouds are hosted. Especially in the multi-tenant environment of SESAME, the adversary *per se* could be a legitimate tenant interacting with network entities by using valid credentials and having privileged *access* to virtualised resources. Also, the emerging *Bring Your Own Device (BYOD)* trend[31] in many enterprises, constitutes many conventional security solutions incapable of protecting the private network. For example, a Trojan horse, that infected an employee's device, can bypass the security of the corporate firewall. Hence, the cloud provider must ensure the physical security of the cloud infrastructure and of the data centres. This can be done, e.g., by following the recommendations from the Cloud Security Alliance [24].
Besides, the selection of suitable cloud provider could be based on formal methodologies to ensure that the security and privacy requirements are met [30]. This effectively means that services offered by cloud providers who do not meet the specified requirements and have not implemented the mandatory security controls, could be restricted or even blocked.

To ensure confidentiality and integrity of UE data, the cryptographic security controls must be in place. This is to say that any adopted public-key scheme that enables the encryption of the communications among CESCM, CESC, UE, and the cloud, must be sufficiently secure. Cryptographic and privacy protection techniques are particularly important in cases where an end user receives service from multiple service or network providers due to mobility or QoE considerations.

An important category of attacks could potentially target the management system. For example, if initiated inside virtualised environments and aims at taking control of the CESCM. Also, the NFVO is an attractive attack target due to being in the "middle" of the system model architecture, as well as other components of the management layer, such as the VNFM. Also, impersonation by the adversary of one of the VNFs or the micro-server when communicating with the management layer could be a potential threat. Considering again the virtualised environment, both host and guest OS may be targeted, and to alleviate the impact of such an attack, adequate isolation must be enforced between guest Virtual Machines (VMs), as well as between the host and guest VMs.
The adversary could attempt to break the isolation by exploiting, e.g., some flaws of the used virtualisation platform [35]. Therefore, appropriate choice of the virtualisation platform that meets the security and privacy requirements is of major importance.

In some cases, to launch an attack against a component, the adversary requires that this component has specific exploitable configuration or runs aspecific software. For example, a precondition for a Denial-of-Service (DoS) attack can be specific configuration of the CESCM with regard to the allocation of resources to tenants. Yet, some flaws in the resource allocation algorithm can allow the adversary to prevent a tenant from accessing its portion of virtual resources.

[31] For more related information, see; *https://en.wikipedia.org/wiki/Bring_your_own_device*.

As security will be a fundamental enabling factor of future 5G networks, we are concerned with identifying and mitigating security threats and vulnerabilities against a broad range of targets at the intersection of SCaaS, NFV, and MEC.
These will have crucial effect on legal and regulatory frameworks as well as on decisions of businesses, governments, and end-users.

# 6. Architecture Extensions

The SESAME architecture envisioned in this deliverable can be extended based on the current trends towards 5G that are identified by the ETSI MEC standardization group. SESAME partner Athonet actively attends the MEC technical meetings and has recently submitted a proposal for decision with respect to a use case, which is of interest for SESAME.

Current virtualization technologies allow bringing mobile core functions close to the mobile edge, thus enabling deploying the service platform alongside the components of the EPC, while still being in proximity to users.

There are several advantages brought by this deployment, since the platform can leverage on many tasks performed by legacy core components (e.g., the PDN Gateway), thus without the need to implementing them. Such tasks are related to, among others, gating, GTP encapsulation, QoS enforcement [20], charging, lawful interception and mobility support.

The straightforward impact on the current SESAME architecture would be to deploy the virtual EPC (vEPC) [21] alongside with the virtual small cell and let the NFV components perform their tasks at the egress point of the core, i.e., the SGi interface[32], instead of the ingress point, i.e., the S1 interface.
In this way, as aforementioned, NFVs can take actions directly on IP packet flows without the need for implementing GTP encapsulation/decapsulation mechanisms as currently required in the SESAME architecture. Looking at the market trends, a further main advantage of the proposed extension architecture; however, is with respect to mobility support, meaning that moving applications might increase the impact of a terminal handover to another base station. Therefore, having a mechanism that supports session continuity for MEC applications is crucial in the design of future mobile system.

[32] See, for example: *http://lteworld.org/ltefaq/what-are-lte-interfaces*.

# 7. Conclusions

In order to address the needs and requirements of robust and agile network management, and building upon the pillars of network functions virtualisation, mobile-edge computing and cognitive management, SESAME's goal is the development and demonstration of an innovative architecture, capable of providing Small Cell coverage to multiple operators "as a Service". To that end, SESAME envisages to virtualise and to partition Small Cell capacity, while at the same time it aims to support enhanced edge cloud services by enriching Small Cells with micro servers. In the previous sections the envisaged use-case scenarios and the proposed system-level architecture have been presented.

This document presents a first approach to the high-level overall architecture of the SESAME system, the entities and the main interfaces/reference points. All SESAME partners contributed to this endeavour, achieving consensus among the consortium members on the initial architectural vision.

Using the overall architecture as reference, the project can proceed to the next tasks, which are the detailed definition of the SESAME layers and subsystems (to be contained in deliverables *D2.3* and *D2.4*) as well as the initiation of the implementation phase. Using an iterative approach, the feedback received from the detailed subsystems' design and specification as well as from the early phases of implementation will help to refine and amend the overall architecture as well.
The outcomes of this refinement will be reflected in the Deliverable *D2.5*.

# 8. References

[1] 3GPP TR 22.828 Study on co-ordinated Packet data network GateWay (PGW) Change for Selected IP.
[2] 3GPP TR 22.864: FS_SMARTER - Network Operation.
[3] 3GPP TR 22.891: Study on New Services and Markets Technology Enablers.
[4] 3GPP TR 23.799: Study on Architecture for Next Generation System.
[5] 3GPP TR 23.829: Local IP Access and Selected IP Traffic Offload (LIPA-SIPTO).
[6] 3GPP TR 23.859: LIPA Mobility and SIPTO at the Local Network.
[7] 3GPP TR 32.842 v13.1.0: "Telecommunication management; Study on network management of virtualized networks (Release 13)", December, 2015.
[8] 3GPP TR 38.913: "Study on Scenarios and Requirements for Next Generation Access Technologies".
[9] 3GPP TS 23.251 v13.1.0: "Network Sharing; Architecture and functional description (Release 13)", March, 2015.
[10] 3GPP TS 28.500: "Telecommunication management; Concept, architecture and requirements for mobile networks that include virtualized network functions".
[11] 3GPP TS 32.101: "Telecommunication management; Principles and high level requirements (Release 12)", September, 2014.
[12] 3GPP TS 32.130 v13.0.0: "Telecommunication management; Network sharing; Concepts and requirements (Release 13)", January, 2016.
[13] 3GPP TS 36.300 v13.2.0: "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (EUTRAN); Overall description; Stage 2 Release 13)", December, 2015.
[14] 3GPP TS 23.251: :Network sharing; Architecture and functional description:
[15] 5G-PPP 5G-EX project, http://www.5gex.eu/.
[16] 5G-PPP 5G-NORMA project, https://5gnorma.5g-ppp.eu/.
[17] 5G-PPP COHERENT project, http://www.ict-coherent.eu/.
[18] 5G-PPP METIS II project, https://metis-ii.5g-ppp.eu/.
[19] 5G-PPP SPEED-5G project, https://speed-5g.eu/.
[20] Aqsa M., Junaid Q., Basharat A., Kok-Lim A.-Y., Ubaid U., QoS in IEEE 802.11-based wireless networks: A contemporary review, Journal of Network and Computer Applications, Volume 55, September 2015, pp.24-46.
[21] Basta, A., Kellerer, W., Hoffmann, M., Hoffmann, K., Schmidt, E.-D., "A Virtual SDN-Enabled LTE EPC Architecture: A Case Study for S-/P-Gateways Functions," in Future Networks and Services (SDN4FNS), 2013 IEEE SDN for , vol., no., pp.1-7, 11-13 Nov.2013.
[22] Anwer B., Benson T., Feamster N., and Levin D., Programming slick network functions. In Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research (SOSR '15). ACM, New York, NY, USA, (2015), Article 14 , 13 pages.
[23] Chourasia, S., Sivalingam, K.M., "SDN based Evolved Packet Core architecture for efficient user mobility support," in Network Softwarization (NetSoft), 2015 1st IEEE Conference on , vol., no., pp.1-5, 13-17 April 2015
[24] Cloud Security Alliance, Security Guidance for Critical Areas of Focus in Cloud Computing V2.1, Dec. 2009.
[25] Deliverable "D2.3: Specification of the CESC components – First Iteration", H2020 SESAME project, April, 2016.
[26] Deliverable "D2.4: Specification of the Infrastructure Virtualisation, Orchestration and Management – First Iteration", H2020 SESAME project, April, 2016.
[27] ETSI GS NFV-MAN 001 (V1.1.1): "Network Function Virtualisation (NFV); Management and Orchestration", December, 2014.
*[28]* ETSI Industry Specification Group Mobile-edge Computing, http://www.etsi.org/technologies-clusters/technologies/mobile-edge-computing.
[29] GSMA, "The Mobile Economy 2013", 2013 http://www.gsmamobileeconomy.com/GSMA%20Mobile%20Economy%202013.pdf.

[30] Mouratidis H., Islam S., Kalloniatis C., and Gritzalis S., "A framework to support selection of cloud providers based on security and privacy requirements," *J. of Syst. and Softw.*, vol. 86, no. 9, 2013, pp. 2276-2293.
[31] Heat Orchestration Template (HOT) Guide, OpenStack, March 2016, http://docs.openstack.org/developer/heat/template_guide/hot_guide.html
[32] Akyildiz I.F., Wang P., Lina S.C., "SoftAir: A software defined networking architecture for 5G wireless systems" in Computer Networks, vol. 85, no.C, pp.1-18, July 2015.
[33] Juju, Canonical Ltd., March 2016, https://jujucharms.com
[34] He J., Song W., Smart routing: Fine-grained stall management of video streams in mobile core networks, Computer Networks, Volume 85, 5 July 2015, pp.51-62.
[35] M.T. Hoesing, "Virtualization security assessment," *Inf. Secur. J.: A Global Perspective*, vol. 18, no. 3, 2009, pp.124-130.
*[36]* MCN Deliverable 2.5 - Overall Architecture Definition, Release 2, available at http://www.mobile-cloud-networking.eu/site/
[37] NFV Management and Orchestration - An Overview, GS NFV-MAN 001 V1.1.1, European Telecommunications Standards Institute (ETSI), 2014
[38] Overlay encapsulation in DC network http://docs.openstack.org/developer/neutron/devref/openvswitch_agent.html.
[39] PricewaterhouseCoopers, "We need to talk about Capex; Benchmarking best practice in telecom capital allocation", 2012.
[40] Riggio R., et al., "Programming Wireless Network Functions", in Proc. Of IEEE NOMS 2016, Istanbul, Turkey
[41] Riggio, R., Bradai A., Rasheed T., Ahmed T., Slawomir K., and Schulz-Zander J., Conference Paper, "Virtual Network Functions Orchestration in Wireless Networks", 11th International Conference on Network and Service Management (CNSM), Barcelona, IFIP/IEEE, November, 2015.
[42] S. Skorobogatov, "Physical attacks and tamper resistance," *Introduction to Hardware Security and Trust*, Springer New York, 2012, pp.143-173.
[43] SCF 017.06.01: "Market drivers for multi-operator small cells", January, 2016.
[44] SCF 019.06.01: "Regulatory issues for multi-operator small cells", January, 2016.
[45] SCF 066.05.01: "Enterprise SON use cases", November, 2013.
[46] SCF 069.06.01: "Enterprises and multi-operator small cells", December, 2013.
[47] SCF 077.05.01: "SON use cases", February, 2014.
[48] SCF 083.05.01: "SON API for small cells", March, 2015.
[49] SCF 106.06.01: "Virtualization for Small Cells: Overview", June, 2015.
[50] SCF 154.05.01.02: "Virtualization in small cell networks. Translating NFV Concepts to SCN Functions", June, 2015.
[51] SCF 159.06.02: "Small Cell Virtualization: Functional Splits and Use Cases", January, 2016.
[52] SCF 160.05.1.01: "Coverage and Capacity Impacts of Virtualization", June, 2015.
[53] SCF 161.06.01: "Network Aspects of Virtualized Small Cells, June, 2015.
[54] Service Function Chaining in Open Daylight using NSH protocol: https://wiki.opendaylight.org/view/Service_Function_Chaining:Main.
[55] Small Cell Forum, http://www.smallcellforum.org/.
[56] Technical Specification Group Services and System Aspects; Network Sharing; Architecture and functional description, 3GPP TS 23.251, V10.1.0, March 2011.
[57] TM Forum, https://www.tmforum.org/.
[58] TR-069, "CPE WAN Management Protocol (CWMP)", Broadband Forum.
[59] TR-196v2, "Femto Access Point Service Data Model", Broadband Forum.
[60] UNIFY Project, D2.2 "Final Architecture", 15.11.2014: http://www.fp7-unify.eu/files/fp7-unify-eu-docs/Results/Deliverables/UNIFY%20Deliverable%202.2%20Final%20Architecture.pdf.
[61] Web Services Business Process Execution Language Version 2.0, OASIS Standard, OASIS, April, 2007
[62] WS Choreography Model Overview, W3C, W3C working draft, March 2004, https://www.w3.org/TR/ws-chor-model/.
[63] WS Services Choreography Description Language Version 1.0, W3C, W3C Candidate Recommendation 9, November 2005, https://www.w3.org/TR/ws-cdl-10/.