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 D7.5

Overall Assessment and Roadmap

Document Number:	H2020-5GPPP-GA No.671596/*WP7/D7.5/31.12.2017*
Contractual Date of Delivery:	*31.12.2017*
Editor:	Maria Belesioti, Evangelos Sfakianakis, Ioannis Chochliouros – Hellenic Telecommunications Organization S.A. (OTE)
Work-package:	*WP7*
Distribution / Type:	Public (PU) / Report (R)
Version:	1.0
Total Number of Pages:	90
File:	*SESAME_Deliverable 7.5_v1.0_Final*

Abstract

This document assesses the overall success of the SESAME Proof of Concept (PoC) demonstrator and, more specifically, how it fulfils the related identified KPIs.

The deliverable aims to "assess" the work that has been performed while implementing and/or deploying the corresponding SESAME architecture into the selected use cases for demos.

The deliverable has been structured upon the basis of the fundamental innovative features of the related SESAME architecture. We have included a suitable approach so that to be able to realize an overall system assessment. Furthermore, we have included a description of the fundamental selected demos together with a roadmap for potential adoption of the SESAME architecture. The document has also included a QoE assessment framework.

5G-PPP Disclaimer:
This *Deliverable* has been prepared by the 5G Initiative, via an inter 5G-PPP project collaboration. As such, the contents represent the consensus achieved between the contributors to the report and do not claim to be the opinion of any specific participant organisation in the 5G-PPP initiative or any individual member organisation of the 5G-Infrastructure Association.

Version History

Version	Date	Comments, Changes, Status	Authors, contributors, reviewers
0.1	20.11.2017	Initial draft by OTE	Maria Belesioti
0.2	21.11.2017	NFVO inputs by i2CAT	Pouria Sayyad Khodashenas
0.3	01.12.2017	Monitoring assessment input by ATOS	Elisa Jimeno
0.4	08.12.2017	Inputs by IPA	Alan Whitehead
0.5	12.12.2017	Inputs by ITL	Antonino Albanese, Claudio Meani, Pietro Paglierani
0.6	12.12.2017	Inputs by UoB	Konstantinos Kosmidis, Haris Mouratidis
0.7	12.12.2017	Updated version and inputs by OTE	Maria Belesioti
0.8	13.12.2017	Inputs by ZHAW	Irena Trajkovska
0.9	14.12.2017	Inputs by EHU	Begoña Blanco, Fidel Liberal
0.10	14.12.2017	Updated version by OTE	Maria Belesioti
0.11	17.12.2017	Inputs/revision by ATN	Daniele Munaretto
0.12	18.12.2017	Inputs by NCSRD	Ioannis Giannoulakis
0.13	1912.2017	Inputs by ORION	Emmanouil Kafetzakis
0.14	19.12.2017	Updated version and inputs by OTE	Maria Belesioti, Evangelos Sfakianakis
0.15	20.12.2017	Inputs by UPC	Jordi Pérez-Romero, Oriol Sallent
0.16	21.12.2017	Inputs by FLE	Mick Wilson, Hui Ziao
0.17	21.12.2017	Inputs by VOSYS	Michele Paolino
0.18	22.12.2017	Inputs by OTE	Evangelos Sfakianakis, Ioannis Chochliouros
0.19	22.12.2017	Inputs by CNET	Cristina Costa
0.20	27.12.2017	Inputs by SMNET	Athanassios Dardamanis
0.21	31.12.2017	Inputs and editorial review by OTE	Ioannis Chochliouros
0.22	08.01.2018	Editorial review by SMNET	Athanassios Dardamanis
1.0	12.01.2018	Final full version of the deliverable – Submission to the Commission	Ioannis Chochliouros

Contributors

First Name	Last Name	Partner	Email
Maria	Belesioti	OTE	mbelesioti@otereseacrh.gr
Evangelos	Sfakianakis	OTE	esfak@oteresearch.gr
Ioannis	Chochliouros	OTE	ichochliouros@oteresearch.gr
Eirini	Vasilaki	OTE	evasilaki@oteresearch.gr
Pouria Sayyad	Khodashenas	i2Cat	Pouria.khodashenas@i2cat.net
Elisa	Jimeno	ATOS	Elisa.jimeno@atos.net
Jordi	Pérez-Romero	UPC	jorperez@tsc.upc.edu
Oriol	Sallent	UPC	sallent@tsc.upc.edu
Konstantinos	Kosmidis	UoB	K.Kosmidis@brighton.ac.uk
Haralambos	Mouratidis	UoB	h.mouratidis@brighton.ac.uk
Antonino	Albanese	ITL	antonino.albanese@italtel.com
Claudio	Meani	ITL	claudio.meani@italtel.com
Pietro	Paglierani	ITL	Pietro.paglierani@italtel.com
Michele	Paolino	VOSYS	m.paolino@virtualopensystems.com
Begoña	Blanco	EHU	begona.blanco@ehu.eus
Fidel	Liberal	EHU	fidel.liberal@ehu.eus
Hui	Xiao	FLE	Hui.Xiao@uk.fujitsu.com
Mick	Wilson	FLE	Mick.Wilson@uk.fujitsu.com
Ioannis	Giannoulakis	NCSRD	giannoul@iit.demokritos.gr
Emmanouil	Kafetzakis	ORION	mkafetz@orioninnovations.gr
Daniele	Munaretto	ATN	daniele.munaretto@athonet.com
Cristina	Costa	CNET	ccosta@fbk.eu
Athanassios	Dardamanis	SMNET	adardamanis@smartnet.gr
Irena	Trajkovska	ZHAW	traj@zhaw.ch
Alan	Whitehead	IPA	alan.whitehead@ipaccess.com
Fidel	Liberal	EHU	fidel.liberal@ehu.eus
Begoña	Blanco	EHU	begona.blanco@ehu.eus

Glossary

Acronym	Explanation
2FA	Two-factor authentication
3GPP	Third Generation Partnership Project
4G	Fourth Generation of Mobile Communications
5G	Fifth Generation of Mobile Communications
AC	Admission Control
AC	Alternating Current
AGW	Access Gateway
AHP	Analytical Hierarchy Process
AI	Artificial Intelligence
ANR	Automatic Neighbour Relation
API	Application Programming Interface
AR	Augmented Reality
ARM	Advanced RISC Machine
BGP	Boarder Gateway Protocol
BF	Broadband Forum
CCIS	Communications in Computer and Information Science
CESC	Cloud Enabled Small Cell
CESCM	CESC Manager
CM	Configuration Management
CP	Control Plane
CPE	Customer Premises Equipment
CPU	Central Processing Unit
cSON	centralised Self-Organized Network
CSP	Cloud Storage Provider
CSP	Communications Service Provider
CWMP	CPE WAN Management Protocol
D2D	Device-to-Device
DC	Data Centre
DC	Direct Current
DDR	Double Data Rate
DDR3	Double Data Rate type three
DDR4	Double Data Rate fourth-generation
DNS	Domain Name System
DoS	Denial of Service
DP	Data Plane
DPDK	Data Plane Development Kit
DPI	Deep Packet Inspection
DSL	Digital Subscriber Line
DSP	Digital Signal Processing
DVD	Digital Video Disc
E2E	End-to-End
EANN	Engineering Applications on Neural Networks
EMS	Element Management System
EPC	Evolved Packet Core
ETSI	European Telecommunications Standards Institute
EuCNC	European Conference on Networks and Communications
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FA	Factor Authentication
FG	Forwarding Graph
FF	Fast Forwarding

FM	Fault Management
FMC	Fixed Mobile Convergence
FPGA	Field Programmable Gate Array
fps	frames per seconds
FTTH	Fiber-to-the-Home
GA	Grant Agreement
GB	Giga Bytes
GbE	Gigabit Ethernet
GDP	Gross Domestic Product
GHz	Giga Hertz
GPRS	General Packet Radio Service
GTP	GPRS Tunnelling Protocol
GPU	Graphics Processing Unit
GS	Group Specification
GUI	Graphics User Interface
GW	Gateway
H2020	Horizon 2020
HD	High Definition
HEVC	High Efficiency Video Coding
HO	Handover
HOT	Heat Orchestration Template
HTTP	Hypertext Transfer Protocol
HW	Hardware
I/O, i/o	Input/Output
IB	Information Block
ICIC	InterCell Interference Coordination
ICT	Information and Communication Technology
ICMP	Internet Control Message Protocol
ID, id	Identifier
IEEE	Institute of Electrical and Electronic Engineers
IETF	Internet Engineering Task Force
IoT	Internet of Things
IP	Internet Protocol
IPS	Intrusion Prevention System
IRP	Integration Reference Point
IS	Intermediate System
IS-IS	Intermediate System to Intermediate System
IT	Information Technology
ITU	International Telecommunication Union
ITU-T	International Telecommunication Union - Telecommunication Standardization Sector
KHz	Kilo Hertz
KPI	Key Performance Indicator
KVM	Kernel-*based* Virtual Machine
LAN	Local Area Network
LCD	Liquid Crystal Display
Light DC	Light Data Centre
LMDS	Local Multipoint Distribution Service
µs	micro-server
MAC	Mandatory Access Control
MAC	Medium Access Control
MANO	Management and Orchestration
MCC	Mobile Country Code
MEC	Mobile Edge Computing
MME	Mobility Management Entity

MNC	Mobile Network Code
MOCN	Multi-Operator Core Network
MP	Management Protocol
MPEG	Moving Pictures Experts Group
MSU	Moscow State University
MTC	Machine Type Communications
NBI	Northbound Interface
NFV	Network Functions Virtualization
NFVI	Network Functions Virtualization Infrastructure
NFVO	Network Function Virtualization Orchestrator
NMS	Network Management System
NOS	Network Orchestration System
NS	Network Service
ODL	OpenDayLight
OPNFV	Open Platform for NFV
OS	Operating System
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First
OVS	OpenvSwitch
OWASP	Open Web Application Security Project
PC	Personal Computer
PCI	Peripheral Component Interconnect
PCIe	Peripheral Component Interconnect Express
PLMN	Public Land Mobile Network
PM	Performance Management
PNF	Physical Network Function
PoC	Proof of Concept
PPP	Public-Private Partnership
PS	Packet Scheduling
PS	Power Supply
PSNR	Peak Signal to Noise Ratio
QoE	Quality of Experience
QoS	Quality of Service
R/W, r/w	Read/Write
R&I	Research and Innovation
RAID	Redundant Array of Independent Disks
RAB	Radio Access Bearer
RAM	Random Access Memory
RAN	Radio Access Network
RFC	Request for Comments
RIA	Research and Innovation Action
RISC	Reduced Instruction Set Computer
RJ	Registered Jack
RRC	Radio Resource Control
RRM	Radio Resources Management
RTD	Research and Technical Development
SATA	Serial Advanced Technology Attachment
SC	Small Cell
SCaaS	Small Cell as a Service
SC-C-VNF	Small Cell-Common-VNF
SCF	Small Cell Forum
SCNO	Small Cell Network Operator
SDN	Software-Defined Networking
SFC	Service Function Chaining
SIB	System Information Block

SIM	Subscriber Identity Module
SLA	Service Level Agreement
SOAP	Simple Object Access Protocol
SoC	System on Chip
SON	Self-Organized Network
SOTA	State-of-the-Art
SQL	Structured Query Language
SRIA	Strategic Research and Innovation Agenda
SRIOV, SR-IOV	Single Root Input Output Virtualization
SS	Solution Set
SSIM	Structural Similarity Index
SSL	Secure Socket Layer
SUT	System under Test
SW	Software
TAC	Tracking Area Code
TAI	Tracking Area Identity
TCP	Transmission Control Protocol
TDE	Technology Development Envelope
TLS	Transport Layer Security
TR	Technical Report
TREC	Text Retrieval Conference
TRL	Technology Readiness Level
TS	Technical Specification
TU	Transcoding Unit
TV	Technology Value
UC	Use Case
UDP	User Datagram Protocol
UE	User Equipment
URL, url	Uniform Resource Locator
USB	Universal Serial Bus
UTRAN	Evolved Universal Terrestrial Radio Access Network
VA	Video Analytics
VGA	Video Graphics Array
VIM	Virtualised Infrastructure Manager
VM	Virtual Machine
VMM	Virtual Machine Monitor
VNF	Virtual Network Function
VNFM	VNF Manager
VSCNO	Virtual Small Cell Network Operator
vTU	virtual Transcoding Unit
VTU	Video Transcoding Unit
WAN	Wide Area Network
WiFi	Wireless Fidelity [An IEEE 802.11 protocols family]
WP	Work Package
WT	Web Toolkit
XSS	Cross-Site Scripting

Table of Contents

List of Figures

List of Tables

Executive Summary

The purpose of this document is to present -as well as to discuss- the results of the various validation tasks that have been performed in the context of the SESAME project (GA No.671596) in order to validate all RTD outcomes.

This deliverable is a report "depicting" the assessment of the technical work and the demonstrations made by all partners of the SESAME project during the project lifecycle, as these partners/beneficiaries have designed and prepared the final demos.

Building upon the specific activities that have been reported in the related SESAME deliverables, this document aims to "assess" the work that has been performed when implementing and deploying the corresponding SESAME architecture into the selected use cases, as the latter have been proposed in the first year (Y1) of SESAME and have been further developed in the continuity of the project. For this reason, we provide a global view of the architecture and of the related demonstrations.

The document is structured into six (-6-) distinct thematic sections:

- *Section 1* is the "introduction" of the document and outlines the purpose and the content of the document. It includes an overview of SESAME architecture and a "recap" of its main terms and concepts.

- *Section 2* ("*Summary of Assessment methodology"*) describes the approach that will be performed for assessisng of the outcomes of the project; this presents an overall system assessment as well as the "key achievements", per separate technical output.

- *Section 3* ("*Demonstration evaluation"*) presents the results taken from the demos that have been built around the selected use cases and the scenarios of the project.

- *Section 4* highlights a roadmap regarding the adoption of the SESAME architecture and beyond.

- *Section 5* includes a summary of the results described in this document.

- The document also includes certain bibliographic references.

- Furthermore, we have included -as *Section 6*- an Annex A which composes a part of informative nature and discusses the QoE assessment framework within SESAME.

1 Introduction

This deliverable provides an "assessment" of some among the SESAME project (Grant Agreement (GA) No.671596) activities and, more specifically, about its architecture and its demos results. The infrastructure employed for each demo and "how the proposed architecture is integrated into the respective infrastructure" is analysed; *additionally*, this document provides an assessment methodology for the validation of the architecture in each of the demonstrations in the SESAME designed use cases.

Within the wider SESAME context, a methodology for use case validation and architecture design verification has been established, in order to succeed an overall assessment. As a consequence, this document aims to assess the benefits of the SESAME-*based* architecture innovations, including approaches to "accommodate" the different network infrastructure (virtual) utilisation at different demos' environments.

1.1 Definitions of Terms and SESAME Concepts

This subsection provides definitions for the essential assessment methodology concepts and the interrelations among them. We have considered the following terms:

- **System Assessment:** It is the application of the methodological framework for the evaluation of the SESAME-based outcomes in three different test cases, in terms of complying with the technical and functional requirements.
- **Test Case: C**onsists on a demo devoted to provide a realistic testbed for the outputs of the SESAME project and proves its advantages in specific application contexts.
- **Key Performance Indicator (KPI):** KPIs are used to evaluate factors that are crucial to the success of the system under consideration.
- **Requirements:** These are related to the functional or the technical objectives that the correspondent system must fulfil, and so, they are classified as "functional" or "technical".
- **Metric:** It is a quantitative measure of the degree to which a system, component or process achieves a given objective or requirement.

In addition, in this subsection we also present a brief description of the main SESAME project concepts, relevant to the final aim of this deliverable:

- **Cloud-Enabled Small Cell (CESC):** It is a multi-operator enabled small cell (SC) that integrates a virtualized execution platform equipped with a micro-server to support the execution of VNFs inside the RAN.
- **Cloud-Enabled Small Cell Manager (CESCM):** It is the set of applications to manage the SESAME network services and the deployment and composition of VNFs.
- **Light Data Centre (Light DC):** This is the SESAME NFVI (Network Functions Virtualization Infrastructure), made up of the aggregation of the different micro-servers of each CESC included in a cluster. The concept of Light DC enables "running" VNFs of the different Virtual Small Cell Network Operators (VSCNOs) in different CESCs in a cloud computing fashion.
- **Monitoring:** Beyond the basic NFV monitoring and management, it is in charge of monitoring and management of some specific metrics related to the network slice in the RAN, including both radio and cloud parameters.

- **Network Function Virtualization Orchestrator (NFVO):** This is the entity in charge of taking the new network slice requests from the VSCNOs and "mapping" them to specific VNF instances and service chaining configurations in the Light DC.

- **Service Function Chaining (SFC):** It is a sequence of VNFs, either SC VNFs or service VNFs, connected through a VNF Forwarding Graph to define a certain Network Service (NS).

- **Virtual Network Function (VNF):** The SESAME project differentiates two types of VNFs, as follows: SC VNFs are the set of SC specific functions that are deployed as VNFs in the Light DC, and Service VNFs are the edge service instances that are also deployed as VNFs in the Light DC.

1.2 SESAME Architecture Overview

Figure 1, below, "depicts" the main components of the SESAME architecture for provisioning SCaaS (Small Cells as-a-Service) with Mobile Edge Computing (MEC) capabilities in multi-tenant environments, as it was defined in the respective SESAME Deliverable D2.5[1].
Speaking in general terms, SESAME scenarios assume a certain venue (e.g. a mall, a stadium, an enterprise, etc.) where a Small Cell Network Operator (SCNO) has deployed a number of small cells (SCs) to provide wireless access to end-users of different tenants, i.e. Virtual Small Cell Network Operators (VSCNO) through a SCaaS model.

The architecture of Figure 1 relies upon the Cloud Enabled Small Cell (CESC) concept, a new multi-operator enabled small cell that integrates a virtualized execution platform for executing novel applications and services inside the access network infrastructure.
A CESC actually consists of a Small Cell Physical Network Function (SC PNF) unit, where a subset of the SC functionality is implemented via "tightly coupled" software (SW) and hardware (HW), and a micro-server that supports the execution of VNFs, which provide the rest of the SC functionality together with other added-value services.

The CESCs support the Multi-Operator Core Network (MOCN) sharing model of 3GPP[2], which allows them to offer access over shared radio channels to multiple operators' core networks (CNs). Accordingly, each CESC is connected to the Evolved Packet Core (EPC) of each tenant through an S1 –C/U interface[3].

The physical aggregation of a set of CESCs, *denoted as a CESC cluster*, gives the possibility to "jointly" operate the computational, storage and networking resources of the micro-servers as a single virtualised execution infrastructure, denoted as the Light Data Centre (Light DC). Aiming to compose the Light DC, the different CESCs are connected to a centralized Ethernet switch.
In the specific SESAME prototyping environment, a micro-server consists of an Advanced Reduced instruction set computer Machine (ARM)-*based* System-on-Chip (SoC) including a multicore 64-bit ARM[4] Central Processing Unit (CPU) with integrated Hardware (HW) accelerators and integrated fabric for Input/Output (I/O) communication.

A number of VNFs can be executed in the Network Function Virtualisation Infrastructure (NFVI) constituted by the Light DC. The functionalities of a VNF are implemented in software modules that run on one -or more- Virtual Machines (VMs). The VMs are instantiated on the physical CPUs of the Light DC

[1] I. Giannoulakis (editor): *Deliverable D2.5: "SESAME Final Architecture and PoC Assessment KPIs"*. H2020/5G-PPP SESAME project, December 2016.
[2] For more details see: 3GPP TS 23.251: *Network Sharing; Architecture and functional description.*
[3] For further details also see: ETSI TS 136 410 v9.1.1 (2011-05): *LTE; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 general aspects and principles (3GPP TS 36.410 version 9.1.1 Release 9)*. Available at: http://www.etsi.org/deliver/etsi_ts/136400_136499/136410/09.01.01_60/ts_136410v090101p.pdf
[4] For more related information, see: http://www.arm.com/products/processors/armv8-architecture.php.

through a hypervisor[5] that partitions and abstracts the underlying physical resources. The use of VM allows hiding the HW infrastructure while offering the same sort of resources (processor, memory/storage, interfaces/ports) of a physical server to Software (SW) developers.

The subset of the SC functionality that is embedded in the SC PNF and the subset that is run externally as VNFs depend on the selected functional split. Different functional splits are discussed in SFC 159.06.02[6], following the layering architecture of the radio interface protocol stack. Each micro-server allows the execution of the VNFs with the SC functionality associated with the locally attached SC PNF as well as other VNFs devoted to provide virtualised service-level functions within the CESC cluster (e.g. deep packet inspection, caching, etc.).

The CESC Manager (CESCM) is the central service management component in the architecture that integrates the traditional 3GPP network management elements and the novel functional blocks of the NFV-MANO (Network Function Virtualization Management and Orchestration) framework[7]. Configuration, Fault and Performance management of the SC PNFs and VNFs is performed through the Element Management System (EMS). The EMS in SESAME is split in three parts, denoted as PNF EMS, the SC EMS and the EMS in Figure 1, which perform, *respectively*, the management of the SC PNF, the SC VNF and the service VNFs. In addition, the Service Level Agreement (SLA) monitoring takes information from the EMS to evaluate the level of conformance between the current service status and the Key Performance Indicators (KPIs) defined in an SLA.

As shown in Figure 1, the PNF EMS and SC EMS include the centralised Self Organized Network functions (cSON) and the centralised components of the hybrid SON functions. In turn, the decentralized SON (dSON) functions, or the decentralised components of the hybrid SON functions, reside at the CESCs.

Within the CESCM, the lifecycle management of the VNFs is carried out by the VNF Manager (VNFM), while the Network Functions Virtualization Orchestrator (NFVO) composes service chains constituted by one or more VNFs, running in one or several CESCs and manages the deployment of VNFs over the Light DC with the support of the Virtualized Infrastructure Manager (VIM).
The VIM creates an abstraction layer to the CESCM of all the physical resources in the Light DC, managing the lifecycle of VMs and providing the virtual links among VMs and VNFs. The VIM includes a Software Defined Network (SDN) controller in charge of setting up the virtual networks on the physical infrastructure, configuring the virtual switches in the micro-server so that the different VNFs can communicate.

The CESCM also includes a portal, which is a control panel web Graphical User Interface (GUI) that serves as the "entry point" for the users, both SCNO and VSCNO, to the CESCM and constitutes the main graphical front-end to access the SESAME platform.

---

[5] A hypervisor or Virtual Machine Monitor (VMM) is computer software, firmware, or hardware that creates and runs virtual machines. A computer on which a hypervisor runs one or more virtual machines is called a *host machine*, and each virtual machine is called a *guest machine*. The hypervisor presents the guest operating systems with a virtual operating platform and manages the execution of the guest operating systems. Multiple instances of a variety of operating systems may share the virtualized hardware resources: for example, Linux, Windows, and macOS instances can all run on a single physical x86 machine. This contrasts with operating-system-level virtualisation, where all instances (usually called *containers*) must share a single kernel, though the guest operating systems can differ in user space, such as different Linux distributions with the same kernel. More details can be found, *inter-alia*, at: *https://en.wikipedia.org/wiki/Hypervisor*.

[6] See: Small Cell Forum (SCF) 159.07.02: "*Small Cell Virtualization: Functional Splits and Use Cases*", January, 2016. Available at: *http://scf.io/en/documents/159_-_Small_cell_virtualization_functional_splits_and_use_cases.php*.

[7] For more details see the framework presented in: European Telecommunications Standards Institute (ETSI) (2014): "*NFV Management and Orchestration - An Overview", GS NFV-MAN 001 v1.1.1.* European Telecommunications Standards Institute, Sophia-Antipolis, France. Available at: *http://www.etsi.org/deliver/etsi_gs/NFV-MAN/001_099/001/01.01.01_60/gs_NFV-MAN001v010101p.pdf*.

Finally, SESAME provides a northbound interface[8] between the CESCM and each tenant operator's Network Management System (NMS). It is also used internally by the CESCM to provide the underlying capabilities of the CESCM Portal.

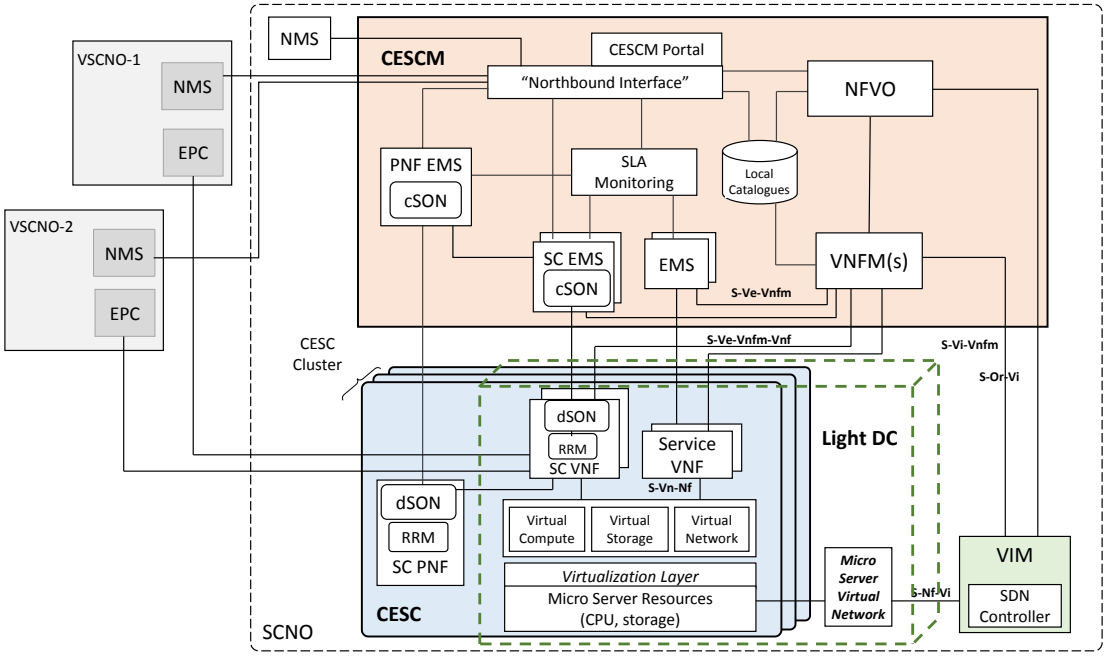

Figure 1: SESAME Overall Architecture

[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. Also, see *for example*: https://en.wikipedia.org/wiki/Northbound_interface.

2 System Assessment

2.1 Summary of Assessment Methodology

The objective of the SESAME assessment process is to assess/evaluate the performance of the various SESAME outcomes, in terms of complying with the technical and functional requirements.
The relevant evaluation process should take the into account that the SESAME-specific architecture will be evaluated in three different demos, appearing as "demo 1", "demo 2" and "demo 3".
The purposes of these demos are the provision of a realistic testbed for the outputs of the SESAME project and the demonstration of the actual advantages of these in specific application contexts.

A methodological framework is proposed for the SESAME assessment, taking into account the features described above. Under the proposed methodological framework, the evaluation of the SESAME outcomes involves the following major steps:

1. Determination of the objectives and the performance expectations for the SESAME outcomes.
2. Identification of measurable and not measurable requirements, for each one of the SESAME outcomes.
3. Identification of methods for assessing the achievement of the requirements.

The assessment of the SESAME outcomes emerges from the requirements specification performed within Task 2.1 in WP2, as explicitly described in the SESAME DoW. The proposed evaluation framework is based on the aforementioned demos, where the SESAME outcomes are thus evaluated and its use cases are assessed.

2.2 Overall SESAME Technology Assessment

The next generation of communication systems, as assessed within the "5G" context, will be the first illustration of a truly converged network environment where wired and wireless communications will make use the same infrastructure, thus "driving further" the future (global) networked society. It offers virtually ubiquitous, ultra-high bandwidth, "connectivity" not only to separate users but also to (Internet-) connected objects.
Therefore, it is assumed that the future 5G infrastructure will "serve" a wide multiplicity of services/applications and domains/sectors also including professional uses (e.g. assisted driving, eHealth, energy management, possibly safety applications, etc.). In fact, 5G is aiming to be reasonably different. It is about more than just "raising the bar" on previous generations or extending them to a certain context[9].
While it cannot be disagreed that wireless -via its successive generations- has brought huge socio-economic value almost beyond measure, no generation has ever set out with this important goal as a priority, as it actually happens for the 5G challenges.
In this view, 5G is quite diverse. It is the "first generation" -or the first organised effort- to clearly "target providing/offering socio-economic benefits" and this will be a "key-goal" to "guide-and-drive" the priorities of the many new 5G capabilities that are anticipated, by both modern economies and societies.

Thus, SESAME in the framework of 5G is not only a single "progression" or a "simple evolutionary step" of mobile broadband networks. Nevertheless, it is expected to "bring" innovative and exceptional network and service capabilities, in common with modern applications and related services/facilities. Mostly, safeguards user experience continuity in challenging situations such as high mobility (e.g. in trains), very "dense" or "sparsely populated" areas, and journeys covered by heterogeneous technologies. Moreover,

[9] Also see the discussion provided in: I.P. Chochliouros, A.S. Spiliopoulou, A. Kostopoulos, M. Belesioti, E. Sfakianakis, et *al.* (2017): *Putting Intelligence at the Network Edge through NFV and Cloud Computing: The SESAME Approach.* In G. Boracchi et *al.* (eds.), EANN 2017, CCIS 744, pp.704-715. Springer International Publishing AG.

mission critical services requiring very high reliability, global coverage and/or very low latency, which are up-to-now handled by specific networks, typically public safety, will become natively supported by the SESAME infrastructure.

In addition, SESAME aims to integrate networking, computing and storage resources into "one programmable and unified infrastructure". This sort of much promising "unification" allows for an optimized -and more enhanced- and fully dynamic usage of all distributed resources, as well as for the anticipated "convergence" of all "underlying" fixed, mobile and broadcast services. Within this scope, SESAME also supports multi-tenancy models, thus enabling operators and other market players to collaborate in new ways, enhancing opportunities for growth and development within a converged environment.

Furthermore, leveraging upon the features of existing cloud computing, SESAME supports further progress of the single digital market, e.g. by "paving the way" potentially for virtual pan-European operators relying on nation-*wide* infrastructures.

SESAME is designed in a way to be a sustainable and fully scalable technology. In view of this vision and aim, the telecommunications industry will compensate incredible usage evolution and growth by radical energy consumption reduction.
In addition, cost reduction through resource optimization allows for supportable business models for all ICT stakeholders to be involved in related actions.
Last but not least, SESAME generates an ecosystem for both technical and business novelty. Since network services rely more and more on software, the creation and growth of startups in the sector will be systematically encouraged. In addition, the 5G infrastructures will deliver network solutions and include vertical markets (such as automotive, energy, food and agriculture, city management, government, healthcare, manufacturing, public transportation, and so forth) this extending all potential beneficial uses.

SESAME within the 5G-PPP framework delivers solutions, architectures, technologies and standards for the ubiquitous 5G communication infrastructures of the next decade.

The following high level Key Performance Indicators (KPIs) are proposed to frame the research activities:

- Providing 1000 times higher wireless area capacity and more varied service capabilities compared to 2010.
- Saving up to 90% of energy per service provided. The main focus will be in mobile communication networks where the dominating energy consumption comes from the radio access network.
- Reducing the average service creation time cycle from 90 hours to 90 minutes.
- Creating a secure, reliable and dependable Internet with a "zero perceived" downtime for services provision.
- Facilitating very dense deployments of wireless communication links to connect over 7 trillion wireless devices serving over 7 billion people.
- Enabling advanced user controlled privacy.

The new 5G systems open new opportunities for efficient services in the business, administrative and private domain, which will make the societies and economies in Europe "more competitive".

Therefore, the development and provision and a new 5G communication infrastructure based on secure networks will be an essential prerequisite for positive economic effects in all sectors.

In this respect, 5G will be quite different. It will be the first generation to explicitly target delivering socio-economic benefits and, as a result, many new 5G capabilities are anticipated.

2.3 CESCM Assessment

2.3.1 CESCM Portal

Within the PoC, the CESCM Portal is provided by two separate interfaces[10]:

- The EMS portal that provides client access to the SC EMS and PNF EMS, *and*;
- The NFV Portal that provides access to the NFVO and VNFM functions.

2.3.1.1 EMS Portal

As previously stated in the Deliverable D3.4[11], in the SESAME PoC the SC EMS and PNF EMS are provided by a single client-server solution that is based on the ip.access Ltd. (IPA) Network Orchestration System (NOS).
It uses Java WebStart[12] to provide the downloading and starting of a GUI management client from a Web page published by the EMS server. Configuration Management (CM), Fault Management (FM) and Performance Management (PM) functions are provided to SCNOs and VSCNOs by the EMS client.
The set of managed objects that may be viewed and modified are controlled by access permissions granted to individual users.

In summary, the feature operates broadly as follows:

- Permissions to sub-sets of managed objects are assigned to one -or more- user groups.

- Permissions for an individual user are determined by the groups to which that specific user belongs to, and these are actually assessed as the "greatest" sort of permission provided by those groups. For example, if a user belongs to two groups, namely A and B, and the one group (e.g., the group A) provides read-only access to a specific managed object whereas the other group (i.e., the group B) provides read-write access, then the user has the "greater" of these two, which is the read-write access.

- By creating appropriate groups, the SCNO is able to:
  - Assign permissions to SCNO users who are able to manage all parts of the network, including those aspects belonging to VSCNOs and those belonging to the SCNO and not accessible to VSCNOs.
  - Assign permissions to VSCNO users such that they are only able to view their own "network slice" and are not able to view -or interact with- the managed objects of other VSCNOs or the SCNO.

In addition to EMS client users, a similar permissions system may be applied to users of the northbound configuration management system.
This is based on the SOAP[13] solution set of the Configuration Management IRP[14], [15],[16],[17]. Note, *however*, that whilst their permissions may be managed in a similar manner, EMS Client user and Northbound CM

[10] See: A. Whitehead (editor): *Deliverable D3.4: "CESC Small Cell prototype and PoC"*. H2020/5G-PPP SESAME project, June 2017.
[11] Ibid.
[12] For more details see, *for example*: WebStartRef - Sun Microsystems (Oracle). Also see: Java Web Start, available at: http://docs.oracle.com/javase/tutorial/deployment/webstart/.
[13] Simple Object Access Protocol. For further information see, *inter-alia*: https://en.wikipedia.org/wiki/SOAP.

users fall into two distinct sub-sets; i.e., an EMS client user may not use his credentials to access the northbound CM interface and *vice versa*.

The SESAME Deliverable D7.3[18] reported on those aspects of EMS testing that had successfully been completed and a small number of outstanding problems.
Those specific problems have now been corrected and EMS testing has been successfully completed.

In summary, the following EMS Portal achievements have been made:

- The EMS Portal allows VSCNOs to "access" the facilities of the EMS client via a graphical user interface and perform Configuration Management, Fault Management and Performance Management tasks such as:
  - Provisioning a new virtual cell , including selecting the service level agreement to be applied to that cell;
  - Temporarily taking a virtual cell out of service;
  - Decommissioning a virtual cell;
  - Monitoring the performance of defined sets of virtual cells against an agreed SLA.

- The EMS successfully implements network isolation such that one VSCNO is unable to view or interact with the managed objects of another VSCNO or of the SCNO. This isolation is provided by means of a set of user permissions that work in an "equivalent fashion" via both the EMS client and the EMS' northbound CM interface.

- Similarly, network isolation is provided via the northbound PM interface where each VSCNO is only able to "view" PM and KPI reports relating to their own network slice. Regarding the Access Gateway (AGW) we have to notice that whilst the PoC demo implements the proposed northbound file structures for PM and KPI reports, currently there is permissions mechanism to prevent one VSCNO from viewing another's data.

- The EMS provides business logic that "links" together the elements of the solution such that the configuration of a virtual cell is propagated to the associated CESC elements (such as the PNF, SC Common VNF and associated SC VNF).

- The EMS successfully provides the ability to represent an SLA as a managed object where the properties of this object specify the KPIs and associated thresholds of the SLA. The EMS SLA Monitoring Service parses PM data received for the PNF in order to: calculate KPI values; assess SLA compliance, and; raise an alarm in the event of SLA breach.

Areas for future research and enhancement include the following sections 2.3.1.2 and 2.3.1.3:

---

[14] 3GPP TS 32.306: "*Configuration Management (CM); Notification Integration Reference Point (IRP): Solution Set (SS) definitions (Release 10)*", 3GPP, September 2010. Available at: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1927.

[15] 3GPP TS 32.316: "*Generic Integration Reference Point (IRP) management; Solution Set (SS) definitions (Release 10)*", 3GPP, June 2010. Available at: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1935.

[16] 3GPP TS 32.606: "*Configuration Management (CM); Basic CM Integration Reference Point (IRP); Solution Set (SS) definitions (Release 10)*", 3GPP, September 2010. Available at: http://www.tech-invite.com/3m32/tinv-3gpp-32-606.html.

[17] 3GPP (2011): *3GPP TS 32.662:* "*Configuration Management (CM); Kernel CM Information Service (IS) (Release 10)*", 3GPP, March 2011. Available at: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2134.

[18] A. Albanese (editor): *Deliverable D7.3:* "*Experimental Integration results of HW/SW modules of the overall SESAME framework*". H2020/5G-PPP SESAME project, September 2017.

2.3.1.2 Virtual Cell Status Reporting

As envisaged in ITU-T Recommendation X.731[19], the operability of a virtual cell can be "reflected" via the *Operational State* and *Availability Status* attributes of the managed object that represents it.
The EMS business logic would be required to link the following aspects in order to provide this status:

- The operability of the PNF hosting the virtual cell. In the case of small cell implementing the TR-069[20] data model, the *Device.Services.FAPService.{i}.FAPControl.LTE.RFTxStatus* parameter indicates whether or not the PNF is transmitting, a value of FALSE indicating that the transmitter is off. If the PNF is not transmitting, then the virtual cell cannot be in service and its *Operational State* should be DISABLED and its *Availability Status* should include the value DEPENDENCY.

- The operability of the SC-VNF serving the virtual cell and, *in particular*, whether or not it has successfully established an S1 connection to the configured MME. If the *Operational State* of the SC-VNF is DISABLED then, *similarly*, the EMS should set the *Operational State* of the virtual cell to DISABLED and its *Availability Status* should include the value DEPENDENCY.

- Currently, if an SLA breach is detected, the EMS SLA Monitoring service may, *as a configurable option*, raise on alarm on the appropriate Monitored SLA object. In addition to an entry in fault management view, the EMS also indicates the presence of such alarms by colour coding the SLA managed object in configuration management view. Whilst this makes it obvious that an SLA has been breached, it is not as obvious as to which virtual cells are affected. This is particularly true of the case where the geographic scope of the SLA makes use of the polygon option. A possible EMS enhancement would be to set the *Availability Status* of affected virtual cells to include the value DEGRADED when an SLA breach is detected.

2.3.1.3 SLA Modification

An obvious omission in the current PoC EMS is that the UE (User Equipment) and throughput caps specified in a Provisioned SLA are only propagated to the associated SC-VNF and SC PNF at virtual cell provisioning time. Additional EMS business logic is required to permit the SCNO to modify the parameters of a Provisioned SLA object and have the changes propagated to each virtual cell that was provisioned with the SLA.

Similarly, the EMS provides the VSCNO with the ability to select one from the set of SLAs created by the SCNO at virtual cell provisioning time[21].

Another useful feature would be the ability to apply a different Provisioned SLA to a virtual cell.

---

[19] International Telecommunication Union – Telecommunications Standardization Sector (ITU-T) (1992, January): Recommendation X.731: Information Technology (IT) – Open Systems Interconnection (OSI) – Systems management: State management function. ITU-T, Geneva.

[20] The TR-069 (Technical Report 069) is a technical specification of the Broadband Forum (BF) that defines an application layer protocol for remote management of customer premises equipment (CPE) connected to an Internet Protocol (IP) network. The CPE WAN Management Protocol (CWMP) defines support functions for auto-configuration, software or firmware image management, software module management, status and performance managements, and diagnostics. For further details see, *inter-alia*: https://en.wikipedia.org/wiki/TR-069.
Also, see: https://www.broadband-forum.org/standards-and-software/major-projects/tr-069-and-its-evolution.

[21] These could for example offer a choice of "Bronze", "Silver" or "Gold" service levels.

2.4 PNF Assessment

The PNF provided by the ip.access E40[22] required little change in order to work in the SESAME PoC demo as the majority of additional functionality over standard MOCN is implemented by the VNFs running in the CESC.

There were two specific areas of change to the PNF:

- Additions to the TR-069[23] data model and supporting code to permit the Automatic Neighbour Relations (ANR) feature to record and present the full list of PLMNs supported by MOCN capable neighbour cells.
- Additional PM counters that record activity on a per-PLMN basis. Currently, these counters are mainly in two areas: RRC setup and GTP usage.

The PNF "meets" the requirements of the PoC, as stated in the SESAME Deliverable D7,2[24]. In particular, we notice the following:

- It is able to be configured with multiple PLMN IDs which are broadcasted in System Information Block SIB1[25].
- UEs with SIMs having one of the broadcasted PLMNs can attach to the PNF, while UEs that do not have a supported PLMN are rejected.
- UEs with different, supported, Home PLMNs are able to simultaneously be attached to the PNF.
- The *Reserved for Operator Use* information element can be selectively set, by configuration management, for each PLMN in the list broadcast in SIB1 and, *when set*, UEs with a Home PLMN that is *reserved* do not attach to the cell.
- The PNF is able to establish an S1 connection to the SC-Common VNF and it correctly declares the list of PLMN IDs that it supports in the S1 SETUP request.
- A UE's *Selected PLMN* is correctly forwarded in the TAI[26] field of an INITIAL UE MESSAGE, allowing the SC-Common VNF to route the UE's signalling to the correct SC VNF[27].
- Data flows can be established, via the SC-Common VNF and serving SC VNF to the EPC and that bi-directional user plane traffic is exchanged.
- Performance Management reports contain the additional per-PLMN counters detailed in Deliverable D7.4[28].

---

22 For further details about the E-40 access point provided by ip.access Ltd., see: http://www.ipaccess.com/en/lte.

23 See: Broadband Forum (2013, November): TR-069: "*CPE WAN Management Protocol, Issue: 1, Amendment 5*". Available at: https://www.broadband-forum.org/technical/download/TR-069_Amendment-5.pdf .

24 A. Whitehead (editor): *Deliverable D7.2:* "*Integrated CESC Prototype Validation*". H2020/5G-PPP SESAME project, July 2017.

25 For further details about SIB1, also see, inter-alia: http://howltestuffworks.blogspot.gr/2011/11/system-information-block-type-1.html.

26 Tracking Area identity (TAI) is the identity used to identify tracking areas. The Tracking Area Identity is constructed from the MCC (Mobile Country Code), MNC (Mobile Network Code) and TAC (Tracking Area Code).

27 Note that user plane traffic bypasses the SC-Common VNF and is exchanged directly between SC VNF and SC PNF.

28 A. Whitehead (editor): *Deliverable D7.4:* "*Integrated Pilot and Evaluation Report*". H2020 SESAME project, December 2017.

- When a radio environment scan is performed, the PNF populates the full list of supported PLMN for each MOCN capable neighbour.

Possible PNF-*related* areas for future research and enhancement include those presented below.

2.4.1 Additional Performance Counters

In addition to the thirteen per-PLMN counters implemented by the current PoC (see Deliverable D7.4[29]), areas where additional counters would prove useful are:

- **Handover (HO) Success Rate**. The users of different PLMNs may be in different locations (particularly enterprise users in office buildings) any may well be handed in from and out to different neighbour cells. Reporting on handover success rate on a per-PLMN basis would prove very useful.

- **RAB Establishment Success Rate**. Again, if the users of different PLMNs are clustered in different locations, they may experience a different quality of service and reporting separately per-PLMN would provide useful diagnostic information.

- **Paging Activity**. Paging activity is an area where a VSCNO with a badly configured EPC could potentially comprise the performance of the PNF for other VSCNOs. A per-PLMN/per-VSNO paging activity report would allow such problems to be identified.

2.4.2 On air PLMN Status Indication

When a virtual cell is administratively locked or the PNF hosting does not have an S1 connection that supports a particular PLMN, then the PLMN ID broadcast by the PNF in SIB1 needs to be either "marked" as unavailable, or omitted from the broadcast list.

In the current SESAME PoC, this is achieved by setting the associated value of *Cell Reserved for Operator Use* to TRUE.

Whilst preventing access by the majority of UEs, strictly speaking, this is a miss-use of feature as UEs of access classes 11 and 15[30] are still able to access the PLMN[31].
A useful extension would be for the PNF to be able to broadcast additional information such as a *PLMN Unavailable* status.

[29] Ibid.

[30] 3GPP (2011): *3GPP TS 22.011:* "*Technical Specification Group Services and System Aspects; Service accessibility (Release 10)*", March 2011. Available at: *https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=566*.

[31] This approach was chosen as the "alternative" mechanism of removing PLMNs from the list having different associated problems that included the identity of the PNF changing (as its Cell ID is within the scope of the first PLMN ID broadcast) and RRC setup attempts being directed towards the wrong EPC (as UEs simply include an index into the list as their *Selected PLMN* and not the PLMN ID itself).

2.5 Light DC Assessment

The Light DC is a computing platform, consisting of micro-servers, distributed across the network edge. It helps bringing data processing closer to the end-user, reducing latency and increasing transfer rates. Moreover, the Light DC energy and space efficiency is achieved by using low power, small form factor micro-server platforms.

The Light DC is the component where all VNF, SC-VNFs and resulting SFC are executed. It provides heterogeneous platform, consisting of ARMv8[32] and x86[33] nodes. Those nodes have been studied and described in deliverable D4.1[34].

Some of them are equipped with different hardware accelerators enabling offloading heavy computational tasks (e.g., video transcoding) from the CPU. This hardware is fully supported by the software baseline providing virtualization, virtualized hardware accelerators, accelerated virtual networking as well as integration with the SESAME VIM of choice – OpenStack[35].

The activities and integration results related to the PoC of the Light DC are described in D4.4[36].
Further tests to be addressed may include:

- Integration with the SDN Controller for establishing the networking paths between VNFs to create the services, *and*;
- integration with the CESCM, as the upper management layer, able to communicate with the VIM to orchestrate the Light DC virtual infrastructure *as a whole*.

2.5.1 ARM

One of the "key innovations" presented by SESAME is the support for ARM architecture[37] in the virtualized infrastructure.

Such innovation is interesting for different reasons:

- Optimized power management;
- better core scalability on a "single die";
- openness of the hardware ecosystem (ARM CPU cores are produced in silicon by different vendors worldwide and are not relying on one single company), *and*;
- European nature of the technology.

For the above, and for other reasons explained in detail in the SESAME Deliverables D4.1[38], D4.2[39], D4.3[40] and D4.4[41], the SESAME architecture envisages an ARM-*based* Light DC whose low level software

[32] Further related information can be found, *for example*, at: http://www.arm.com/products/processors/armv8-architecture.php.
[33] For further details see, *among others*: https://en.wikipedia.org/wiki/X86.
[34] A. Albanese (editor): *Deliverable D4.1:* "*Light DC architecture design*". H2020/5G-PPP SESAME project, June 2016.
[35] For more details see: https://www.openstack.org/.
[36] A. Albanese (editor): *Deliverable D4.4:* "*Light DC Prototype*". H2020/5G-PPP SESAME project, June 2017.
[37] For more details also see, *inter-alia*: https://en.wikipedia.org/wiki/ARM_architecture.
[38] A. Albanese (editor): *Deliverable D4.1:* "*Light DC architecture design*". H2020 SESAME project, June 2016.
[39] P. Bliznakov, M. Paolino, S. Pinneterre and D. Raho (editors.): *Deliverable D4.2:* "*Virtualisation extensions for Acceleration of Light DC capabilities*". H2020/5G-PPP SESAME project, December 2016.
[40] I. Giannoulakis (editor): *Deliverable D4.3*: "*Techniques for efficient VNF Deployment with relevant VIM extensions, Evaluation framework*". H2020/5G-PPP SESAME project, June 2017.
[41] A. Albanese (editor): *Deliverable D4.4:* "*Light DC Prototype*". H2020/5G-PPP SESAME project, June 2017.

infrastructure (Linux distribution, virtual switch, OpenStack support, etc.) have been developed and tested within the framework of WP4 in SESAME.

After the development and benchmark activities performed on this platform during SESAME, we can mention that "*today the support for the ARM architecture in the NFV ecosystem is not at the same level of the competition*", i.e., Intel[42]. This is also true for the overall performance of the two solutions.

The most high-level distinction between the chips is based on power and speed. ARM chips are great for low-power environments but are typically slower, while the x86 chips work quickly but are not as power-conscious.
That basic generalization is changing, *though*, as Intel attempts to produce low-powered versions of its x86 chips, and ARM chips, like the latest ARMv8 processor produced by Qualcomm[43], begin to overtake laptop chips.

We therefore think that SESAME that has been a pioneer in this area and that the expertise, the solutions and the technologies built will be beneficial in the short-term future.

2.5.2 x86

The Intel server is the GOMA[44] FlexPAC[45] Industrial portable workstation based on Intel Xeon E5-2630v3[46] CPU. It is used in the SESAME PoC to host a HW-*accelerated* virtual Transcoding Unit, accelerated through an NVIDIA QUADRO M4000[47] GPU.
This server has been chosen for practical reasons and not for low power considerations. Anyway, for what we said previously about SW portability between x86 platforms, this choice does not affect the SESAME PoC.

FlexPAC Industrial portable workstation configured with:

- Aluminium Chassis including 17.3" LCD Display (Resolution 1920*1080).
- US Keyboard with integrated touchpad.
- CPU Intel Xeon E5-2630v3 2.4GHz, 8 Core, 16 Threads.
- 64 GB DDR4 RAM (max. 256 GB).
- Intel® C612 chipset[48].
-
- 2x PCI-e[49] 3.0 x16, (one reserved for GPU).
- 1x PCI-e 3.0 x8.
- 1x PCI-e 3.0 x4.
- 1x PCI-e 2.0 x4 (When CPU 2 is installed).

[42] https://www.intel.com/content/www/us/en/homepage.html.
[43] See: https://www.qualcomm.com/products/qualcomm-centriq-2400-processor.
[44] For more details see: https://www.gomaelettronica.it/en/server-and-workstation-portables-high-density-storage-server-intel-i7-i5-i3-series.
[45] For more details see: http://www.acmeportable.com/products/flexpac.
[46] For more details see: https://ark.intel.com/products/83356/Intel-Xeon-Processor-E5-2630-v3-20M-Cache-2_40-GHz.
[47] For more detailed information, also see: http://www.pny.com/nvidia-quadro-m4000.
[48] For more details see: https://ark.intel.com/products/81759/Intel-C612-Chipset.
[49] For more relevant information also see, *for example*: https://en.wikipedia.org/wiki/PCI_Express.

- 10x SATA3[50].
- 2x RJ45 GbE LAN ports.
- 4x USB3.0, 2x USB2.0, RAID[51].
- 1 x HDD 4 TerabByte SATA 3.5" in Drive bay.
- 3x 3.5" Removable Drive bay.
- DVD R/W.

It should be noted that in the SESAME distributed architecture the real important requirements are those of the overall CESC cluster. In fact, the SCNO should have the flexibility to use different micro-servers, each of them not necessarily having the possibility to show all the HW requirements needed by the Light DC. For example, let us suppose there are some VNFs that need a server equipped with a GPU PCIe card. These types of cards are quite expensive and pose severe physical and power constraints to the hosting server.
Anyway, if multiple VNFs running in different micro-servers need this HW resource, the computing power of a single GPU can be distributed among all the requesting VNFs by associating each VNF to a GPU slice. The same could be said for other HW accelerators (taking advantage, for example, of SRIOV[52] feature) or for storage. This distributed model, leveraging on the use of different types of complementary micro-servers, lowers the overall costs and improves performance, flexibility and global efficiency.

Following this particular vision, also the physical co-location of the micro-server with the Small Cell can be seen as "*not strictly necessary*".
In fact, the Ethernet link connecting the PNF to the micro-server can be up to 100m long and, *if needed*, it is not excluded the possibility to use some centralised computing resource (e.g. directly connected to a cluster switch) to be shared into the cluster.

The introduction of HW Accelerators in the SESAME architecture has been made, being aware that the research of a single processing platform able to exhibit maximum efficiency with all possible computational workloads is inevitably doomed to fail.

There will always be applications that can achieve the best performance on CPUs, while others will run optimally on GPUs, or on hardware accelerators embedded in a SoC, or on different programmable devices, such as DSPs or FPGAs. But, this is exactly the motivation why the proposition of a multi-accelerated architecture can "fit" a Mobile Edge environment where, *for a given performance*, minimizing space, infrastructure costs and energy consumption is a "must". In this context, the choice of relying on an architecture based on ARMv8, instead of more powerful x86 processor as Intel Xeon has a clear meaning.

Starting from the consideration that SESAME architecture does not exclude the possibility to deploy in its Light DC one or more x86 nodes, having distributed heavier workloads to more efficient GPUs, DSPs, FPGAs[53] (as well as HW accelerators inside SoCs), so much offloads CPU that most of its resources turn out to be overabundant. For this reason, in SESAME micro-server a x86 processor like Xeon[54], which is usually employed in servers, can conveniently be replaced by a less power-hungry ARMv8 processor.

[50] More related informative details can be found, *for example*, at: https://en.wikipedia.org/wiki/Serial_ATA.
[51] For more details see, *inter-alia*: https://en.wikipedia.org/wiki/RAID.
[52] For more details see, *inter-alia*: https://en.wikipedia.org/wiki/Single-root_input/output_virtualization.
[53] Also see the interesting discussion in: https://www.electronicsweekly.com/news/products/fpga-news/dsp-versus-fpga-2012-05/.
[54] For more details, also see *among others*: https://en.wikipedia.org/wiki/Xeon.

2.6 Monitoring Assessment

In the virtualized environment where an End-to-End (E2E) service is provided, the service can be "composed" by different infrastructure that will have to take into account the overall assessment of the whole infrastructure.
Nowadays, it is absolutely required to have a continuous monitoring system running in the network, systems and applications. This is, the only way to make an efficient maintenance and prevent disruptions in a proactive way; however, this requires a lot of time and resources to detect more hidden problems over a large period of time.

Total visibility of the infrastructure is a main requisite to prevent issues and it is very useful when making informed decisions to enhance the future system performance of the environment.

The main advantages to monitor systems are:

- Detect problems, before there is an adverse effect;
- Detect problems that do not allow applications within the expected performance;
- Collect behavioral data, when something stops working;
- Establish comparison in measurement between different environments or situations;
- Check the "core" of the problem when something fails;
- Prevent to reduce downtimes;
- Taking maximum advantage over hardware's capabilities;
- Access to system status in real time;
- Create inventories.

All this features can be achieved if the system is able to generate the following actions:

- Obtain performance information periodically;
- Store monitoring data for future use when there is a problem;
- Make general system status more visible for system's administrator;
- Collect more data over a system when a failure occurs;
- Track changes made to the system.

In the SESAME architecture we can consider three different environments to monitor; this will allow having an overall overview of the service provided.

In the following parts we present and discuss the environments and specific metrics:

- **Cloud Environment:** The monitoring process allows analysing information from the dynamic infrastructure. The cloud infrastructure is based on real and virtualized environment(s), such as servers, containers and virtual machines. All this sort of information is combined, in order to provide

full insights into customer experience, application performance management, and infrastructure monitoring. The metrics that are retrieved from the cloud infrastructure as described as follows:

- CPU (system, user, nice, iowait, steal, idle, irq, softirq, guest)[55].
- Memory Load.
- Disk Space Used in percent.
- Disk Utilization per Device.
- Disk IOS per device (read, write).
- Disk Throughput per Device (read, write).
- Context Switches.
- Network Traffic (In, Out).
- Netstat (Established).
- UDP stats (InDatagrams, InErrors, OutDatagrams, NoPorts).

- **Radio environments:** The radio parameters monitored are related to the Small Cell performance. The information is retrieved from the NOS system that is the one in charged to "gather" all the SCs performance. In the radio environment there are two time-*related* elements, that is: the granularity period is the interval in which the measurements are taken, and; the reporting period, is the interval in which one or several reports are generated; this last one is the selected for the current implementation.The parameters that are generated from the SCs are the following:
  - Sc_availability: Availability of the PNF measured as a percentage.
  - UplinkOctets: The total number of uplink octets transmitted by the PNF.
  - DownlinkOctets: The total number of downlink octets received by the PNF.
  - PERC_CallDropRate: Call drop rate of the PNF measured as a percentage.

- **Network environment:** The monitoring of the network parameters are more related to the NS compliance with the SLA. The control of the data plane flow in SDN features provides fine-granular overview and control of the network applications. The metrics produced by the SDN controller in the flow management are the following:
  - Flow table statistics per node *(control, compute, neutron and netfloc).*
  - Aggregate flow statistics per host.
  - Port statistics per host (e.g., netfloc, compute, control and neutron).
  - Service Function Chain-*related* flow statistics per host (SFC flows priority=20).

The assessment of the monitoring infrastructure can be grouped by service, as the new virtualized infrastructure can be assembled by location, data center or service function.

[55] Also, see: *https://www.opsdash.com/blog/cpu-usage-linux.html*.

This information will help to identify errors and, at the same time, it enables the detection and diagnosis of availability and performance problems across the entire stack.
More detailed information of the SLA evaluation based on this metrics can be found in the Deliverable D5.2[56] and Deliverable D7.4[57].

[56] E. Jimeno (editor): *Deliverable D5.2:* "*VIM and CESC Implementation*". H2020/5G-PPP SESAME project, September 2017.

[57] A. Whitehead (editor): *Deliverable D7.4:* "*Integrated Pilot and Evaluation Report*". H2020 SESAME project, December 2017.

2.7 NFVO Assessment

To correctly assess the level of Virtual Network Function (VNF) orchestration, it is important to come up with a grading mechanism. However, considering the fact that network virtualization/softwarization is a continuous effort pushed by motivations such as agility and cost reduction, drawing sharp "Technology Readiness Level (TRL)" lines would not be very easy.

In this section, we followed the Oracle proposal[58]: it is a multi-level orchestration and automation capacity for grading VNFs[59]. Such a normalized definition across the industry creates a common language/understanding, which will "boost" the technology adaptation/advancement.
For example, it helps VNF providers to understand the level of automation/orchestration required by customers: Telcos / Cloud Storage Providers (CSPs). Practically, it gives a "means" to Telcos/CSPs to assess the level of VNF provider's automation and orchestration readiness.

Open source[60] community could align their efforts with an orchestration level. Oracle suggested the following division:

- **Level 1**: Basic virtualization support: it represents the most basic virtualization capability, i.e. being able to have VNF(s) over the major hypervisor solutions, e.g. KVM[61], VMWare[62], etc. Here there is no VNF chaining is envisioned. The idea is only to be able to have standalone VNFs.

- **Level 2**: Basic automation deployment, i.e. on-boarding VNF(s) via templates (e.g. Heat[63], Tosca[64], etc.). It implicates for partial automated deployment (manual procedures may be needed).

- **Level 3**: Advanced automated deployment, i.e. on-boarding with metadata/scripts for bootstrapping VNF automatically ready for service configuration.

- **Level 4**: Basic service orchestration, i.e. VNFs fully operational and configured for service (chain of VNFs). It implicates for scale in/out support as well as for programmatically exposed KPIs.

- **Level 5**: Full service orchestration and lifecycle management support, i.e. VNF test certification. Automated upgrade. It implicates for automatic scale as well as for healing/maintenance support.

Having the presented VNF orchestration rating in mind, SESAME during its lifetime targeted few proof of concept (PoC) showcases to verify its contribution on the NFVO technology readiness. During its first year (Y1), SESAME showed two separate cases:

1. VNFs instantiation over the ARM platform: the choice of non-x86 architectures permits increased energy efficiency and better trade-off between cost and performance for the type of processing jobs expected (mainly light workloads). In this showcase (presented on the SESAME Y1 Review Meeting), SESAME verified *Level 1* of orchestration over the ARM platforms.

2. Automated deployment of the virtual watermarking service: on this showcase, automated lifecycle management of a VNF service, i.e. vWatermarking, has been proven. To do so, first the SESAME team worked on the forming the SESAME VNF templates[65]. The SESAME template was used to on-boarding vWatermarking (*Level 2*). A preliminary version of the demo has been presented at EuCNC

[58] See, *for example*: https://docs.oracle.com/cd/E18727_01/doc.121/e13442/T292653T292656.htm.
[59] S. Esfandiari, "What is your VNF orchestration level?" SDN NFV World Congress, The Hague, October 2017. Available at: *https://www.layer123.com/download&doc=Oracle-1017-Esfandiari-What_is_your_VNF*.
[60] For further details also see: *https://opensource.org/*.
[61] KVM (Kernel-*based* Virtual Machine) is a full virtualization solution for Linux on x86 hardware containing virtualization extensions. For further details, also see: *https://www.linux-kvm.org/page/Main_Page*.
[62] For further details see, *for example*: *https://www.vmware.com/*.
[63] For more details see: *https://wiki.openstack.org/wiki/Heat*.
[64] For more details also see, *inter-alia*: *https://docs.openstack.org/tosca-parser/latest/*.
[65] P. Paglierani (ed.): *Deliverable D5.1* "*Description of CESC abstraction model*". H2020 SESAME project, June 2016.

2016[66]. Next, for the SESAME Y1 Review Meeting, an updated version of the showcase verifying VNF on-boarding using the metadata/scripts for bootstrapping automatically has been prepared and showed (*Level 3*).

Built on top of the achievements of the past, on the Final Review Meeting SESAME targeted more advance PoC, i.e. to ensure the Quality of Service (QoS) / agreed SLA per tenant bases. It is an important aspect of the service lifecycle management which represents *Level 4* of NFVO technology readiness.
To deliver the job correctly, the orchestrator needs to simultaneously take into account both the radio status (i.e., volume of traffic, geographical distribution of traffic, etc.) and the cloud capabilities (i.e., available IT resources, VM-to-VM communication requirements, etc.) for all the actions related to the service lifecycle management. It means forming a QoS insurance feedback loop.
The actual implementation details of the SESAME solution have been reported in D6.4[67] but from a functional perspective, this demo verifies VNFs fully operational and configured for service, scale in/out support and KPI programmability (*Level 4*).

Since SESAME is a special project, in a sense that it focuses on a joint cloud-radio environment, it is expected that in future, other-EU funded projects, for example phase-II 5G-PPP projects, and/or open source communities shall adopt the SESAME solution and "push" the work forward toward higher NFVO TRLs.

[66] More detailed information about the wider scope of the EuCNC-2016 Conference can be found at: *https://www.eucnc.eu/2016/www.eucnc.eu/index78f5.html?q=node/93*.

[67] P.S. Khodashenas (editor): *Deliverable D6.4:* "*Orchestrator Prototype*". H2020/5G-PPP SESAME project, September 2017.

2.8 VNFs Assessment

Quality of Service and Quality of Experience are "key" characteristics of VSCNOs environments. If we consider performance management across the SESAME architecture, we have different building blocks to consider: the NFV Infrastructure (NFVI), the Management and Orchestration (MANO) stack, the different Virtual Network Functions (VNFs) and Network Services (NSs).
A number of metrics need to be defined when designing each one of these components; then some methods need to be implemented to collect these metrics appropriately, and "suitable" interfaces need to be available so that to "carry" the results across the architecture to different "authorized and subscribed consumers" of the metrics.
Then, performance measurements can be done either as part of pre-validation, with simulated steady traffic, or with peak of traffic. But performance measurements can also be done upon a live environment on an ongoing basis or on-demand, to check the behaviour of the network.

Two types of measurements are typically performed: Quality of Service (QoS), ensuring that the network behaves according to expectations, or; Quality of Experience (QoE), ensuring that the user perception of the network and of the service quality is according to expectations. These different concepts have been described in some initial specifications in ETSI NFV[68], then have further been refined in other specifications that will be detailed below. In parallel, a number of tools have been designed by the open source community, in particular in the context of the OPNFV[69] collaborative project. Other standard organizations have also specified benchmarking, for instance IETF.

Strong collaboration has occurred between those different entities and contributors to ensure consistency and complementary work. As NFV is touching on many different areas such as Telco Cloud, fixed and mobile network functions, customer premises environment, management platform and processes, service deployment and operation, many different types of metrics have to be defined and collected.
It became obvious that a number of metrics could be leveraged from existing standards in some of these areas, but also that some overlaps or inconsistencies existed when mapping those metrics to the ETSI NFV architecture.

As a result, an initiative was triggered across the industry in order to align metrics for NFV across key stakeholders. In parallel, a few telecom operators advanced in NFV deployment, such as Verizon[70], have issued requirements in terms of metrics they want suppliers to provide. Finally, as technology evolves, with new hardware, networking and virtualization capabilities, metrics and measurement methods and tools get also to evolve. New usage also drives new performance requirements that "drive" new technologies, metrics and measurement capabilities. In conclusion, NFV metrics and performance management is a long journey and this deliverable is just giving an introduction and update on some of the current highlights.

2.8.1 VNF benchmarking

In some ways, benchmarking a VNF should be the same as benchmarking a traditional device. For example, a virtualized router should be benchmarked similarly to a traditional router, by measuring its

[68] For more details, see the context discussed in: *http://www.etsi.org/technologies-clusters/technologies/nfv*.
[69] Open Platform for NFV (OPNFV) facilitates the development and evolution of NFV components across various open source ecosystems. Through system level integration, deployment and testing, OPNFV creates a reference NFV platform to accelerate the transformation of enterprise and service provider networks. For further details, also see: *https://www.opnfv.org/*.
[70] See, *for example*, the context provided in: *https://www.openstack.org/news/view/215/verizon-launches-industry-leading-large-openstack-nfv-deployment*.

forwarding performance using RFC 2544[71], as well as by measuring the scale and performance of key routing protocols such as BGP[72], IS-IS[73] and OSPF[74].

Traditional benchmarking should also be applied to virtual BNGs, virtual CPEs, and virtual firewalls/IPSs[75]. For a (virtual) load balancer, which is the example for this case study, benchmarking typically means assessing the maximum rate of HTTP requests the load balancer can sustain. Such benchmarking has been well-defined for several years, and is now embodied in products with scenario-*based* methodologies for rapid benchmarking of both traditional and virtualized devices.

2.8.2 Challenges Unique to VNF Benchmarking

Beyond traditional benchmarking, VNF benchmarking is different from traditional benchmarking in three distinct ways, that is: complexity, abstraction, and concurrency.

Complexity: Virtual environments are more complex since they include more components, such as hypervisors, virtual switches, etc. These additional components introduce potential areas of weakness and must be considered while benchmarking VNFs.

Abstraction Virtualization means that the data plane (DP) and control plane (CP) have been abstracted from the executing hardware. Abstraction introduces three challenges when benchmarking VNFs:

- The impact of the quantity and type of resources which have been allocated to the VNF. The variety of virtual resources continues to expand, for example, with virtual FPGAs and virtual GPUs now being available.
- The impact of certain technologies, such as DPDK[76], SR-IOV and CPU pinning, which are meant to address performance concerns associated with virtualization.
- The time to spin up, or instantiate, a VNF (in a traditional network element, the device simply exists and does not need to be created).

Concurrency: In real-world deployments, VNFs will often co-exist and possibly interact with other VNFs. Other VNFs will have their own performance constraints, and furthermore may also be competing for the same resources.

Historically, performance measurement of network devices could be described as “black box testing”, meaning that the testing was conducted at the external interfaces of the devices, without much thought into the inner workings of the device itself, other than some basic statistics and configuration issues. In the virtual world, the boundaries are much “blurrier”: VNFs are not standalone devices, but rather components in a “complex” ecosystem where resources can be shared across multiple consumers. Consequently, performance testing in virtual environments should best be thought of as “white box testing”, where measurements are made for each component in the ecosystem.

The challenge of abstraction really means performing benchmarking with different quantities and types of resources allocated to the VNF. For example, does doubling the amount of virtual memory double the scale of a routing protocol in a virtual router?

---

[71] S. Bradnerand, and J. McQuaid (1999, March): *Request for Comments (RFC) 2544: “Benchmarking Methodology for Network Interconnect Devices”*. IETF. Available at: https://www.ietf.org/rfc/rfc2544.txt.

[72] See, *for example*: https://en.wikipedia.org/wiki/Border_Gateway_Protocol.

[73] For further details see, *for example*: https://en.wikipedia.org/wiki/IS-IS.

[74] See, *for example*: https://en.wikipedia.org/wiki/Open_Shortest_Path_First.

[75] Also, see: https://www.paloaltonetworks.com/cyberpedia/what-is-an-intrusion-prevention-system-ips.

[76] For more details also see, *inter-alia*: https://en.wikipedia.org/wiki/Data_Plane_Development_Kit.

Abstraction challenges include the benefits and pitfalls of technologies such as DPDK, which can accelerate forwarding performance, but may also increase packet delay, especially for long packets.

We can retrieve useful and informative metrics from multiple places in the virtual infrastructure. We distinguish the following:

- From the hypervisor: CPU Utilization.
- From the virtual switch: System stats, virtual NIC stats, physical NIC stats.
- From the VNF or Operating System instance: CPU Utilization, Memory Utilization, Disk I/O Rate, Memory access Latency, Cache read/write latency, Instructions processed per second, Throughput for memory access.

2.8.3 vDPI

Packet classification is an essential part of most -if not of all- network functions and also important in the SESAME ecosystem. It is the process of associating packets with identifiers by analysing the layers of the protocol stack up to, but not including, the application layer (see Figure 2).

Actions taken by a VNF are based upon these identifiers. Since different applications can use the same protocol parameters from the perspective of the transport and network layer (i.e., same IP addresses, protocol IDs and ports), it is impossible to "reliably distinguish" between them with this type of packet classification.

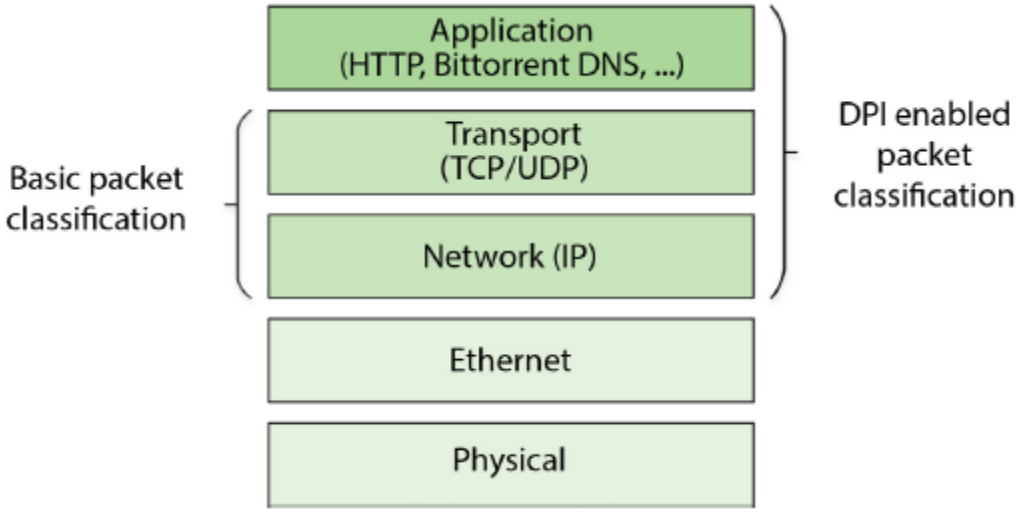

Figure 2: Layers of the Protocol Stack

Characterization is typically done by applying stateless load on a system under test (SUT) running the VNF workload. The constructed traffic contains packets based on a fixed set of template packets with randomized bit-patterns or a range of values written at pre-determined offsets. The transmission rate is fixed, possibly below line-rate (i.e., at 85% of line-rate). This is done by repeatedly inserting periods of silence on the wire. A characterization report then details the rates at which the VNF is able to perform its functions for the provided traffic. Other characterization details, like latency, could also be reported.

DPI (Deep Packet Inspection) inspects packets in context of flows. In many cases, it needs to inspect multiple packets before the classification is available. Clearly, loading the SUT that runs the DPI workload with the packets that do not carry any state (referred to as "packet blasting") repeatedly is not sufficient to characterize performance. Instead, the SUT needs to be loaded with traffic consisting of flows (referred

to as "flow blasting"). The important parameters for a given traffic profile are the maximum setup rate (reached mainly at the initial ramp up phase), the total number of concurrent connections and total bandwidth. The traffic profile itself also has an influence on the resource requirements for DPI.

SESAME focused on the DPI engine as an SDK solution meant for integration due to its broad applicability. For this, a prototype VNF was developed in which the SDK was used. The Prototype analyses packets (DPI path) while, *at the same time*, forwarding it through the network (data path). The VNF does not take any action on the classification result, but still performs all the necessary steps for the classification. One way the classification result could be used is in a VNF that implements traffic shaping. Here, the classification result could be used to determine which of the queues within the traffic shaping algorithm will be used to buffer the given packet. A use-case (UC) for this is a VNF that prioritizes *Skype* over HTTP traffic.

The results in this report show DPI performance in function of the number of cores assigned to DPI tasks. The bars refer to the highest link utilization for which the DPI cores were handling at least 99.999% of all packets. Figure 3 illustrates the DPI performance for a VNF with 1, 2, 3, 4 and 7 DPI cores (2, 4, 6, 8 and 14 hyper-threads, *respectively*). The fact that the maximum link utilization is below 100% has to do with the constraints of the traffic parameters.

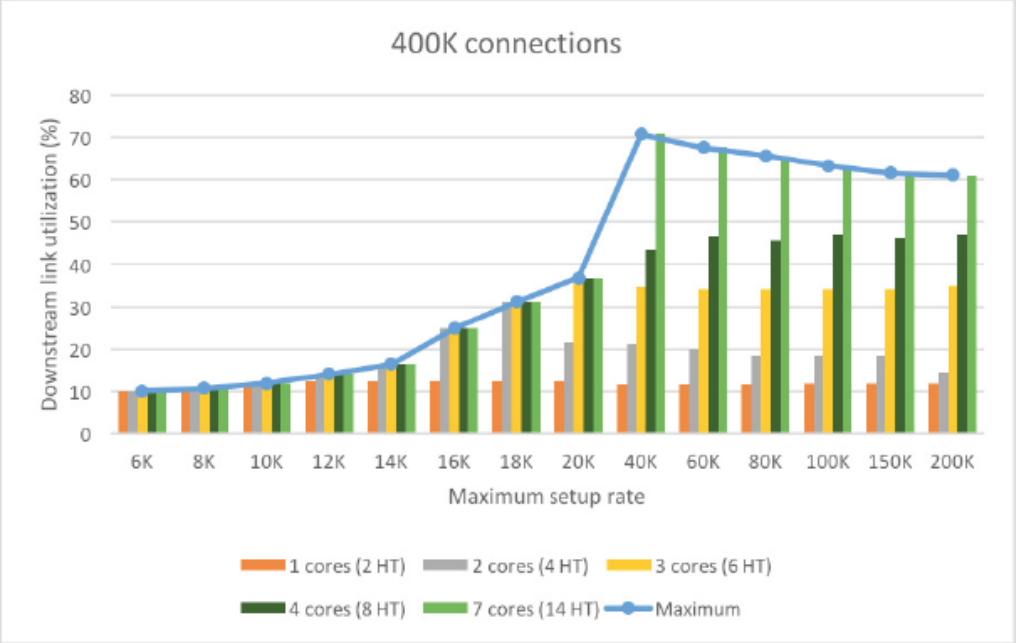

Figure 3: DPI classification performance for a VNF configured with 1, 2, 4 and 7 DPI cores. (The number of concurrent connections in steady state set to 400K)

Due to the interdependences of all the parameters, the approach is first to numerically find the solution space boundaries for the maximum setup rate, maximum number of concurrent connections and the per connection bitrate. The first two parameters can be chosen arbitrarily. The last parameter is specified indirectly. A transformation function takes as "input" the distribution of connection bitrates and a scaling factor. The function scales all connections bitrates by the scaling factor, unless the connection bitrate is limited by the link speed. This means that the bit-rate ratios are maintained, whenever possible. As a consequence, the link utilization is increased. The reported link utilizations in this report are measured link utilizations and are the result of applying the transformation described above.

Figure 4 shows how the downstream bandwidth evolves over time for a maximum setup rate of 150K connections per second and the number of concurrent connections set to 800K. The first phase, P1, is the phase during which the number of concurrent connections for the test has not yet been reached. This phase will take at least "number of concurrent connections" divided by the "maximum setup rate" seconds. It will take longer in most cases due to connections being terminated during the phase itself. The

second phase, P2, is a 10 second period (determined experimentally) during which the bandwidth peaks due to the following two reasons. First, the higher the setup rate, the more synchronized the connections are. This means that a higher setup rate will create a higher bandwidth peak. Second, on average, there is an increase in per-connection bitrate after the setup has completed since the setup is an exchange of a few small packets without any real data, i.e., the three-way TCP handshake.

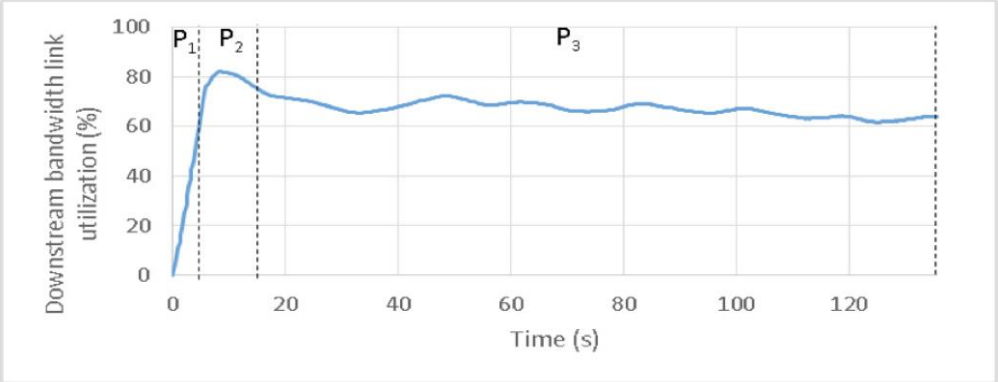

Figure 4: Bandwidth in function of time

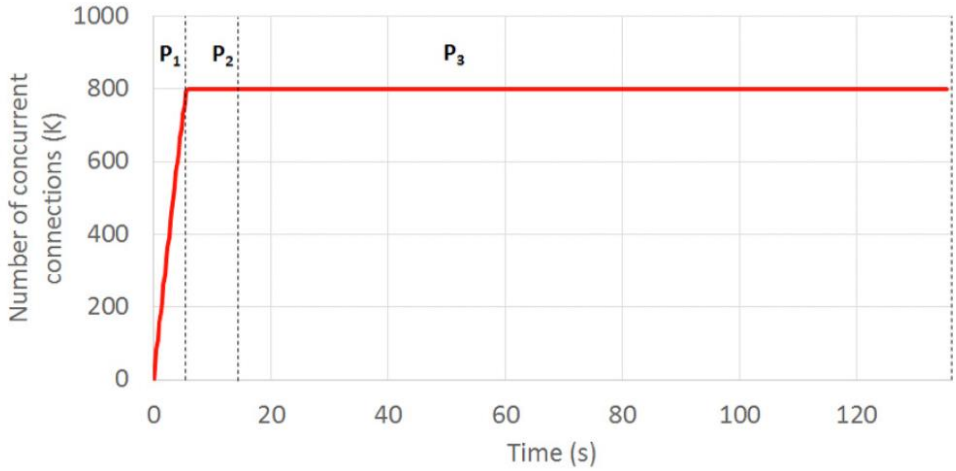

Figure 5: Number of concurrent connections in function of time during the test

The actual data transfer follows the setup (see the HTTP example from Figure 3). After P2, the bandwidth is in a steady state. This phase is the measurement phase for which the downstream bandwidth is reported in the results. It lasts for 120 seconds. The peak in P2 is the reason why the link utilization during steady state is lower than during the peak. The number of concurrent connections is shown in Figure 5. Note that the data shown by Figure 4 and Figure 5 is measured data.

In summary, since vDPI operates at the 4th layer and beyond in the protocol stack, the SUT must be loaded with state-full packet streams, based on a traffic profile extracted from real networks. The three parameters being changed for a traffic profile are the maximum setup rate, maximum concurrent connections and the link utilization.

2.8.4 VwM

Watermarking techniques can be very useful and convenient and this is why vWatermaking has been selected as one of the SESAME VNFs.
In fact, Watermarking techniques have evolved rapidly and succeeded to "embed" ownership data in a wide range of digital media such as Documents, sound tracks, images, video, file systems, etc.

In the SESAME architecture, the traffic exchanged between the EPC and the small cell, either data or control signals, is encapsulated by using GTP.
On the other hand, the OVS[77] requires pure IP packets, so the traffic must be de-capsulated from its GTP headers and then forwarded to the inner IP packets, which contain the actual video service data, to the upper modules.

The vWM can stamp or convert video streams from one video format to another. It can be implemented as a VNF providing optimized video transcoding and watermarking function, for the benefit of many other VNFs, to create enhanced services. Depending on the type of application that should be provided, the source video stream could originate from a file within a storage facility, as well as coming in form of packetized network stream from another VNF.
Moreover, the requested transcoding service could be mono-directional, as in video stream distribution-like applications, or bi-directional, like in videoconferencing.

The original videos have been selected from TREC Video Retrieval Evaluation[78] (TRECVID) collection]. The transcoding experiments have been performed with three scene of two minutes duration each one taken from the beginning of three original videos from three cameras.
All videos have the same initial format (MPEG-2) and the same encoding details. The bit rate of input videos is variable with mean value 6002 kbps, the frame rate equals to 25 fps and the resolution is 720 × 576. The vWM application was used for the video transcoding operation, while the MSU (Moscow State University) Video Quality Measurement Tool[79] has been applied for measuring video quality. Each video has been transcoded in three different bit rates: 128 kbps, 256 kbps and 512 kbps.

For saving on bit rate, videos have been scaled down to a smaller size by simply lowering the video screen resolution.
The three original videos have been captured in a 4:3 ratio. In order to maintain the original image aspect ratio in the resizing, the videos have been transcoded to 160×120, 320 × 240, and 640 × 280 screen resolutions.

Figure 6 presents the Peak Signal to Noise Ratio[80] (PSNR) metric for videos of different resolution, while the subsequent Figure 7 demonstrates the SSIM (Structural Similarity Index[81]) metric of the aforementioned video.
Regarding the video quality measurements, the three videos demonstrated comparable behaviour in both PSNR and SSIM metrics.

[77] See: *http://openvswitch.org/*.
[78] For further details, also see: *http://trecvid.nist.gov/*.
[79] For more details about the MSU Video Quality Measurement Tool also see, *inter-alia*: *http://www.compression.ru/video/quality_measure/video_measurement_tool.html*.
[80] For further details also see, *inter-alia*: *https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio*.
[81] For further details also see, *inter-alia*: *https://en.wikipedia.org/wiki/Structural_similarity*.

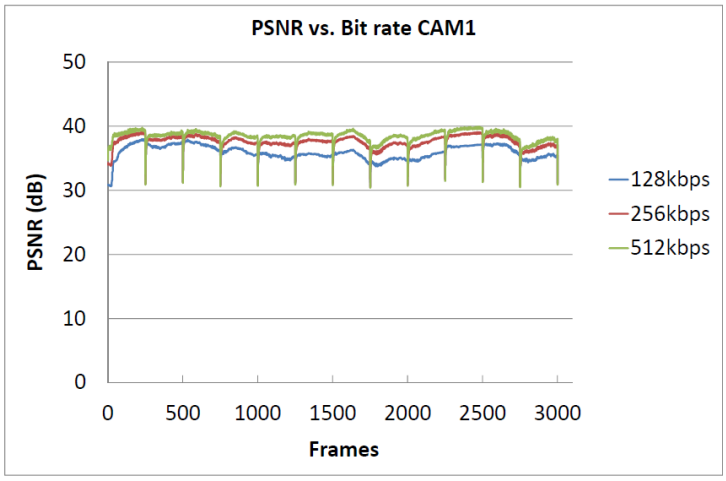

Figure 6: PSNR metric

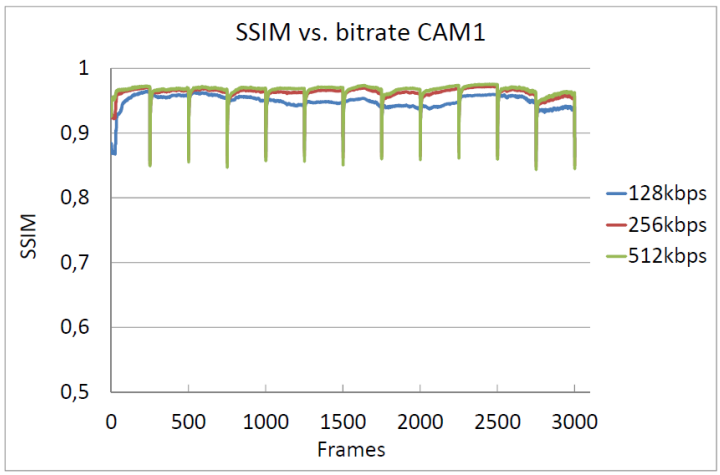

Figure 7: SSIM Metric

2.8.5 vTranscoding Unit - vTU

In the SESAME project, many tests were carried out to achieve a full performance characterization of the VTU, both for the SW-*only* version, and the GPU-*accelerated* one.

The SW-*only* version of VTU was tested on two different micro-servers, based on ARM and Intel x86 CPUs, *respectively*. The ARM-*based* micro-server is a NXP commercial evaluation board equipped with a LS2085A processor[82] (with 8x A57 cores @1.8 GHz) and 16GB DDR4 system memory. The Intel-*based* server is a commercial platform (GOMA FlexPAC Industrial portable workstation) equipped with an Intel Xeon E5-2630v3 2.4GHz, 8 Core CPU and 64 GB DDR4 RAM.

The GPU-*accelerated* version of VTU was tested on the same GOMA server, this time equipped with one NVIDIA Quadro M4000 GPU. During all the tests it was verified that the System Memory (RAM) was not completely used, to be sure that the results were not influenced by the different quantity of RAM installed in Intel-*based*, rather than ARM-*based* micro-servers.

In the following, only some meaningful results are presented and discussed, for the sake of brevity. In all described tests, the same H.264[83] Full HD video file (1080x1920 resolution) has been used as input.
In particular, Figure 8 shows the results obtained with the VTU featuring the H.264 transcoding (expressed in frames per second (fps)) without HW acceleration (SW-*only*) and with HW acceleration (using a GPU). The processing implies decoding from the input format to the one required as output.
The VTU provided four different video resolutions as output, in four different transcoding tests: VGA (480x640 pixel), HD480 (480x852 pixel), HD720 (720x1280 pixel), HD1080 (1080x1920 pixel). The vertical axis represents the output resolution, while the horizontal axis indicates the achieved output frame-rate in frames per second. The Encoder used in SW-*only* VTU for H.264 is X264[84]. The Encoder used by the GPU for H.264 and H.265 is NVIDIA NVENC[85].

Figure 8 collects the results of the tests achieved with H.264 encoders. Only one session was launched for each test. As one can easily see, the performance for the three HW platforms are very different. In particular, the ARM performance is very poor, to the point that it is not conceivable a utilization in a real-time scenario. The improvement achieved by using the GPU compared to a SW-*only* solution is remarkable in all cases. This confirms the need of GPU acceleration especially in modern and future scenarios where even higher video resolutions are going to be used. Another important aspect to emphasize is related to the occupation of compute resources during transcoding. Although in SW-*only* mode CPU resources were completely occupied, (all the CPU cores were running at 100%), using the GPU, both CPU and GPU resources were only partially used.
This fact led to a second set of tests in which the multi-session performance was analysed. In this new set of tests, the focus was on a single case, i.e., H.264 HD1080, launching 2, 4, 8, and 16 concurrent transcoding sessions.
The results are reported in Figure 9 and in Figure 10.

Considering the SW-*only* implementation (Figure 9), the performance of each single session decreases with the total number of executing sessions. Again, the ARM performance appears very low, being almost a fourth of the Intel one.

[82] More details can be found at: *https://www.nxp.com/products/no-longer-manufactured/qoriq-layerscape-2085a-and-2045a-multicore-communications-processors:LS2085A*.
[83] More details can be found at: *https://en.wikipedia.org/wiki/H.264/MPEG-4_AVC* .
[84] Also see: *https://www.videolan.org/developers/x264.html*.
[85] NVidia NVENC is a feature in its graphics cards that performs H.264 video encoding, offloading this compute-intensive task from the CPU. It was introduced with the Kepler-*based* GeForce 600 series in *March 2012*. For more informative details also see *inter-alia*: *https://en.wikipedia.org/wiki/Nvidia_NVENC*.

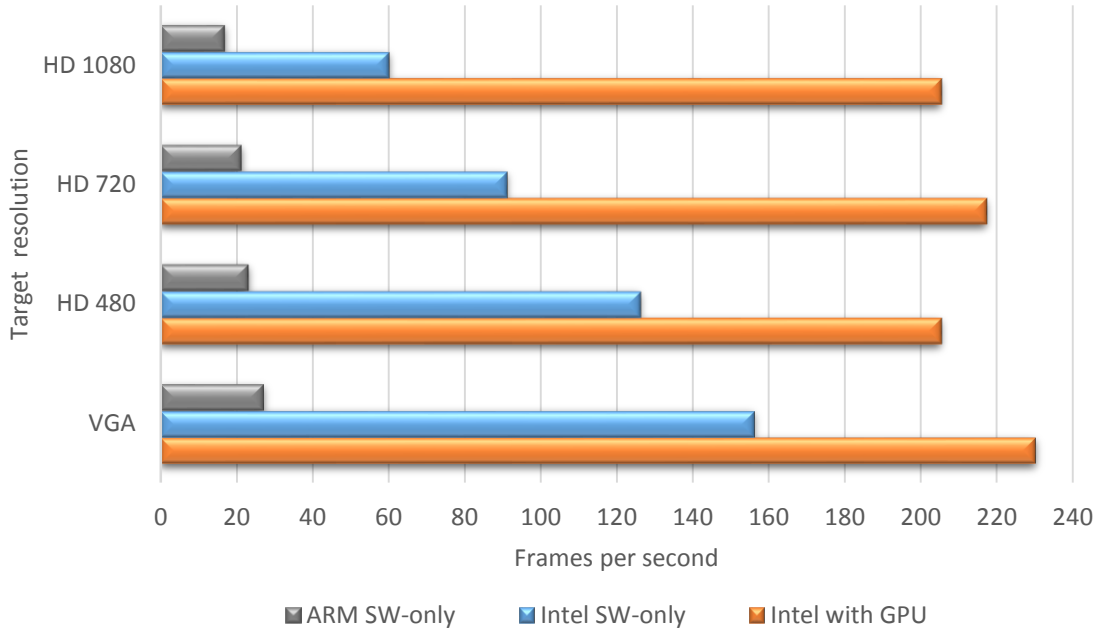

Figure 8: H.264 single session encoding performance (higher is better) measured on three different HW platforms, for different output resolutions

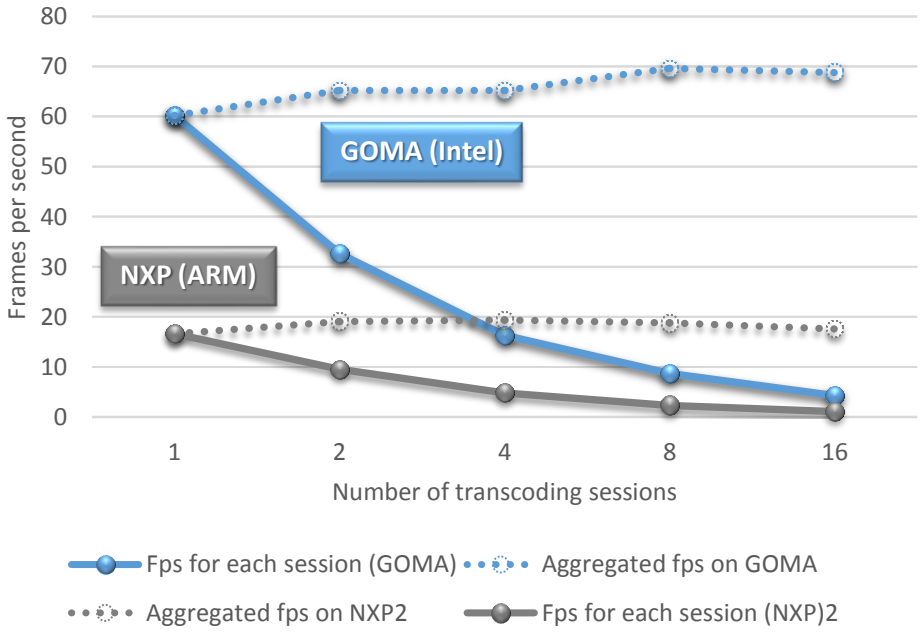

Figure 9: H.264 HD1080 encoding, SW-*only*, in multi-session transcoding tests (higher is better). Blue lines refer to Intel, grey to ARM. Performance refers to each single session (solid line) and to aggregated sessions (dotted lines)

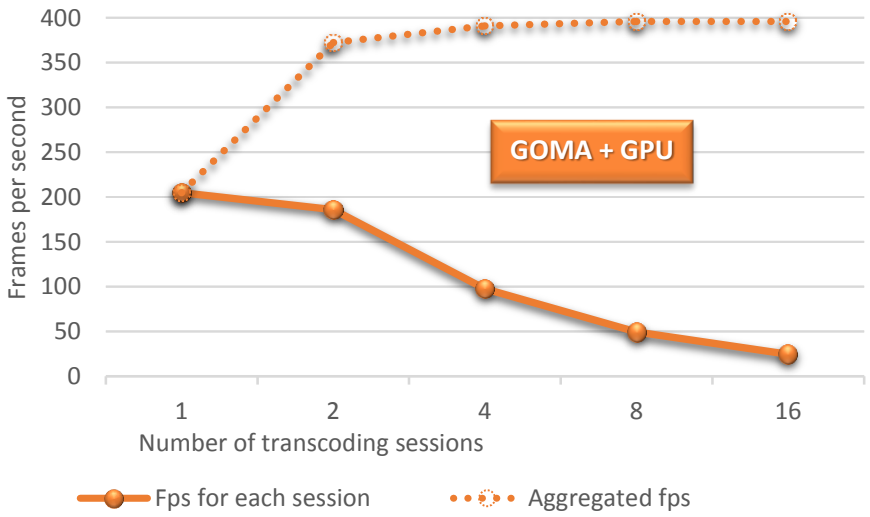

Figure 10: H.264 HD1080 encoding, using a GPU, in multi-session transcoding tests (higher is better). Performance refers to each single session (solid line) and to aggregated sessions (dotted lines)

Comparing the global performance of 1 session to that with 2-4-8-16 concurrent sessions (aggregated fps) the result is very similar as it can be seen from the dotted lines. This can be easily justified considering that the CPU occupation during the processing is always around 100% also running a single session. The same is not true when using the GPU (Figure 10).
In this case, the CPU is only partially used because the workload is mainly offloaded to the GPU whose resources are, *in turn*, not fully used (Figure 11). When using the GPU with 16 concurrent sessions we reach 24.75 fps for each session, for a total of 16x24.75=396 aggregated fps. The 396/69 ratio brings to a 5.7x gain in performance when using the GPU with respect to a SW-*only* solution.
Furthermore, during the GPU test with 16 transcoding sessions the CPU was running at 70%, giving it the possibility to run other tasks. This was not possible with SW-*only* solution, because in such a case the CPU was always 100% occupied.

It is interesting to analyse what are the power figures of the three HW platforms used to "run" the tests (multi-session performance). Figure 12 shows the power consumption, expressed in Watt.
The measures have been carried out in DC, testing the current flowing on the reference voltages of the motherboards (12V, 5V. 3.3V), to the end of excluding the contribution of the main AC-DC power supply and having a more comparable setup between platforms. We used a current clump with a resolution of 0.1A, resulting in a maximum uncertainty of 1.2W (on 12V voltage rail).

As expected, the NXP micro-server exhibits the best results, starting from 31W in idle state and going to 48W when running the VTU application. It is worth noting that 9.5W of the 31W are dedicated to the fans that always run at maximum speed; if the NXP micro-server could properly control the fan speed, its power consumption would improve significantly. Regarding the behavior of the Intel platform, with and without GPU, the results could appear unexpected because there are conditions in which the presence of the GPU does not increase the total power consumption, but rather decreases it. This can be explained considering Figure 121, which shows the percentage of CPU and GPU resources occupied during transcoding (dotted lines). Let us consider, *for example*, the situation with one session. Intel platform consumes 119W without GPU, and 95W with GPU. However, while without (w/o) GPU the CPU is running

at 100% (not reported in the figures), with GPU transcoding requires only 20% of CPU and GPU resources (as in Figure 12). This is the reason why the power consumption with GPU is in some case even better (lower) than the one w/o GPU.

Comparing the efficiency of the two solutions in term of performance/watt (Figure 13), we see another important advantage of using the GPU. In fact, in the case of 16 concurrent sessions (H.264-HD1080), the gain in efficiency using Intel + GPU is 5.4x compared to Intel (SW-*only*).
This could be, in some way, expected. Less obvious is the greater efficiency of Intel with respect to ARM (for this particular application).

Finally, Figure 14 shows the performance of the VTU when the transcoding is made starting from the same input file used so far, but converting it to a H.265 format[86]. This transcoding operation is significantly more complex from the computational point of view than the previous one (H.264), as one can see from the reduced performance respect to Figure 8.
We did not report the ARM performance, because it is too low to be considered. Intel performance is very low too, and the reduction in performance (and efficiency) with respect to the GPU reaches 10x.

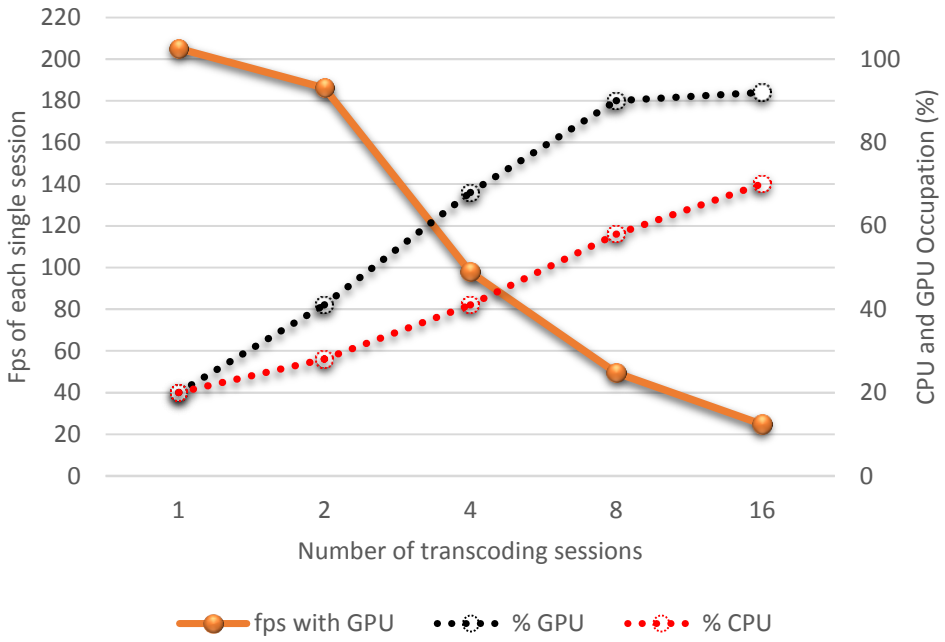

Figure 11: H.264 HD1080 encoding, with GPU in multi-session transcoding tests (performance related to each single session) with percentage of CPU and GPU resources utilization

[86] More details can be found, *for example*, at: https://en.wikipedia.org/wiki/High_Efficiency_Video_Coding.

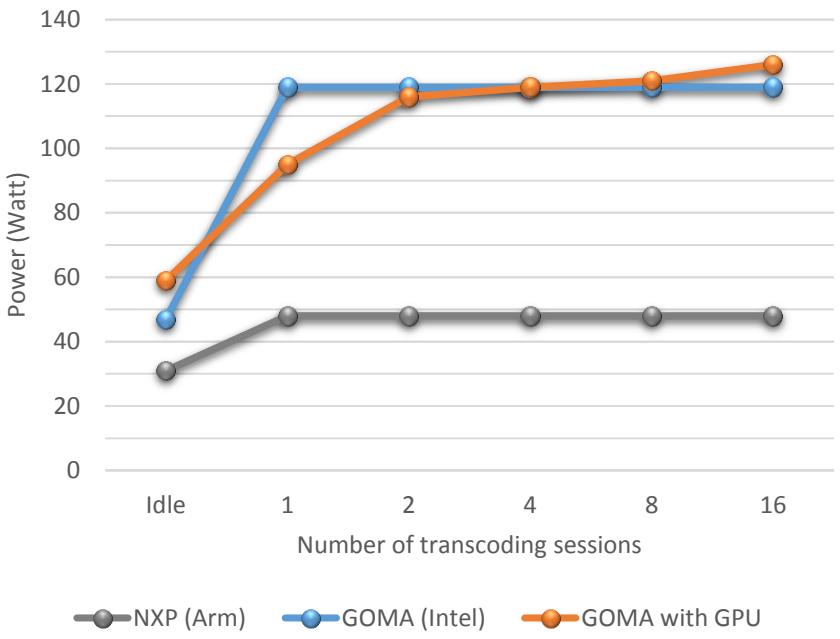

Figure 12: Power consumption (lower is better) of the three HW platforms running the H.264 HD1080 encoding multi-session transcoding tests

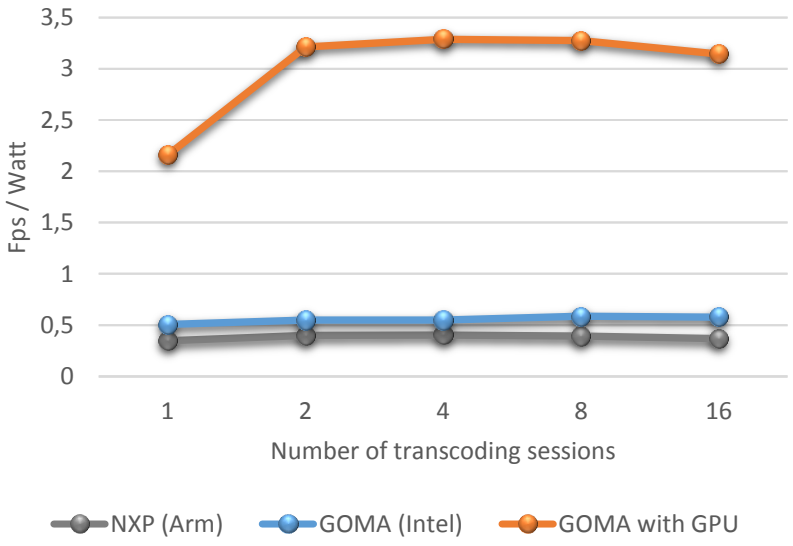

Figure 13: Efficiency of the three HW platforms (expressed in performance/power) for H.264 HD1080 encoding in multi-session transcoding tests (higher is better)

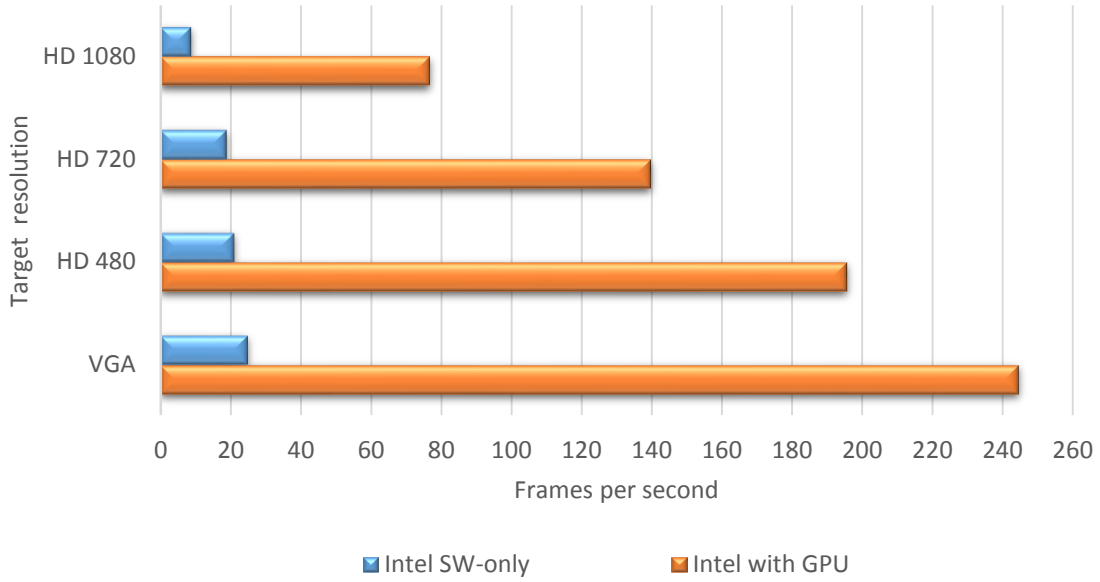

Figure 14: H.265 single session encoding performance (higher is better) measured on two different HW platforms, for different output resolution

2.8.6 Video Analytics (VA)

The two real-time video analytics-*based* VNFs (VA VNFs) enable two types of VA-*based* services, namely Augmented Reality (AR) service and real-time remote control of IoT devices (Smart IoT) services.
For more detailed information regarding the descriptions and implementations of the two VA VNFs and the descriptions of the PoCs for the two services, the reader is referred to SESAME Deliverables D4.3[87] and D7.3[88], *respectively*.
The information provided here focuses upon the assessment results of the two VA VNFs.

The AR VA VNF is integrated with and tested by using the SESAME testbed at NCSRD and demonstrated at EuCNC 2017[89].
The test results showed that, compared to hosting the AR VA VNF at a third party remote cloud server, the deployment of the AR service at the edge light DC can ensure that the end-to-end (E2E) service latency is within 100 ms, and the receiving device can receive the processed video data with unnoticeable delay compared to the real-life situations, whereas the remote cloud deployment on average introduces a two frame delay compared to the real-time situations. The frame rate of the streaming camera being used in the tests is 30 fps.

The test results of using remote cloud computing and edge computing for the AR service are shown in Figure 15 and Figure 16, *respectively*.

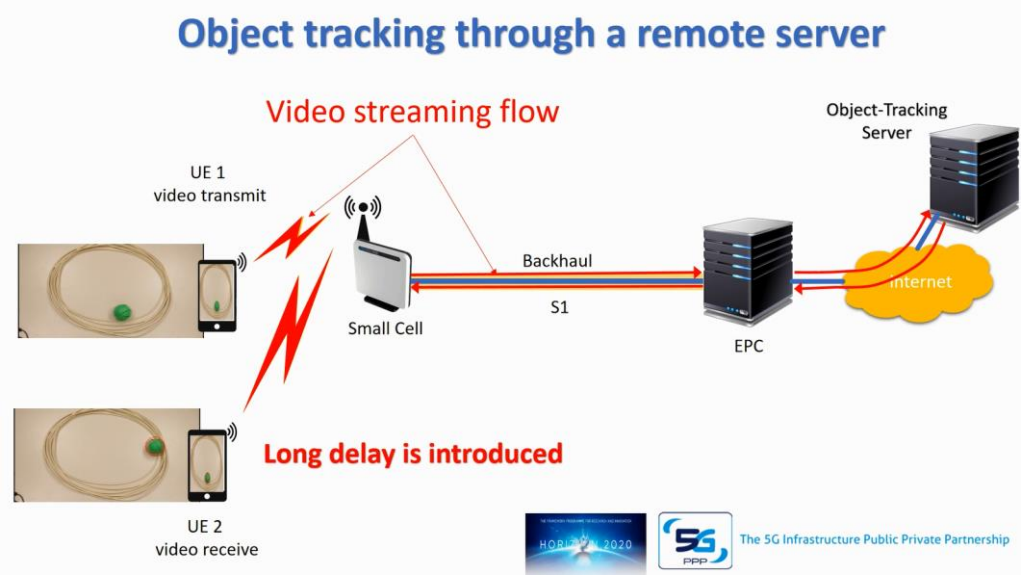

Figure 15: Architecture and performance of the object-tracking - based AR service, hosted at a remote server

[87] I. Giannoulakis (editor): *Deliverable D4.3:* "*Techniques for efficient VNF Deployment with relevant VIM extensions, Evaluation framework*". H2020/5G-PPP SESAME project, June 2017.

[88] A. Whitehead (editor): *Deliverable D7.3:* "*Experimental Integration results of HW/SW modules of the overall SESAME framework*". H2020/5G-PPP SESAME project, September 2017.

[89] More details about the wider scope of the EuCNC-2017 Conference can be found at: *https://www.eucnc.eu/2017/www.eucnc.eu/index78f5.html?q=node/93*.

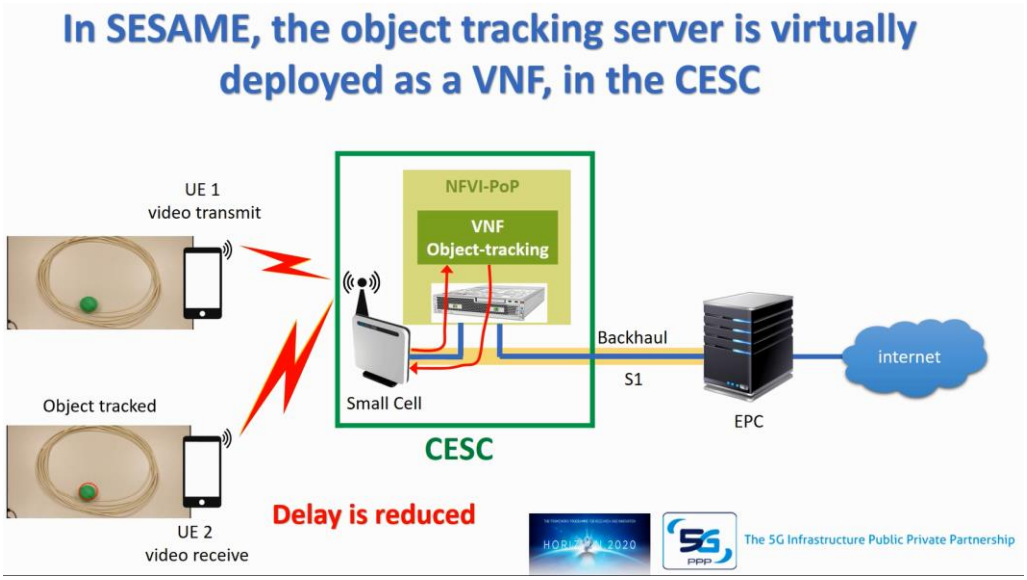

Figure 16: Architecture and performance of the object-tracking based AR service hosted at a mobile network edge server

The Smart IoT VA VNF is tested by using a WiFi-*based* testing system developed by FLE. The set-up of the testing system is illustrated in Figure 17.
The test results show that when a predefined object of interest moves across multiple cameras views, the pans and tilts of the available monitoring cameras are intelligently and remotely controlled by the Smart IoT VA VNF based on the real-time video streamed from a monitoring camera to the Smart IoT VA VNF.

The timeliness and validity of the pan and tilt adjustment decisions made by the Smart IoT VA VNF and the corresponding actions taken by the relevant IoT devices ensure the continuous seamless tracking of an object of interest achieving millisecond level response time to the real-time changing situations.

For the detailed algorithms for detecting the need for the video feed switching or camera reconfiguration, the reader is referred to SESAME Deliverable D4.3.

Basically the movement trajectory of an object of interest is tracked and predicted; if the estimated time for the tracked object to leave the current camera's coverage is slightly larger than the required time budget to complete the video feeding switching or camera reconfiguration process, then a decision is made to trigger the control decision making process, which will start to find the most appropriate camera or camera settings to enable the continuous tracking of the object of interest.
The performances of the movement tracking algorithm and that of the movement prediction algorithm are shown in Figure 18.

Figure 18 shows the results of the tracked and predicted trajectories of an object of interest (in this case a rolling ball) when it is moving within the view field of a monitoring camera.
The magenta dots represent the left and right boundaries of the monitoring camera's view field, and the blue dashed line denotes the boresight of the monitoring camera.

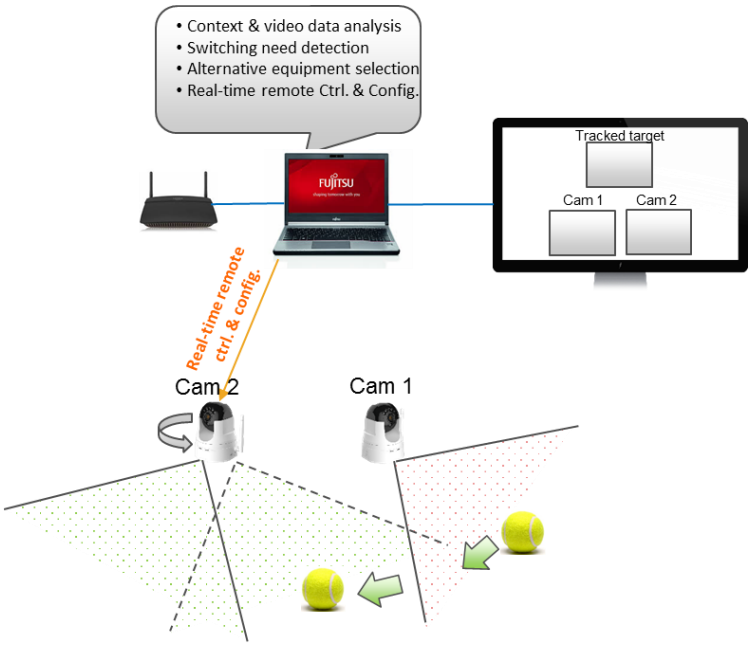

Figure 17: Illustration of the WiFi-*based* testing system for the Smart IoT VA VNF

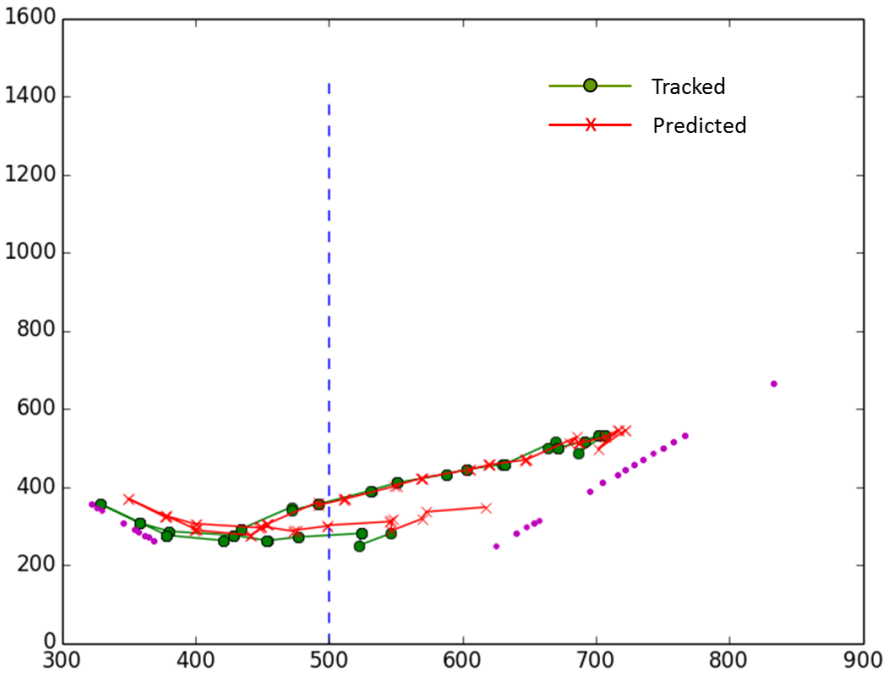

Figure 18: Tracked and predicted trajectories of a moving object

2.9 Service Chain Provisioning and Performance Measurements

The Service Function Chain (SFC) is a service supported by Netfloc. Throughout the development of Netfloc[90] and the chaining algorithm, several assessment measurements had been performed in order to provide the answers to the following questions:

(1) What is the advantage of the networking abstraction model used by Netfloc with respect to native OpenStack networking?

(2) What are the achievements in terms of protocol complexity and overheads?

(3) What is the performance of the chain implementation with different traffic patterns?

(4) What is the provisioning timeframe of the service chain over a simple SDN-*enabled* OpenStack environment?

In order to evaluate the functional operation of the service chain solution several experiments, there have been performed, being run in a real Cloud-NFV environment.
The experiments referred to the performance and scalability of different network service chains. These results have validated the SFC algorithm and its applicability in a real NFV environment and also answer the questions (1), (2) and (3). The experimental measurements focused on two main aspects:

1. *Characterization of the SFC deployment in terms of feasibility and validity.*

2. *Performance comparison results of the following QoS: throughput, packet loss, latency and jitter for use-case chain services.*

Further in-depth information of the results from the assessment results, can be found at *http://www.sciencedirect.com/science/article/pii/S092054891730017X*.
Overall, the summarized results are the following:

- The throughput rises linearly with packet size for all traffic types (ICMP[91], UDP and TCP).
- Service chain decreases the throughput, with more impact on TCP than UDP, both TCP and UDP values being bounded to the infrastructure limitations.
- UDP packet loss decreases with lower packet sizes.
- The chain affects a sort of traffic with low packet sizes.
- TCP reaches higher latency values for packets beyond 1000 bytes.
- Highest jitter values are measured in the case of chained TCP traffic.
- UDP jitter values are negligible.
- Similar like the latency, it is the traffic type that defines the quality within the chain, not the replicated chain scenario.

As a discussion it was observed that several impact sources are responsible for degraded traffic quality within the chains and this includes: network outage (physical and virtual), SDN controller limitations, and VNFs' service type constraints, occasional ARP[92] message exchange within the scenario, etc. Furthermore,

[90] For further details, also see *for example*: *https://github.com/icclab/netfloc*.
[91] For more details see, *among others*: *https://en.wikipedia.org/wiki/Internet_Control_Message_Protocol*.
[92] For more details see, *among others*: *https://en.wikipedia.org/wiki/Address_Resolution_Protocol*.

it was shown that scaling the chain service does not affect significantly the overall traffic quality when compared to single chain topology.
These preliminary scalability results appointed two main reasons as boundaries for the traffic overall quality: (1) environment setup (system limitation, physical and virtual switch interfaces, hard-wire thresholds), and; (2) scenario specifics (behavior of OpenStack, Netfloc, the VNFs) as well as the effect of each individual network function along the virtual traffic path (with the specific actions they perform to the packets).
As a conclusion it was proved a correct chain algorithm and decent traffic quality results. The option to improve an NVF-SDN use-case is by using more powerful equipment and advanced throughput technologies, such as DPDK.

To answer the question (3) and as it was discussed in the T-NOVA project Deliverable D7.2 (*http://cordis.europa.eu/project/rcn/189104_en.html*) the SFC creation can be invoked in two ways: directly via Heat in OpenStack and through the SESAME orchestrator TeNOR[93].
In the first case, the resource provisioning and service deployment time is dependent on the performance of Heat, except for the last step of chain creation via Netfloc, which consists in API call invoked from TeNOR to the Netfloc Chain APIs with a duration in a scale of seconds.
When TeNOR deploys the service via Netfloc, the process takes the resource mapping and instantiation of OpenStack networks, ports, subnets and VNFs. As a last step, the Netfloc API create chain is invoked from the service template used by TeNOR.

Overall the following are the measured times for the service-related events:

- Netfloc startup time: running ./bin/karaf: – 1min40sec–1min45sec.
- OpenStack resource creation: networks and VNF VMs all specified via HOT[94]: Heat stack create of demo_create.
- Chain creation: create chain resource from same heat template – 60sec.

The expected output after the listed events is couple of created VNFs created, Neutron[95] port IDs and Chain ID returned in heat stack information and service chain virtual path established in the network.

[93] Also see: *https://github.com/T-NOVA/TeNOR/wiki*.
[94] More details about the Heat Orchestration Template (HOT) can be found, *for example*, at: *https://docs.openstack.org/heat/pike/template_guide/hot_guide.html*.
[95] For more details see, *inter-alia*: *https://wiki.openstack.org/wiki/Neutron*.

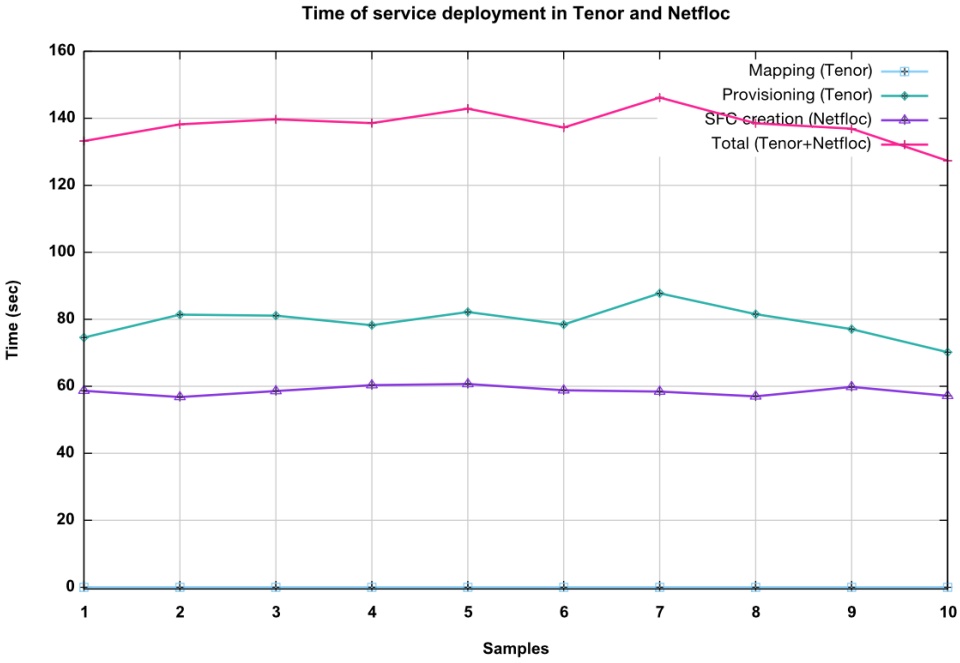

Figure 19: Resource mapping, provisioning and service chain creation time statistics

Figure 19, *as above*, depicts the average time taken during several samples of service deployment, and this has been tested in the NCSRD's testbed, used as well for the T-NOVA project (*http://www.t-nova.eu/*.).

2.9.1 Netfloc exporter for Prometheus and Grafana – metrics description

With the objective to provide advanced assessment of the SFC and further visibility of the network traffic controlled by Netfloc, the Netfloc-Prometheus exporter[96] has been implemented.

In this section we include the metrics that have been exposed in Grafana[97] and used in the assessment of the SFC service as represented in Prometheus[98]:

Flow table statistics per node
****

Netfloc active flows per node (control, compute, neutron and netfloc).
Ex: openflow_132129486422869 is the openflow identification of the node.

netfloc_active_flows{node_label="control_openflow_132129486422869"} 13.0
netfloc_active_flows{node_label="compute_openflow_255726480544624"} 13.0
netfloc_active_flows{node_label="neutron_openflow_255726480164143"} 17.0

Netfloc packets lookedup per node
netfloc_packets_looked_up{node_label="control_openflow_132129486422869"} 16.0
netfloc_packets_looked_up{node_label="compute_openflow_255726480544624"} 44881.0
netfloc_packets_looked_up{node_label="neutron_openflow_255726480164143"} 251111.0

Netfloc packets matched per node
netfloc_packets_matched{node_label="control_openflow_132129486422869"} 8.0
netfloc_packets_matched{node_label="compute_openflow_255726480544624"} 44873.0
netfloc_packets_matched{node_label="neutron_openflow_255726480164143"} 251103.0

Aggregate flow statistics per host
****
This shows how many packets, bytes and flows are there in the physical host.

Netfloc byte count per node
netfloc_byte_count{node_label="control_openflow_132129486422869"} 0.0
netfloc_byte_count{node_label="compute_openflow_255726480544624"} 0.0
netfloc_byte_count{node_label="neutron_openflow_255726480164143"} 3456.0

Netfloc flow count per node
netfloc_flow_count{node_label="control_openflow_132129486422869"} 2.0
netfloc_flow_count{node_label="compute_openflow_255726480544624"} 2.0
netfloc_flow_count{node_label="neutron_openflow_255726480164143"} 6.0

Netfloc packet count per node
netfloc_packet_count{node_label="control_openflow_132129486422869"} 0.0
netfloc_packet_count{node_label="compute_openflow_255726480544624"} 0.0
netfloc_packet_count{node_label="neutron_openflow_255726480164143"} 42.0

Port statistics per host (e.g., netfloc, compute, control and neutron)
****

[96] See: *https://github.com/icclab/netfloc-prometheus-exporter*.
[97] See: *https://grafana.com/*.
[98] Also, see: *https://prometheus.io/docs/visualization/grafana/*.

This section of metrics shows the packets (transmitted and received) and bytes (transmitted and received) for each port in each host.
For example, the example below shows that the control node has 2 ports (eth1_1, "br_int_LOCAL) and there have been 3604 and 8 packets transmitted via those ports are respectively.

Netfloc packets transmitted per node and per port
netfloc_packets_transmitted{node="control",port="eth1_1"} 3604.0
netfloc_packets_transmitted{node="control",port="br_int_LOCAL"} 8.0
netfloc_packets_transmitted{node="compute",port="eth1_80"} 3628.0
netfloc_packets_transmitted{node="compute",port="br_int_LOCAL"} 44873.0
netfloc_packets_transmitted{node="neutron",port="qg_46e7cfae_6e_135"} 0.0
netfloc_packets_transmitted{node="neutron",port="tape601145a_ec_2"} 24.0
netfloc_packets_transmitted{node="neutron",port="tap47277d52_53_3"} 19.0
netfloc_packets_transmitted{node="neutron",port="br_int_LOCAL"} 158.0
netfloc_packets_transmitted{node="neutron",port="tap0c65a8f4_cf_1"} 31.0
netfloc_packets_transmitted{node="neutron",port="tap0e190da8_ff_4"} 8.0
netfloc_packets_transmitted{node="neutron",port="eth1_101"} 3643.0
netfloc_packets_transmitted{node="neutron",port="qr_4d0877db_d6_5"} 0.0

Netfloc bytes transmitted per node and per port
netfloc_bytes_transmitted{node="control",port="eth1_1"} 680.0
netfloc_bytes_transmitted{node="control",port="br_int_LOCAL"} 648.0
netfloc_bytes_transmitted{node="compute",port="eth1_80"} 558.0
netfloc_bytes_transmitted{node="compute",port="br_int_LOCAL"} 648.0
netfloc_bytes_transmitted{node="neutron",port="qg_46e7cfae_6e_135"} 0.0
netfloc_bytes_transmitted{node="neutron",port="tape601145a_ec_2"} 648.0
netfloc_bytes_transmitted{node="neutron",port="tap47277d52_53_3"} 648.0
netfloc_bytes_transmitted{node="neutron",port="br_int_LOCAL"} 648.0
netfloc_bytes_transmitted{node="neutron",port="tap0c65a8f4_cf_1"} 648.0
netfloc_bytes_transmitted{node="neutron",port="tap0e190da8_ff_4"} 648.0
netfloc_bytes_transmitted{node="neutron",port="eth1_101"} 648.0
netfloc_bytes_transmitted{node="neutron",port="qr_4d0877db_d6_5"} 864.0

Netfloc packets received per node and per port
netfloc_packets_received{node="control",port="eth1_1"} 8.0
netfloc_packets_received{node="control",port="br_int_LOCAL"} 8.0
netfloc_packets_received{node="compute",port="eth1_80"} 7.0
netfloc_packets_received{node="compute",port="br_int_LOCAL"} 8.0
netfloc_packets_received{node="neutron",port="qg_46e7cfae_6e_135"} 0.0
netfloc_packets_received{node="neutron",port="tape601145a_ec_2"} 8.0
netfloc_packets_received{node="neutron",port="tap47277d52_53_3"} 8.0
netfloc_packets_received{node="neutron",port="br_int_LOCAL"} 8.0
netfloc_packets_received{node="neutron",port="tap0c65a8f4_cf_1"} 8.0
netfloc_packets_received{node="neutron",port="tap0e190da8_ff_4"} 8.0
netfloc_packets_received{node="neutron",port="eth1_101"} 8.0
netfloc_packets_received{node="neutron",port="qr_4d0877db_d6_5"} 10.0

Netfloc bytes received per node and per port
netfloc_bytes_received{node="control",port="eth1_1"} 680.0
netfloc_bytes_received{node="control",port="br_int_LOCAL"} 648.0
netfloc_bytes_received{node="compute",port="eth1_80"} 558.0
netfloc_bytes_received{node="compute",port="br_int_LOCAL"} 648.0
netfloc_bytes_received{node="neutron",port="qg_46e7cfae_6e_135"} 0.0
netfloc_bytes_received{node="neutron",port="tape601145a_ec_2"} 648.0
netfloc_bytes_received{node="neutron",port="tap47277d52_53_3"} 648.0
netfloc_bytes_received{node="neutron",port="br_int_LOCAL"} 648.0
netfloc_bytes_received{node="neutron",port="tap0c65a8f4_cf_1"} 648.0

netfloc_bytes_received{node="neutron",port="tap0e190da8_ff_4"} 648.0
netfloc_bytes_received{node="neutron",port="eth1_101"} 648.0
netfloc_bytes_received{node="neutron",port="qr_4d0877db_d6_5"} 864.0

Service Function Chain-related flow statistics per host (SFC flows priority=20)
****

This section contains the metrics for the SFC service. In particular, it shows the packets and flows associated to SFC which in the implementation of Netfloc is tagged with priority=20. So for example: *ServiceChainEndRewrite_2_1_00_00_e8_94_f6_08_53_70* is the flow rule with ID ServiceChainEndRewrite, and refers to the 2nd chain (SFC number 2). The rest of the label states the hop count (_1_) and the mac address (00_00_e8_94_f6_08_53_70 === 00:00:e8:94:f6:08:53:70) of the port where the flow rules is applied. What most matters in this case is the chain ID number as distinciation between 2 or more different service chains, (if we have 2 service chains it will be 1 and 2).

In SDN connotation, this means that on a port with mac address 00:00:e8:94:f6:08:53:70, the ServiceChainEndRewrite rule from the SFC service 2 was matched 78 times.

Netfloc flow packet count per node
netfloc_flow_packet_count{flow="ServiceChainEndRewrite_2_1_00_00_e8_94_f6_08_53_70",node="compute"} 0.0
netfloc_flow_packet_count{flow="ServiceChainRewrite_2_0_00_00_e8_94_f6_08_53_70",node="compute"} 0.0
netfloc_flow_packet_count{flow="ServiceChainEndRewrite_2_1_00_00_e8_94_f6_08_53_70",node="compute"} 0.0
netfloc_flow_packet_count{flow="ServiceChainRewrite_2_0_00_00_e8_94_f6_08_53_70",node="compute"} 0.0

Netfloc flow duration per node
netfloc_flow_duration{flow="ServiceChainEndRewrite_2_1_00_00_e8_94_f6_08_53_70",node="compute"} 78.0
netfloc_flow_duration{flow="ServiceChainRewrite_2_0_00_00_e8_94_f6_08_53_70",node="compute"} 78.0
netfloc_flow_duration{flow="ServiceChainEndRewrite_2_1_00_00_e8_94_f6_08_53_70",node="compute"} 78.0
netfloc_flow_duration{flow="ServiceChainRewrite_2_0_00_00_e8_94_f6_08_53_70",node="compute"} 78.0

Netfloc flow byte count per node
netfloc_flow_byte_count{flow="ServiceChainEndRewrite_2_1_00_00_e8_94_f6_08_53_70",node="compute"} 0.0
netfloc_flow_byte_count{flow="ServiceChainRewrite_2_0_00_00_e8_94_f6_08_53_70",node="compute"} 0.0
netfloc_flow_byte_count{flow="ServiceChainEndRewrite_2_1_00_00_e8_94_f6_08_53_70",node="compute"} 0.0
netfloc_flow_byte_count{flow="ServiceChainRewrite_2_0_00_00_e8_94_f6_08_53_70",node="compute"} 0.0

This list of metrics, despite being represented in the Grafana GUI, they are used in the Netfloc GUI, Netflogi[99]. The idea is to rely on SDN-*based* streaming telemetry for the purpose of advanced monitoring of NFV services in cloud and 5G environment.
This has been tested and is fully compatible with the SESAME CESCM monitoring portal in order to provide value-added SLA monitoring and alerting solution within SESAME.

[99] For more details, also see: https://blog.zhaw.ch/icclab/release-of-netflogi-a-graphical-interface-for-netfloc/.

2.10 Security Risk Assessment

The aim of this section is not to provide a security analysis focusing on vulnerabilities and potential attacks (such analysis is presented in the SESAME Deliverable 5.3) but rather to discuss high level security risks that are related to SESAME. Moreover, we provide some high level countermeasures to minimize the identified risks.

In the context of 5G architectures, to better understand potential risks, it is important to have a clear understanding of the relevant technological context. For example, virtualization is an innovative technology and method of deploying and implementing 5G Networks but from a risk perspective it is important to focus on the volatility of data.
On the other hand, many sophisticated attacks, malware and advanced persistent threats utilize vulnerabilities that are related to the Virtual Machine Bypass vulnerabilities and tries to jump through VMs or find open ports of more important parts of the infrastructure. Another important factor to consider is that most approaches are "vendor-*oriented*", meaning that a huge part of the maintenance of the system is bound to be depended on vendors and their patches for fixes of future vulnerabilities.

There are several risk analysis approaches. For our analysis we make use of the OWASP approach[100], where it is based on standard methodologies and is customized for application security. In that approach, risk is defined by the following formula:

*Risk = Likelihood * Impact*

According to the OWASP method which is based on standard methodologies and is customized for application security, we identify the following four (4) distinct steps as of defining a risk:

Step 1: Identifying a Risk. The first step focuses on the discovery and identification of security risks. This involves collection of information about threat agents involved, attacks that might be used, vulnerabilities involved and the impact of a successful exploit on the infrastructure. There may be multiple possible groups of attackers, or even multiple possible business impacts. In general, the aim is to "address" the worst-case scenario, as that will result in the highest overall risk.

Step 2: Factors for Estimating Likelihood. Once a risk has been identified, the next step is to identify the severity of it. In doing so, an estimation of the "likelihood" is necessary. In an abstract level, this is a rough measure of how likely this particular vulnerability is to be uncovered and exploited by an attacker. Generally, identifying whether the probability is low, medium or high and it is acceptable. There are many factors that can be utilized in order to identify the probability. A set of factors that this analysis is focused upon, is the threat agents involved. The objective is to calculate the probability of a successful attack from a group of possible attackers. It is worth to be mentioned that there may be multiple threat agents that can exploit a particular vulnerability. As a result, like in step 1, the aim is to "address" the worst-case scenario. For example, an insider may be a much more likely attacker than an anonymous outsider, but it depends on a number of factors.

Step 3: Factors for Estimating Impact. The impact of a successful attack should be categorized in two types of impacts. The first is the "technical impact" on the application -infrastructure in general- the data it uses, and the functions and features it implements. The other is the "business impact" on the business and company operating the application and uses the infrastructure. However, an analyst may not have access to all necessary information required to figure out the business consequences of a successful exploit. In this scenario, providing a detailed report about the technical risk will result to "make decisions" and suggestions and having an overview about the business risk.

Step 4: Determining Severity of the Risk. In step 4, the estimation of the probability and the estimation of the impact are "put together" to calculate an overall severity for a specific risk. This is done by computing

[100] OWASP-Risk Rating Methodology; See: https://www.owasp.org/index.php/OWASP_Risk_Rating_Methodology.

whether the likelihood is low, medium or high and then we do the same for the impact. Researchers should determine through the identification of threat factors and identify the most serious and important ones that visualize the result. In due time, there is a possibility for the researchers to find that their initial impressions were incorrect by considering aspects of the risk that were not obvious.

We consider the risk of stolen source code, data security breach for personal, financial, customer data medium high risks, because they are serious risks and have serious impacts but there is a potential probability to "skip" the countermeasures for financial issues of the involved legal entity, or "company". These risks have an important impact to the related company and the reputation of it, but they are risks that do not threat the organization in an immediate state of starting the business goal. Same goes of the risk equipment uncertainty.

Malware, internal network break in from outside, low average security knowledge attack are considered as high risks; for such, countermeasures and policies should be established and should be considered on the list of implementation and to the politics of the related company. Impact from these risks can be crucial for the continuity of the web application and cause financial loss for the related company.
Other risks that considered high priority and belong to the human factor risks are social engineering and productivity uncertainty.

Finally, the risk of funding and the risk human error identification even though have an important impact if not planned, are two risks well established in SESAME and they will continue to be as such in the near future.

2.11 Potential Countermeasures

Below there is a more detailed explanation of the countermeasures that potentially could be taken to mitigate the risks mentioned before. It is worth to mention that the countermeasures discussed here are more detailed and the numbers that are shown in the table below is according to these explanations (listed as 1-9), as in the Table 1 are presented in a more generic way. This section is made to "emphasize" on the implementation part and make suggestions and discussion.

1. VM isolation: The ability to perform isolation of the temporal behavior of multiple VMs among each other is called temporal or performance VM isolation, despite the fact that VMs can be run on the same physical host and sharing resources such as processors, memory and disks.

2. Privilege Escalation Restrictions: In order to mitigate privilege escalation techniques below is a list of potential security measures:

- Applications should be configured with least privileges in order to mitigate the capability for buffer overflow exploits that aims to elevate users to root and making it harder to execute known privileged techniques in the memory;
- Measures for prevention of data or code execution;
- Digital signed kernel mode code;
- Software Updates - Patches;
- Usage of antivirus and intrusion detection software with regular updates;
- Use of an operating system with Mandatory Access Controls[101] (MAC) such as SELinux[102];
- Encryption and usage of cryptographic protocols of software and assets of an infrastructure.

Table 1: Risks and potential Countermeasures

Risk	Probability	Impact	Countermeasures
Source code stolen by External attacker or insider	Medium High	Medium High-loss of competitive advantage, possible negative media exposure	9-Training of the employees-written and implemented information security policy
Virus Worm or Trojan infections	High	Medium-disruption, system restoration needed, unauthorized possible negative media exposure	2-Antivirus software/IDS/Firewall
Low average security knowledge	High	Medium-unauthorized access, media exposure, security bypass, data leakage.	9-Training of the employees / written and implemented information security policy
Denial of service (DoS) attack(s)	High	Medium productivity loss, negative media exposure, disruption, possible restoration system needed.	1,6-Software Updates/VM Isolation/

[101] For further clarification, also see, *for example*: https://en.wikipedia.org/wiki/Mandatory_access_control.
[102] For more details see, *inter-alia*: https://en.wikipedia.org/wiki/Security-Enhanced_Linux.

Data Security breach	Medium High	High unauthorized access, media exposure, security bypass, data leakage.	7-Outsourcing security assessments
Internal network break-in from outside	High	Medium High-disruption, systems restoration needed, possible negative media exposure.	2-Monitoring equipment/IDS/Firewall
Lack of security awareness	High	Medium -unauthorized access, media exposure, security bypass, data leakage.	9-Frequent threat and risk analysis
Social Engineering	High	High-disruption, systems restoration needed, possible negative media exposure.	9-Training of staff - written and implemented information security policy
Human error identification	Low	Low- Identify what errors can occur.	4-Outsourcing security assessments
Equipment Cost Uncertainty	Medium	Medium- Equipment costs are not well established but should be regulated by competitive market representatives.	9-Monitor mechanisms of the procedures by appropriate personnel
Productivity Uncertainty	High	High-The planned rate of progress needed to reach completion as planned is extremely aggressive or no benchmark experience is available to judge the reasonableness of the planned progress rate.	9-Frequent meetings and establishment of regular updates of the project's progress
Funding Risk	Low	Low- Project schedule targets may not be met because the projected funding needed to conduct the planned activities is not available when needed. In turn, schedule delays caused by underfunding can produce a need for increased funds.	9-Monitor mechanisms of the procedures by appropriate personnel

3. Access Control - Authorization: Authorization is a mechanism -or a procedure- where access to a specific service should be allowed, or rejected. Authorization is not similar to authentication. Authentication is providing and validating identity. Authorization includes the set of rules that decides what services and functions the user can entry, giving the appropriate access rights after authentication is successful.

4. Password policy - two factor authentication: Two-factor authentication[103] (also known as 2FA) is a mechanism or a procedure of validating user's identity by applying a combination of two different assets.

[103] Also, see *among others*: *http://searchsecurity.techtarget.com/definition/two-factor-authentication*.

5. Data Confidentiality - Encryption: Data Confidentiality means preserving from information disclosure performed from unauthorized third parties. Encryption is considered one of the most important measures for protecting data confidentiality. Some examples of security protocols and measures are SSL/TLS, enforcing file permissions and access control list to restrict access to sensitive information.

6. Software Update: Usually a Denial of Service attack abuses vulnerability in the software of a device and exploits this vulnerability in order to use this device to perform denial of service attacks to specific targets and services to great extent. Updating and patching this software in order to "fix" the specific vulnerability is considered as a mitigation technique for such kind of attacks, if redirection of the Denial of Service is not an option.

7. Http sessions: Potential countermeasure is session tracking using HTTP sessions API. A new session is created each time a user logs in and all the information required is stored in the session object. The HTTP session keeps track of the user status and whenever the user signs out of the system, the entire data in the session object is destroyed so that other users cannot access sensitive information.
Moreover, we are not storing and saving the entire credit card number in the session, only the last four digits are stored and saved. So, even if the user does not close the browser, others cannot access his personal details.
Preventing SQL Injection: The main way to protect a website again SQL injection attacks is to ensure that SQL statements and variables received and requested from the user are developed and constructed in such way that we can say that we have a reduce of such a risk. This can be done by removing any characters that can be used by some malicious users or targeted attacks from intruders to program and construct their programmed SQL statements to be queried on the database. Finally, an implementation with layers of abstraction between user input and the SQL statement being queried and protected on a database preventing XSS (Cross-Site Scripting[104]). The use of POST requests makes a site more secure from XSS attack than using GET[105] requests. So web developers should use POST[106] requests as much as possible to strengthen and protect their websites against XSS attacks.
It is known that XSS is the most common security vulnerability and most common type of an attack in a web application according to Google and Google's Security Team. Another method of protecting against XSS attacks is to not allow any HTML markup to be entered into forms on a website unless it is definitely and positively necessary. Any HTML markup can simply be taken out by the programmer by processing the incoming data.

8. Preventing Session Hijacking Secure Socket Layer (SSL)
Using SSL is suggested by many security researchers as a method of protecting an application from session hijacking. This is happening because any data being sent between the client and the server is encrypted so users are protected against man in the middle attacks. Disabling the use of session IDs as $_GET variables.
Session Timeout cookies (cookies used to hold the session ID): Cookies are destined to remain while users browsing with the web browser, until a session is terminated. It is advisable to override this and set a certain amount of time so as a session will last for this period of time. This method makes it difficult for an attacker to hijack sessions as by the time they have collected the data the session could have become invalid and irrational. However, a real time attack and intrusion can sometimes prove that this method of protection does not defend and guard from harm against this type of attack.
Regenerating Session IDs: It is the process when users signs in or signs out of the system, the session ID they use will change and transit. This has as a result to make the previous session ID inoperative and the new one unknown and unidentified. This method of protection prevents intruders from using session fixation.

9. Training of personnel - Raise security awareness
Security awareness training is the procedure and measures taken by a legal entity in order to educate employees about computer security. A good security awareness program should educate employees about information security policies and their implementations and procedures for working and cooperating with information technology (IT) regulations.

[104] For more details see: *https://en.wikipedia.org/wiki/Cross-site_scripting*.
[105] Also, see *inter-alia*: *https://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol#Request_methods*.
[106] Also, see *inter-alia*: *https://en.wikipedia.org/wiki/POST_(HTTP)*.

Employees should receive information about who to contact if they discover a security threat and be taught that data as a valuable corporate asset. Regular training is particularly necessary in organizations with high turnover rates and those that rely heavily on contract or temporary staff. Confirming how well the awareness program is working can be difficult and one way to measure and evaluate it is the number of incidents over time.

2.12 Self-X Functions Assessment

The assessment of the Self-X functions in SESAME has been carried out primarily in the context of the use case "*Optimized Radio Network Capacity Planning and Operation mechanisms of the Small Cell Network Operator*" defined in the Deliverable D2.1[107].
This use case intends to highlight the SESAME functionalities that allow achieving an advanced planning of the required radio capacity to be proactively provisioned by CESCs at any time and at any place and an optimized operation of the CESCs relying on the use of "Self-X" functionalities, also referred to as Self-Organizing Network (SON) functionalities, to configure the radio parameters.

The capability to share the CESCs among multiple tenants is an important differential characteristic of the SESAME scenarios. This sharing is sustained in network slicing principles, so that each VSCNO is provided with a Radio Access Network (RAN) slice, that is an instantiated logical RAN able to provide a specific capacity with potentially specific capabilities (ranging from just different network policy settings to the support of different protocols and features), realised together with other slices on the common physical infrastructure.

The way how the RAN slicing is implemented impacts on how the network should be planned and operated. Then, given the relevance of RAN slicing, the SESAME project has studied this problem in a multi-cell RAN, presenting four different alternatives that differ in the Radio Resource Management (RRM) functions used as a support for splitting the radio resources among the tenants.
These alternatives are: the "RAN slicing at spectrum planning" (i.e., assigning different carriers to each tenant in all the cells of a scenario); the "RAN slicing at InterCell Interference Coordination (ICIC)" level (i.e., assigning different resource blocks to each tenant in all the cells of a scenario); the "RAN slicing at packet scheduling (PS)" (i.e., the scheduling process of each cell decides how many resource blocks are assigned to each tenant), and; the "RAN at admission control (AC)" (i.e., the admission control process of each cell decides how many radio access bearers are admitted for each tenant).
The different RAN slicing approaches have been compared both qualitatively and quantitatively[108] analysing aspects such as the degree of radio-electrical isolation and traffic isolation, the granularity and flexibility for assigning resources to tenants and the capability to customize the RRM and associated Self-X functions, on a per-tenant basis.

Whatever the RAN slicing approach is considered, the SCNO faces the challenge to carry out the network planning stage in order to decide how many small cells have to be deployed in a certain area and how many frequency channels need to be assigned to each one.
The multi-tenancy nature in SESAME scenarios makes this problem more difficult to tackle, since the planning needs to consider the Service Level Agreement (SLA) with each tenant as it will impact on the amount of resources needed by each of them.

In this respect, SESAME has proposed a functional architecture for automating the planning process considering that the RAN slicing allows sharing a given carrier among multiple tenants.
The architecture includes various entities in order to: (i) Translate the SLA capacity terms into detailed planning specifications (i.e., demanded capacity per tenant needed in each small cell); (ii) check whether or not the deployed capacity fits the tenants' demand, and; (iii) identify the required changes in the network layout and channel allocation. The detection of capacity issues is periodically checked in order to automatically trigger specific planning actions, such as adding/removing a channel or deploying/relocating a small cell. The proposed approach makes use of specific capabilities, such as the measurements collection, the resources utilization monitoring and the determination of SLA compliance. The assessment of the proposed framework has shown that capacity overprovisioning can be minimized by reducing

[107] M. Belesioti and I. Chochliouros (editors): *Deliverable D2.1:* "*System Use Cases and Requirements*". H2020/5G-PPP SESAME project, December 2015.
[108] J. Pérez-Romero (editor): *Deliverable 3.2:* "*Self-X features and virtualised CESC multi-tenancy techniques evaluation*". H2020/5G-PPP SESAME project, March, 2017.

uncertainties about the spatial and temporal correlations between the new tenant's traffic and the actual traffic in the network[109],[110].

After planning and deploying the network, the RRM and Self-X functions will be in charge of the dynamic operation to ensure an optimized performance.
In this respect, SESAME has focused on controlling the share of radio resources among tenants at admission control level, through the development of a self-optimized multi-tenant AC algorithm that considers:
(i) A cell-level capacity check, which ensures that the cell has sufficient physical resource blocks to admit a new radio access bearer (RAB), *and*;
(ii) a per-tenant capacity share check, which ensures that the resources used by the radio access bearers of one tenant do not exceed the capacity share of this tenant.
A hybrid SON technique that accounts for unused capacity left by the tenants and that copes with heterogeneities in the spatial traffic distribution across different cells has been proposed and assessed[111],[112]. The self-optimized multi-tenant AC algorithm makes use of specific capabilities such as the measurements collection on a per-tenant basis or the reconfiguration of parameters.

The automation of the network planning and optimization processes by means of Self-X functionalities can be enhanced through the inclusion of knowledge discovery capabilities that smartly process input data from the environment and come up with knowledge that can be formalized in terms of models and/or structured metrics that represent the network behavior. This allows gaining in-depth and detailed knowledge about the whole ecosystem, understanding hidden patterns, data structures and relationships, and using them for a more efficient network management.
How and at what extent the Artificial Intelligence (AI) can be used to support the optimization decisions made by Self-X has been a topic of discussion in SESAME and a framework for knowledge discovery has been developed to sustain such discussion identifying different categories of knowledge models at cell level, user level and cell cluster level, together with different AI-*based* tools for knowledge discovery relying on classification, prediction and clustering processes. The applicability of each knowledge model to the different Self-X functions was identified as well.

The introduction of the knowledge discovery capabilities in the EMS has been demonstrated in the Deliverable D4.3[113], based on processing the Performance Management (PM) reports provided by the ip.access Ltd. Network Orchestration System (NOS) in the form of XML files generated according to a certain reporting interval.
Then, focusing on the cell-level knowledge models, the knowledge discovery has been particularised to support the energy saving Self-X function through a classification process of the time domain traffic pattern of the cells for identifying candidate cells to switch-off at certain time periods.
In turn, from the perspective of user-level knowledge models, in the SESAME Deliverable D3.2[114] a clustering methodology was presented for characterizing the QoS of individual users by exploiting the predictability of the user daily motifs, which allows the operator to better identify situations with "poor" user performance that could not be detected by classical network centric optimisation mechanisms relying on aggregate statistics for all the users of a cell.

[109] Ibid.
[110] A. Betzler (editor): *Deliverable 3.3:* "*Framework of a distributed network management system capable to host and run Self-X functionalities*". H2020/5G-PPP SESAME project, June 2017.
[111] Ibid.
[112] Ibid.
[113] I. Giannoulakis (editor): *Deliverable D4.3:* "*Techniques for efficient VNF Deployment with relevant VIM extensions, Evaluation framework*". H2020/5G-PPP SESAME project, June 2017.
[114] J. Pérez-Romero (editor): *Deliverable 3.2:* "*Self-X features and virtualised CESC multi-tenancy techniques evaluation*". H2020/5G-PPP SESAME project, March 2017.

2.13 VNF Placement Assessment

In the context of SESAME, a placement algorithm is one of the "key" components to manage the virtualized resources whit the aim of meeting the service agreements (SLAs).
The placement module interacts with NFV elements such as the SLAs catalogue, NFV orchestrator and the Virtual Infrastructure Manager. The information required to feed the placement algorithm is requested to those services.
From the SLA catalogue it is needed to gather information about the clients, their services and the performance level that has to be accomplished. From the NFVO, the current state of the network services is requested. And finally, an interaction with the VIM is needed to fetch the computational resources as well as its utilization.

To assess the accuracy of the model, a multi-level grading is proposed:

- **Level 1**: The basic model considers that each VNF uses a fixed amount of resources and the latency of each link is also fixed. This means that the model does not introduce variability due to the change on the load produced per VNF.

- **Level 2:** The next iteration adds variability in the load, in order to insert more complexity to the placement algorithm. In this iteration for each VDU of a VNF exists a model that reflects the amount of resources used for a specific usage rate. This includes parameters as CPU usage and computation latency. That will vary depending on the type of network service it is been provisioning, being the most representative parameter the traffic load (aggregated bitrate in the server).
  In addition, per each connection link the model has to consider the latency caused by the link-load. This model has to keep in mind parameters such as the total flow going through the link, the total capacity of each link and the overload that the current network service will add in it.

- **Level 3:** In this iteration the model adds a usage margin to the resources that the network service will use at maximum usage rate. Hence, techniques as Robust Optimization are used in order to guarantee the achievement of the service contains even in worst-case scenarios. In the proposed model deviation/protection parameter has been used.

Having the presented placement algorithm in mind, SESAME during its life time targeted few proof of concept (PoC) showcases to verify the functionality of the proposed solution.

- Integration of the placement algorithm into OpenStack VIM: the proposed placement algorithm has been integrated in a new service capable of interacting with the OpenStack components. The presented module considers all the above requirements (Level 3).

This module represents a proof of concept that set the implementation steps of the placement algorithm inside an OpenStack architecture. This module is indispensable to ensure the Quality of Service (QoS) / agreed SLA per tenant bases.

Based on pre-established modes of resource utilization and service provisioning, it is able to determine if a network service is able to be instantiated in the best position for the provider fulfilling the service requirements.

2.14 Key achievements

The key achievements of the project highlight the advances brought by the SESAME system and components to the 5G and Small Cell scenario.
The solutions developed during the SESAME project allow achieving significant and immediate advantages in the 5G ecosystem, in full compliance with SESAME initial objectives and its technical and functional requirements (requirements specification were performed within Task 2.1 in WP2).

Key achievements are brought by the SESAME architecture *"as a whole"*, as well by each of its components:

- The **CESCM**, a solution for management of edge services, fully interoperable with SESAME components.
- The **Light DC**, a virtualized cloud environment at the network's edge.
- The **Monitoring System** required for efficient (re)configuration of the system and to apply resource optimization algorithms through Self-x functionalities.
- The **SESAME NFVO**, in the edge virtual service orchestration.
- The **edge VNFS** developed during the project, and available in the SESAME's catalogue.
- **SFC**, supporting the chaining of multiple VNFs within a service without impacting negatively the performance.
- **Security features**, allowing the integration of multiple virtual operators (multi-tenancy) sharing the CESC provider's infrastructure, and the isolated and secure provision of vertical services.

The key achievements of the project are described in the table below (Table 2).
They will be assessed in three realistic test beds, which will show the advantages brought by SESAME in specific application contexts. With the demos, also SESAME use cases will be assessed.

Table 2: SESAME Objectives vs. Achievements performed and related KPIs

SESAME Objective	Achievement and Assessment KPI
To provide **multi-tenancy** support. The architecture should support the simultaneous operation of different tenants (e.g., VSCNOs with corresponding vEPCs) sharing the same infrastructure.	The architecture provides tools and features for the management of a multi-tenant infrastructure that leverages on an edge-computing cloud and on the deployment of Self-x and optimization procedures. The capability to share the CESCs among multiple tenants is demonstrated in the final demo, where multiple tenants have access to the Small Cell. **KPI:** Number of tenants (i.e. VSCNO) supported by the SESAME system.
To provide **Self-x features** that support autonomic network self-planning, self-optimisation and self-healing with the aim to optimise the resource	SESAME supports Self-x features for autonomic capabilities for optimized Radio Network capacity planning and operation mechanisms of the SCNO (Small Cell Network Operator). The proposed functional architecture supports the automation of the planning process and the sharing a given carrier among multiple tenants

assignment among tenants.	(RAN sharing). **KPI:** CESC set-up time and CESC cluster configuration. The Self-x features are measured by the capability of adaptive resource provisioning (i.e. rebalancing or repurposing) of network resources through Self-x features, and orchestration capabilities.
To deploy a **virtualized cloud environment** at the network's edge, supporting multiple virtualized services.	SESAME throughout its components support of multiple virtualized services in the Light DC which architecture allow their deployment over a fast time scale. The Light DC lowers the service latency by bringing and managing services closer to the end-users. The Light DC is made of the interconnection of micro-servers in each CESC inside a cluster is the environment for deploying VNFs and network services. This is a key feature of SESAME that unveils a platform capable of moving VNFs from the network core to the edge of the network. A placement module is provided and interacts with NFV elements such as the SLAs catalogue, NFVO and VIM. **KPI:** Number of different VNFs implemented Different types of VNFs deployed: e.g., vWatermarking, vGTP, etc. Number of VNFs or network services deployed in the CESC and CESC cluster
To support the **interoperability between SESAME system components,** through the exposure of management interfaces offering accessibility and configurability.	In SESAME components implement appropriate interfaces to support the management, accessibility and configurability: i) **CESCM** interfaces enable tenants to request network services (i.e. dashboard in the portal) and new resource instances; ii) **CESC** interfaces allow the management of the micro-server environment (e.g. VNF management) and small cell; iii) **VIM** acts as the intermediary element between the CESCM and the CESC in charge of managing the NFVI (e.g. VNF Forwarding Graph, or NFVFG). **KPI:** Accessibility of information from the SESAME system through management interfaces (i.e.: CESCM, VIM, CESC, EMS) and number programmable services/functions.
To implement the **monitoring of physical and virtual resources** required for efficient (re)configuration of the system in order to collect operative and performance measurements of the architecture components, and to apply resource optimization algorithms (e.g. through Self-x).	The resources monitored include both physical (e.g., CESC, network links) and virtual (e.g. VNFs) resources, allowing total visibility of SESAME infrastructure. Metrics specific for the three environments in which SESAME operates (Cloud, Radio and Network) were created. **KPI:** Amount of network traffic, number of connected UEs, small cell related performance indicators within a single CESC and/or CESC cluster on a per tenant VSCNO basis.

Monitoring of tenant metrics are also needed to determine the compliance with agreed SLAs.	
To support the **dynamic configuration of virtual resources**	The CESCM components (such as the NFVO, VNFM, VIM and the SDN controllers) are able to adjust and update the VNFs' parameters and network services configuration dynamically. **KPI:** Time to configure or reconfigure, boot and deploy individual VNFs or network services, memory and RAM requirements.
To provide **hardware and network acceleration** mechanisms for improving the VNFs performance.	SESAME adopts state-of-the-art (SOTA) solutions by means of both hardware and network accelerators. These include the use of a heterogeneous hardware architecture based on ARM SoC, and in particular on multi-core A53 ARMv8 64-bit processor, GPUs, DSPs and FPGAs. **KPI:** VNF performance in terms of processing speed and resource consumption. **KPI for Demo:** Use of HW acceleration in the deployment of services. Technology used for the micro-server (ARM and x86).
To provide **VNFs and service functions chaining** at the networks edge for the creation of composed services.	SESAME supports the chaining of multiple VNFs within a service without impacting negatively the performance. **KPI:** Number of available VNFs in the SESAME's catalogue, number of VNFs within a service chain that can be supported without impacting negatively the performance (e.g. network delay), number of available services.
Security and privacy: Secure access to all SESAME components has to be provided, as well as guaranteed data protection on a *per tenant* and *per slice* basis.	SESAME provides the secure integration of multiple virtual operators (multi-tenancy) sharing the CESC provider's infrastructure, allowing isolated and secure provision of vertical services for massive amount of connected UEs **KPI:** Type of security features developed or adopted for the SESAME system.

3 Demonstration evaluation

3.1 Test cases

3.1.1 Demo 1

This demo is focusing on Service Function Chaining (SFC), which is one of the very interesting features of Network Functions Virtualisation (NFV) relying on Edge Cloud Computing and SDN capabilities.

By utilizing a Light DC infrastructure integrated with Cloud-Enabled Small Cells, the orchestration, service setup and delivery, can be performed quite fast and to the edge of the network, thus improving the latency and being able to deliver various differentiated services over a smartly versatile and adaptable infrastructure.

The deployed scenario demonstrates the orchestration of a service chain over a single network segment with virtualization capabilities, but could be expanded to multiple such segments if an accommodating NFVO can support such functionality, as in the case of the SESAME project[115] and in SESAME demo 2.

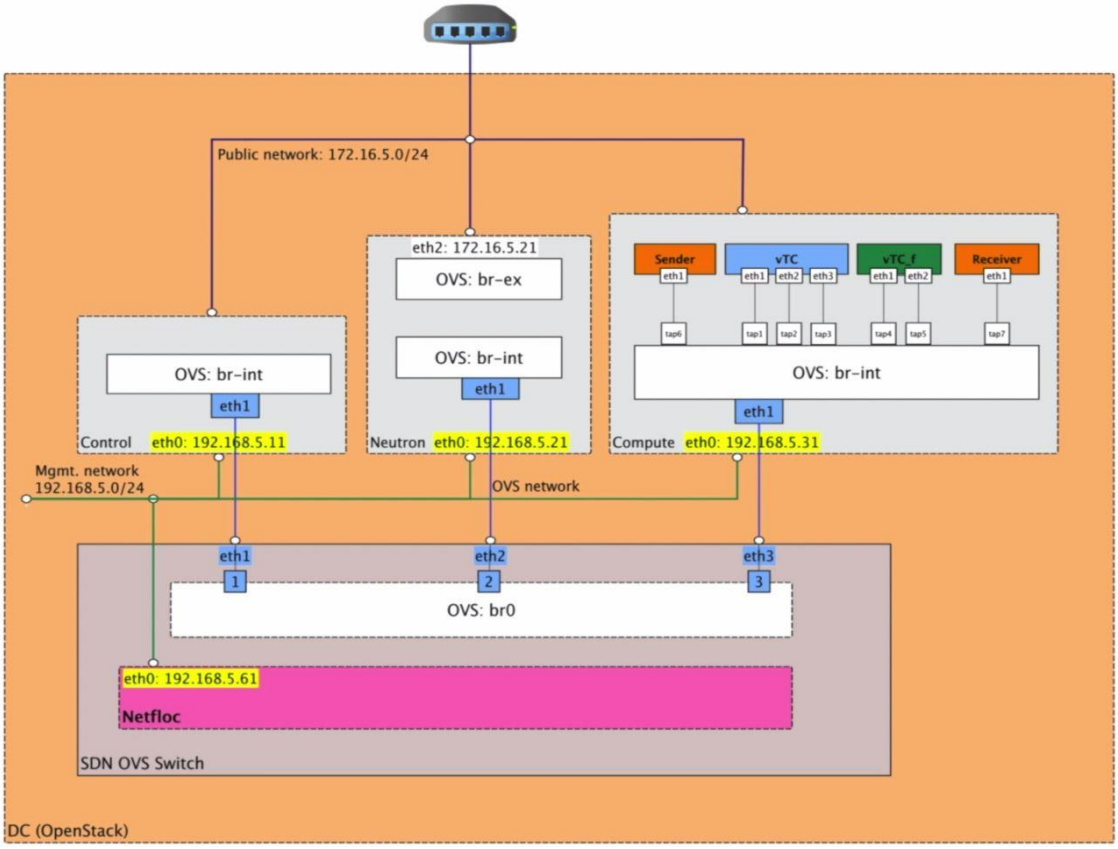

Figure 20: Demo 1 topology

The topology of the above figure (Figure 20) uses four nodes:

- OpenStack Controller 1, which contains all basic OpenStack services apart from Networking.

[115] I. Trajkovska (editor): Deliverable D6.3: "*Service Management and Orchestration functions, including VNF models - Final*". H2020/5G-PPP SESAME project, September 2017.

- OpenStack Networking (Neutron) Controller.
- OpenStack Compute node: which hosts the VMs.
- SFC controller, which hosts Netfloc and operates as an SDN switch. Additionally this host contains the Netfloc exporter, which exports metrics to Grafana[116], also hosted in this host.
- All nodes are connected by using Open vSwitch[117].

The service chain VNFs used in this scenario is Traffic Classification and forwarding, where the first performs deep packet inspection (DPI) and the second traffic analysis and forwarding to relevant ports.

Initially the OpenStack Heat[118] provisions the resources and then Netfloc[119] creates the service chain. Once the service chain is deployed, the VNFs are initiated and traffic is sent, which is captured by the Traffic Classifier and then fed to the forwarder, which rewrites the relevant MAC address accordingly and then routes the packets.
Meanwhile, the metrics are also fed through the Prometheus Netfloc exporter to Grafana, for real time in depth monitoring of the user traffic and the Service Function chain dedicated flows, as it is shown in the figure below (Figure 21).

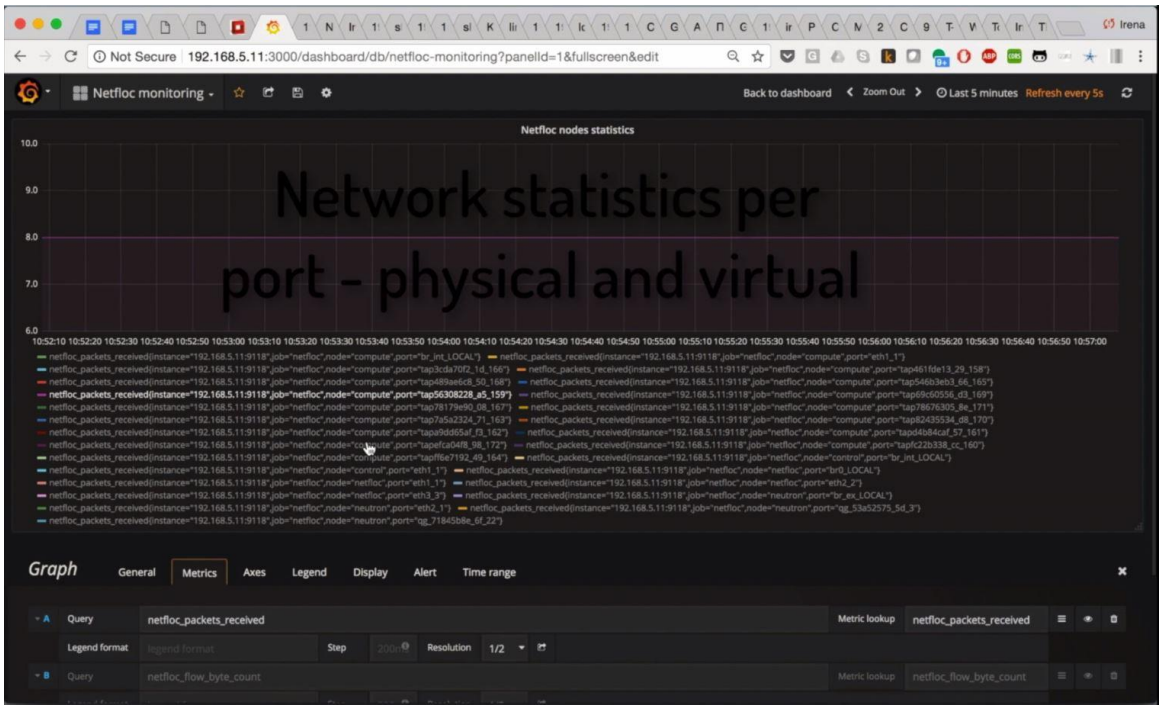

Figure 21: SFC metrics in Grafana

3.1.2 Demo 2

One of the key innovations in SESAME is to provide advanced SDN-*geared* network intelligence and applications relying on Network Functions Virtualisation (NFV) and Edge Cloud Computing technologies. The concept of Cloud-Enabled Small Cell (CESC) with integrated Light Data Center (DC) cloud platform is the underlying infrastructure for the execution of novel applications and services inside the access network infrastructure.

[116] See: *https://grafana.com/.*
[117] See: *http://openvswitch.org/.*
[118] For more details see, *among others*: *https://wiki.openstack.org/wiki/Heat*.
[119] For more details see: *http://icclab.github.io/netfloc/*.

The focus of this demonstration scenario is placed upon elaborating the network slicing concept applied, via using MOCN to a single CESC, which is deployed by an infrastructure provider, that is the Small Cell Network Operator (SCNO), over a certain geographical area and is shared among several tenants, that is the Virtual Small Cell Network Operators (VSCNO), that offer mobile communications services to their own customers.
The sharing is sustained in network slicing principles, so that each VSCNO is provided with a network part, that is an instantiated logical segment able to provide a specific capacity with potentially specific capabilities (ranging from just different network policy settings to the support of different protocols and features), realised together with other slices on the common physical infrastructure. Moreover, this scenario will serve as a Proof of Concept of the proposed CESC provider supporting different tenants' requirements over the virtualised infrastructure (Small Cell + Light DC).
In this sense, the resource abstraction model and its use by the VIM and CESCM will be shown.

To demonstrate the capabilities of the platform, let us assume that an end-user requests a video from a server on the Internet. The VSCNO decides to impose a video encoding format for all video traffic to its end-users.
When the answer to the user request reaches the micro-server, it is redirected to the SC VNF where the data is de-capsulated from the GTP; then it is transferred to a virtual Transcoding Unit (vTU) where the processing is offloaded to a hardware accelerator.
Once the transcoding is complete, the data is sent back to the SC VNF for encapsulation and then to the SC common VNF in order to be relayed to the end-user.

This SESAME PoC also illustrates the establishment of the complete chain of monitoring, decision-making and reaction. In this case, CESCM as a module with the over top view of both radio and cloud side of the ecosystem will monitor cloud/radio parameters (e.g., CPU/RAM usage, call drop rate, etc.).
If a violation occurs, CESCM via processing the monitoring inputs will be able to detect and then appropriately react upon the situation. The decision-making process might be a simple threshold checking or a complicated multi parameter cognitive method.
In the same way, the reaction ranges from the complete NS scaling, to the NS scaling up/down in/out, to the service function chain changes, to the change on a radio parameter (e.g., dedicated bandwidth to a VSCNO). For instance, assume a case where the agreed SLA with a VSCNO determines the support of a certain number of hits on an edge service.
Then, because of a flashy event more hits for that specific service comes, the system should be able detect the case and try to keep the SLA promise by taking an appropriate reaction.

3.1.3 Demo 3

The testbed used for Demo 3 is located at Italtel's (ITL) labs and it will used to demonstrate several outcomes of the Light DC development carried out in SESAME WP4:

1. Hypervisor, Hardware acceleration layer and Accelerated Virtual networking (VOSYSwitch) integration on the NXP platform.

2. OpenStack integration (the Kilo Release[120] is used): computing and networking managed by OpenStack controller in case of Hybrid node (Intel/ARM).

3. VNFs deployment in a compute node based on:
   a) Intel, CPU-*only*;
   b) Intel, CPU+GPU;
   c) ARM, CPU-*only*.

   VTU (Video Transcoding Unit) performance comparison by using configurations a)), b)) and c)), in case of running single session of HD720 (720x1280 pixel) H.264 encoding.

[120] For further details, see: *https://www.openstack.org/software/kilo/*.

The testbed architecture used for Demo 3 is depicted in Figure 22.

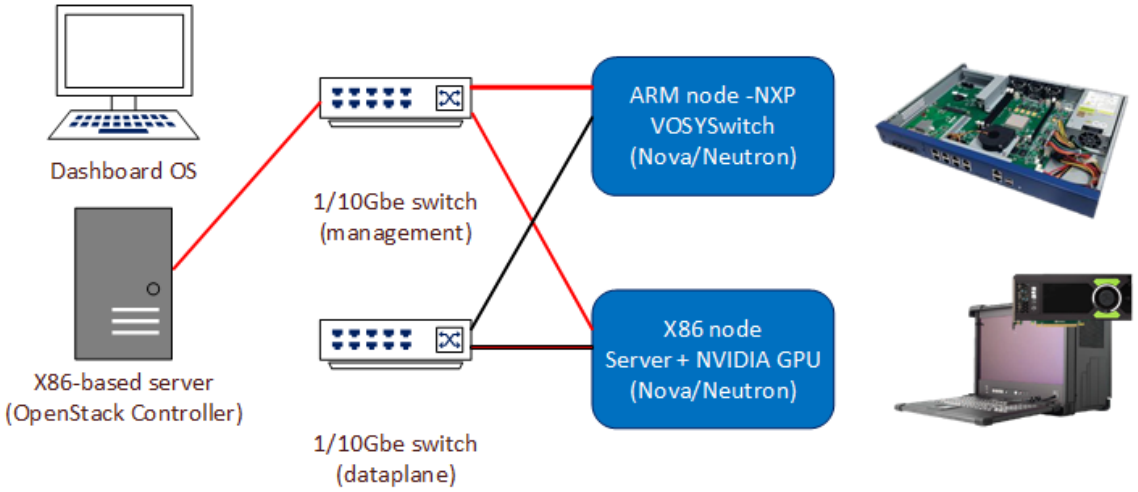

Figure 22: Testbed for Demo 3

4 Roadmap

4.1 General Concerns

The recent advances in Information Technologies (IT), the dispersal of ultra-broadband (fixed and radio) connectivity, the continuous reduction of hardware costs and the wider and wider availability of open source software solutions, are creating the conditions for introducing a deep innovation in the architectural design and in the operations of future telecommunications networks and services.

We are witnessing a period of rapidly growing interest on the part of industry and academia in Software-Defined Networks (SDN) and Network Function Virtualization (NFV). The growing interest in these paradigms (re-proposing principles that have been "well-known") is most probably motivated by the novelty of the overall context, specifically their techno-economic sustainability and high-level performance.
Thanks to these techno-economic trends, SDN and NFV principles will soon impact not only current telecommunications fixed and mobile networks, but also service and application platforms. In fact, SDN and NFV, together with Cloud, Edge and Fog Computing, can be seen as "facets" of a broad innovation wave, called as *"softwarization",* which will contribute to automating processes, optimising costs, reducing time-to-market, providing better services. At the same time, the Internet of Things (IoT), Tactile Internet, Machine Type Communications (MTC), Cloud Manufacturing, Cloud Robotics, etc. will generate a new plethora of services and applications, ranging from industrial (e.g., Industry 4.0) and mission critical ones to precision agriculture, to Smart Cities, etc.
5G (which is not just one-step beyond 4G) will be the main "collector" of this coming wave of innovation, bringing a number of different technologies to maturation, convergence and socio-economic impact, thus accelerating the transition towards the Digital Society and the Digital Economy. Europe should be prepared to face this important transformation, ready to capture all the socio-economic opportunities that it will bring.
In order to "bring" this vision into reality, Europe needs developing and aggregating a "critical mass" of 5G (experimental) facilities, capable of exploiting synergies and collaborations, both within and outside Europe, e.g., with similar initiatives worldwide, *to the extent possible*[121].

4.2 Approach

In order to provide a reliable response to the challenges previously identified and so to "affect" the design and the structuring of the future 5G-*oriented* network architecture, the EU-*funded* SESAME project has strongly promoted the concept of the SCs virtualization and their utilisation and partitioning into logically isolated "slices", able to be offered, *upon demand*, to multiple operators/tenants.
The main aspect of this innovative feature (i.e.: "*multitenancy*") is about the capability to "accommodate" multiple operators under the same infrastructure, by satisfying the profile and requirements of each operator separately. In this context, operators can "add" to their services portfolio the option of being asset providers (infrastructure, network functions and platform(s), all offered "as-a-service") for other operators and/or other actors/players like integrators[122]. This significantly reduces the cost of deployed infrastructure (i.e., cost of ownership, maintenance, etc.), since hosted SCs can be treated as *"an operating resource instead of a capital expenditure"*.
Under this specific perspective, the creation of neutral host solutions comes to "address" also the economic viability of telecom investments.

[121] There are several ways for pursuing this aggregation of 5G experimental facilities; for example by interconnecting test-beds and field-trials at different levels (hierarchically), by making interoperable open and closed experimental islands, by exploiting the service "platforms of platforms" on top of virtual labs, etc.

[122] Leveraging on this important design principle will "push the single digital market further", also "paving the way" for virtual pan-European operators relying on nationally deployed infrastructures.

Furthermore, *until today*, network equipment deployed at the edge and access network part had well specified "hard-wired" functionalities that were not possible to be re-purposed.
With the advent of Cloud Computing, SDN and NFV, the idea to have general-purpose computing and storage assets at the edge of mobile networks has further matured[123].

In this direction, new industry initiatives have already introduced the concept of Mobile-Edge Computing (MEC)[124] and the related "key" market drivers. To further enhance the virtualisation capabilities of the Small Cell deployment, so that to include not only network capacity resources but also edge processing capabilities, a micro scale virtualised execution infrastructure has been proposed by SESAME, in the form of a "Light Data Centre" (Light DC). The Light DC concept can be considered in order to build a clustered infrastructure with high manageability and can further be optimised to reduce power consumption, cabling, space and cost. This aspect not only optimizes end-users' experience(s) with respect to performance issues, but also "promotes" new opportunities (i.e. via the provision of an ecosystem with novel services residing inside the network infrastructure).

To realise SESAME's vision the "*Cloud-Enabled Small Cells*" *(CESCs)* has been designed, developed and implemented with the aim of providing access to network capacity, coupled with MEC resources in a single device. These resources can be offered on-demand to interested Communications Service Providers (CSPs), profiling both access and edge computation resources to satisfy the specific CSPs' needs. From the perspective of service provisioning, the proposed approach can be used to provide edge cloud capabilities and so to enable accelerated services, content and application due to the increased network responsiveness.

Operators may provide the network's edge (i.e., the Light DC) to third-party partners, allowing the rapid deployment of cutting-edge services to users and enterprises, translating to added-value and creating opportunities for vendors, service providers and operators by enabling them complementary and advantageous positions. Besides, the Light DC also "enables" the rapid on-demand deployment of cutting-edge network services in the form of VNFs, such as data processors, security appliances, proxies, media transcoders, M2M gateways, etc., close to the mobile nodes. Locating virtual service processing nodes closer to users reduces latency, improves throughput, and reduces traffic in the network core. By "slicing down" a single CESC (or a CESC cluster) it could be feasible to furnish and allocate resources to different operators.

Furthermore, solutions for aggregation of data, transcoding of video content with optimised delivery in edge networks and caching at the very edge of the network, also enable a reduction in transport time and, *therefore*, are to provide a vital route to successfully reducing service-level latency. For this, a CESC platform also includes functions for cache management and placement, for edge-optimised content delivery, aggregation at the edge through placing scenario-*specific* computation closer to the infrastructure (such as 5G point of attachments) and other relevant entities.
From an engineering point of view, another key challenging issue of the original SESAME approach was the overall optimal configuration of the CESC infrastructure, including resource allocation and re-configuration in case of faults or abrupt demand changes. Given the multi-layer nature of the SESAME architecture as well as the heterogeneous networking and computing assets which are involved and assigned to multiple tenants with diverse needs and constraints, the optimal configuration problem become significantly complex. Typically, this problem can be addressed by "moving" some of these dependencies and layer interactions into a central place, in order to do a *per-box* configuration. Apart from having to understand all layer dependencies and interactions, which becomes impossible for a human network operator with the current infrastructure interdependencies, such an approach thought to be "tailored" only to specific circumstances. To that end, SESAME has provided an added value characteristic where it has enabled autonomic networking towards providing holistic *self-x* characteristics (where "*x*" could stand for "*configuring*", "*healing*", "*optimising*", etc.) to the proposed CESC devices.

[123] Hossain, E., and Hasan, M. (2015, June): 5G cellular: key enabling technologies and research challenges. *IEEE Instrumentation and Measurement Magazine, 18*(3), pp.11-21.
[124] European Telecommunications Standards Institute (ETSI) (2014, September): *Mobile-Edge Computing, Introductory Technical White Pape*r. Sophia-Antipolis, France: ETSI.

In this way, SESAME has become capable of providing a distributed network management system that can be developed even from third-party providers.

The key innovations proposed by the SESAME project focus upon the novel concepts of virtualising Small Cell networks by substantially evolving the SC concept under the paradigms of a multi-operator (multi-tenancy) enabling framework and an edge-*based,* virtualised execution environment.
The proposed CESC includes a multi-tenancy platform able to provide the radio access to support the required wireless capacity in a certain area.
In addition, cloud-*based* computation resources are provided through a virtualised execution platform (i.e. the Light DC) which is used to support the required VNFs, implementing the different features/capabilities of the Small Cells and the cognitive/self-x management operations, as well as the computing support for the mobile edge applications of the end-users.
The overall management and coordination of VNFs is assigned to higher layers. Depending on the actual virtualisation capabilities, clusters can be assigned to one or more Virtual Infrastructure Managers (VIMs) –*i.e., entities that will be responsible for managing the virtualised resources required for proper VNF deployment*.
Monitoring data from all active VIMs will be combined for managing the whole process of VNF restructuring (e.g., migration, rescaling, etc.) in a dynamic and efficient way, cooperating with the CESC Manager (CESCM).
The CESCM is responsible for coordinating and supervising the use, the performance and the delivery of services. It "controls" the interactions between the infrastructure (CESC) provider/owner and the network operator. Accordingly, the Service Level Agreement (SLA) negotiation (encompassing billing issues, accounting and so on) interacts through appropriate open software, with the existing support system of the telecom operator.
Also, on an architectural basis, CESCM will encompass monitoring and analytics, as fundamental tools for efficiently managing the network.

The proposed solutions also focus on extending the portability of virtual functions from closed vertically integrated architectures to open hardware and software architectures.
SESAME's Light DC is based on heterogeneous, parallel architectures, featuring low-power 64-bit processors supported by several hardware accelerators. This infrastructure can be used for the deployment of VNFs, for executing small-size applications, offloading end-user mobile applications and functions, plus for the support of more powerful self-coordinating functionalities.
SESAME's scope does not really restrict to CESC development only, but comes to address a number of important challenges in the network management field, such as providing multi-tenancy through virtualisation techniques, developing novel edge-computing architectures and deploying self-x and optimisation procedures directly to network's edge.

4.3 SESAME Innovations

Europe is home of a growing number of 5G private experimentation and trials (pre-commercial and commercial) involving a diversity of stakeholders and particularly network operators, manufacturers/vendors and some vertical actors . The central target of related trials was -and still is- is to demonstrate the high data rates and low latency communications, which are "key features" for 5G technology. Several "initial" 5G-*related* trials have been focused on enabling technologies related to: (i) The radio interface (high throughput, millimetre-waves and other new large spectrum bands, antenna technologies, etc.); (ii) the network architecture (emphasizing upon virtualization, cloudification, network slicing, edge computing as well other related aspects), and; (iii) the introduction of new technologies dedicated to specific use cases.
It is foreseen that when the maturity level of 5G features increases, more direct vertical stakeholders will be included for trials. Some of the 5G trials could include "joint work" on experimentation platforms that could become open to new ecosystems, in order to develop 5G applications and services in the context of the digital transformation of vertical industries.

SESAME has purely followed the above aspect as, in fact, most of his priorities have been around the above mentioned initiative (i)-(iii). SESAME has identified specific KPIs as well as has developed specific solution to fulfil the identified innovative features. Benefits have already been summarized as in *Section 4.2*, above.
Via the project selected use case (as is the context of the entire descriptive framework of WP2 and, *in particular*, according to the findings discussed/summarised in D2.5) and; via the selected demos together with all PoC-*related* activities (as discussed within the entire descriptive framework of WP7 and, *in particular*, according to the findings of the deliverables D7.1[125], D7.2, D7.3 and D7.4), SESAME has promoted several innovative solutions for the benefit of the 5G-*oriented* market.
Most -if not all- of these solutions become quite "attractive" and beneficial for the involved market actors/players. In particular, the selected demos 1-3 discussed in Section 3 of the present deliverable together with other technical findings of the project seem as capable of "promoting innovation of technical and market aspects" and of "introducing benefits to the involved actors"[126].

Section 2.1.14 of the present deliverable has highlighted the key achievements of the SESAME project as these have been brought by the SESAME system.
The solutions developed during the SESAME project allow achieving significant and immediate advantages in the 5G ecosystem, in full compliance with SESAME initial objectives and its technical and functional requirements. (Detailed information is given in the contents of Table 2).

The SESAME Deliverable D8.4[127] has provided information of "prime" importance as of market analysis, roadmapping and potential business modelling, affecting the entire SESAME progress.
Furthermore, the SESAME Deliverable D8.7[128] in its dedicated *Section 10 ("Exploration Activities and Updated Plans")* has identified intended as well potential forthcoming activities to take place by the interested SESAME partners, so that to promote and/or to implement, *per case*, various SESAME-*related* findings and innovations.

5G networks are expected to provide improved performance, high economic and social value[129]. The 5G-PPP association has identified a series of requirements and KPIs[130] for 5G networks. Although these requirements cannot be achieved in a single case, there are several initiatives and projects trying to address a sub-set of them each time through different architectures, configurations and technologies. Among others, the most frequently discussed performance indicators are ultra-high data rates, very low latency and increased reliability. Simultaneously, 5G networks arise new opportunities to existing as well as to new players by lowering the barriers to entry and by facilitating the development of new advanced applications. In addition, 5G networks will influence the so called as "verticals" that are likely to enter the value chain and generate revenues. This will result in the creation of numerous new job positions contributing to countries' economy by increasing their GDPs. Finally, it is anticipated that 5G networks will influence / improve several societal aspects such as medical care, transportation, human wellbeing, entertainment etc. However, in order to achieve such changes, the social acceptance of 5G innovations is required.
It is thus evident that 5G networks success depends on several different factors (technical, economic and social) leading to increased complexity. By taking into account, the high number of unknowns, concepts and technologies, it is straightforward to show that 5G investments is a very risky venture. In order to

[125] I. Giannoulakis and A. Kourtis (editors.): *Deliverable D7.1:* "*Proof-of-Concept Integration and Validation Plan*". H2020/5G-PPP SESAME project, December 2016.
[126] Also see: NetWorld 2020 - Expert Advisory Group of the European Technology Platform Networld 2020 (2016, July*): Strategic Research and innovation Agenda (SRIA)*. For more details see: *https://www.networld2020.eu/sria-and-whitepapers/. (New SRIA on pervasive Mobile Services)*.
[127] I. Neokosmidis (editor): *Deliverable D8.4:* "*Market Analysis, Roadmapping and Business Modelling Report*". H2020 SESAME project, December 2016.
[128] I. Neokosmidis (editor): *Deliverable D8.7:* "*Techno-Economic Analysis and Commercialisation Plans*". H2020 SESAME project, December 2017.
[129] European Commission (2014): *Identification and quantification of key socio-economic data for the Strategic Planning of 5G introduction in Europe. (SMART 2014/0008 Study)*. See: *https://ec.europa.eu/digital-single-market/en/news/identification-and-quantification-key-socio-economic-data-strategic-planning-5g-introduction*.
[130] For further details see, *for example*: *https://5g-ppp.eu/kpis/*.

navigate such a multi-dimensional landscape and reach a viable economic solution and successful innovations, a clear technology, economic and regulatory roadmap is needed.

SESAME / 5G networks roadmap is aiming to identify and prioritize the factors that will influence their market adoption and evolution. This assessment has been performed through a number of conducted surveys using the Fuzzy Analytic Hierarchy Process (AHP) methodology[131].
More specifically and according to the SESAME D8.4, as of roadmapping activity, factors affecting market adoption and evolution of SESAME and 5G networks have been identified and assessed by experts within SESAME project, via the usage of the Fuzzy AHP. A survey has been conducted there, in order to evaluate the relative importance of these factors and classify them. The survey has been completed by 16 experts from several European countries belonging to a variety of different sectors including industry, research institutes and academia and having a professional background in telecommunication technologies. The derived results showed that Business is rated as the most important criterion for 5G deployment followed in turn by Performance, Flexibility and Acceptance; whilst Technology has the lowest weight. Based on global priorities, the most weighted sub-criterion has been identified as that of cost reduction, followed by low latency, optimized and more dynamic usage of all distributed resources, high reliability and new market opportunities.

The obtained results, as depicted in all related SESAME deliverable fulfilling the context depicted in Table 2 of the present deliverable, will be a valuable tool for SESAME partners, stakeholders and decision-makers in order to adapt their business strategies towards 5G business opportunities.

[131] For further details see, *among others*: Wang, Y.-M., and Chin, K.-S. (2011, June): Fuzzy analytic hierarchy process: A logarithmic fuzzy preference programming methodology. *International Journal of Approximate Reasoning 52*(4), pp.541-553.

5 Conclusions

In this document, we have presented the overall assessment and its methodology of the advances coming from the SESAME project "as a whole" as well as from its individual components (CESCM, Light DC, PNF, Monitoring, NFVO, VNFs, SFC, etc.) according to the developed architecture and the corresponding results, providing verification to the corresponding architecture.

Additionally, there has been performed a high-level security evaluation and propositions on "how to tackle possible threats".

Apart from the above, we have described three different demo test cases, presenting different aspects of the various components of the SESAME architecture, with the infrastructure employed for each demo and by providing info about how the proposed architecture is integrated into the respective infrastructure. The two demos (i.e., demo 1 and demo 3) are focusing upon more specific functionality (SFC, HW acceleration, ARM integration, vTU) have required for "partial integration" with the wider SESAME framework; together with a large scale demo (as in demo 2) they provide an almost complete representation of the whole SESAME architecture.

The present work also aims to "act" as a demonstration of a proof-of-concept (PoC) of the operation of the SESAME architecture and, *additionally*, as a sort of "guide" for any "third person" (legal or physical one) who desires deploying the SESAME architecture as well as checking SESAME operation and ITS advantages.

Finally, a roadmap has also been discussed on how an organisation can follow the 5G advances by adopting the SESAME architecture "as a whole" or its individual components according to their needs and the next steps that could be further taken to "advance" the SESAME architecture as well.

Bibliographic References

The following documents and/or websites are proposed for further reading.

Documents:

- 3GPP TS 22.011: "*Technical Specification Group Services and System Aspects; Service accessibility (Release 10)*", December 2012.

- 3GPP TS 23.251 v13.1.0: "*Network Sharing; Architecture and functional description (Release 13)*", March 2015.

- 3GPP TS 32.306: "*Configuration Management (CM); Notification Integration Reference Point (IRP): Solution Set (SS) definitions (Release 10)*", September 2010.

- 3GPP TS 32.316: "*Generic Integration Reference Point (IRP) management; Solution Set (SS) definitions (Release 10)*", June 2017.

- 3GPP TS 32.435 v14.0.0: "Performance measurement; eXtensible Markup Language (XML) file format definition (Release 14)", April 2017.

- 3GPP TS 32.606: "Configuration Management (CM); Basic CM Integration Reference Point (IRP); Solution Set (SS) definitions (Release 10)", September 2010.

- 3GPP TS 32.662: "*Configuration Management (CM); Kernel CM Information Service (IS) (Release 10)*", March 2011.

- Bradnerand, S., and McQuaid, J. (1999, March): *Request for Comments (RFC) 2544: "Benchmarking Methodology for Network Interconnect Devices"*.

- Broadband Forum (2013, November): *TR-069: "CPE WAN Management Protocol, Issue: 1, Amendment 5*".

- Cheung, N.M, Fan, X., Au, O.C., and Kung, M.C. (2010, March): Video Coding on Multicore Graphics Processors. *IEEE Signal Processing Magazine*, *27*(2), pp.79-89.

- Chochliouros, I.P., Spiliopoulou, A.S., Kostopoulos, A., Belesioti, M., Sfakianakis, E., et *al.* (2017): *Putting Intelligence at the Network Edge through NFV and Cloud Computing: The SESAME Approach.* In G. Boracchi et *al.* (eds.), EANN 2017, CCIS 744, pp.704-715. Springer International Publishing AG.

- Comi, P., Secondo-Crosta, P. Beccari, M., et *al.* (2016): "Hardware-accelerated high-resolution video coding in Virtual Network Functions". In Proceedings of the European Conference on Networks and Communications 2016 (EuCNC-2016), pp.32-36, Athens, Greece, *June 27-30, 2016*.

- European Commission (2014): *Identification and quantification of key socio-economic data for the Strategic Planning of 5G introduction in Europe. (SMART 2014/0008 Study)*.

- European Telecommunications Standards Institute (ETSI) (2014): "*NFV Management and Orchestration - An Overview", GS NFV-MAN 001 v1.1.1*". European Telecommunications Standards Institute, Sophia-Antipolis, France.

- European Telecommunications Standards Institute (ETSI) (2014, September): "*Mobile-Edge Computing, Introductory Technical White Paper*". Sophia-Antipolis, France.

- Hossain, E., and Hasan, M. (2015, June): 5G cellular: key enabling technologies and research challenges. *IEEE Instrumentation and Measurement Magazine, 18*(3), pp.11-21.

- ITU-T (1992, January): "*Recommendation X.731: Information Technology (IT) – Open Systems Interconnection (OSI) – Systems management: State management function*".

- Paglierani, P., Grossi, G., Pedersini, F., and Petrini, A. (2016): "GPU-based VP8 encoding: Performance in native and virtualized environments". In Proceedings of the International Conference on Telecommunications and Multimedia 2016 (TEMU-2016), pp.1-5, Heraklion, Greece, *July 25-27, 2016*.

- Pérez-Romero, J., Sánchez-González, J., Sallent, O., and Agustí, R. (2016): "On Learning and Exploiting Time Domain Traffic Patterns in Cellular Radio Access Networks". In Proceedings of the 12th Int. Conf. on Machine Learning and Data Mining (MLDM), pp. 501-515, Springer, New York (2016).

- RapidMiner Studio, *http://www.rapidminer.com*.

- SESAME Deliverable 2.1: "*System Use Cases and Requirements*". H2020/5G-PPP SESAME project, December 2015.

- SESAME Deliverable 2.2: "*Overall System architecture and Interfaces – First Iteration*". H2020/5G-PPP SESAME project, April 2016.

- SESAME Deliverable 2.5: "*SESAME Final Architecture and PoC Assessment KPIs*". H2020/5G-PPP SESAME project, December 2016.

- SESAME Deliverable 3.1: "*CESC Prototype design specifications and initial studies on Self-X and virtualization aspects*". H2020/5G-PPP SESAME project, June 2016.

- SESAME Deliverable 3.2: "*Self-X features and virtualised CESC multi-tenancy techniques evaluation*". H2020/5G-PPP SESAME project, March 2017.

- SESAME Deliverable 3.3: "*Framework of a distributed network management system capable to host and run Self-X functionalities*". H2020/5G-PPP SESAME project, June 2017.

- SESAME Deliverable D3.4: *"CESC Small Cell prototype and PoC"*. H2020/5G-PPP SESAME project, June 2017.

- SESAME Deliverable 4.1: "*Light DC architecture design*". H2020/5G-PPP SESAME Project, June 2016.

- SESAME Deliverable 4.2: "*Virtualisation extensions for Acceleration of Light DC capabilities*". H2020/5G-PPP SESAME project, December 2016.

- SESAME Deliverable 4.3: "*Techniques for efficient VNF Deployment with relevant VIM extensions, Evaluation framework*". H2020/5G-PPP SESAME project, June 2017.

- SESAME Deliverable 4.4: "Light DC Prototype". H2020/5G-PPP SESAME project, *June 2017*.

- SESAME *Deliverable D5.1* "*Description of CESC abstraction model*". H2020 SESAME project, June 2016.

- SESAME Deliverable 5.2: "*VIM and CESC Implementation*". H2020/5G-PPP SESAME project, September 2017.

- SESAME Deliverable 6.3: "*Service Management and Orchestration functions, including VNF models – Final*". H2020/5G-PPP SESAME project, September 2017.

- SESAME Deliverable 6.4: "*Orchestrator Prototype*". H2020/5G-PPP SESAME project, September 2017.
- SESAME Deliverable D7.1: "*Proof-of-Concept Integration and Validation Plan*". H2020/5G-PPP SESAME project, December 2016.
- SESAME Deliverable D7.2: "*Integrated CESC Prototype Validation*". H2020/5G-PPP SESAME project, July 2017.
- SESAME Deliverable D7.3: "*Experimental Integration results of HW/SW modules of the overall SESAME framework*". H2020/5G-PPP SESAME project, September 2017.
- SESAME Deliverable D7.4: "*Integrated Pilot and Evaluation Report*". H2020 SESAME project, December 2017.
- SESAME Deliverable D8.4: "*Market Analysis, Roadmapping and Business Modelling Report*". H2020 SESAME project, December 2016.
- SESAME Deliverable 8.7: "*Techno-Economic Analysis and Commercialisation Plans*". H2020 SESAME project, December 2017.
- Small Cell Forum (SFC) (2016, January): *SCF 159.06.02:* "*Small Cell Virtualization: Functional Splits and Use Cases*".
- Wilson, R.A. and Keil, F.C. (1999): *The MIT Encyclopedia of the Cognitive Sciences*. MIT Press.

Websites:

- ARM web page: *https://www.arm.com/products/system-ip/amba-specifications.php*.
- Ceph web page: *http://ceph.com/*.
- NXP web page: *http://www.nxp.com/products/microcontrollers-and-processors/arm-processors/qoriq-arm-processors:QORIQ-ARM*.
- OpenStack wiki for Cinder: *https://wiki.openstack.org/wiki/Cinder*.
- OpenStack wiki for Neutron component: *https://wiki.openstack.org/wiki/Neutron*.
- OpenStack wiki for Nova component: *https://wiki.openstack.org/wiki/Nova*.
- OpenStack wiki for Neutron component: *https://wiki.openstack.org/wiki/Neutron*.
- NVIDIA NVENC Programming Guide: *https://developer.nvidia.com/nvenc-programming-guide 2017.03.27*.
- WebM Video Hardware RTLs. Available from: *https://www.webmproject.org/hardware/ 2017.03.27*.
- Wiki page: *https://en.wikipedia.org/wiki/Logical_Volume_Manager_(Linux)*.

6 Annex A: Quality of Experience (QoE) Assessment Framework

6.1 Introduction

5G is a compact two-lettered word encapsulating a very large concept. Indeed, capturing a definitive understanding of 5G (short for "fifth generation") networking is elusive, since 5G covers so many different aspects. The "fifth generation" of telecommunications systems, or "5G", will be the most critical building block of our "digital society" in the next decade. 5G networks are expected to be the perfect means to accommodate the high number of heterogeneous connected devices, increasing traffic needs and tight performance requirements. It is anticipated that 5G networks will offer ultra-high data rates (e.g. 1-10 Gb/s to end-users), very low latency (<1msec), high density of heterogeneous devices and increased reliability.

The next generation of communication systems, as assessed within the "5G" context, will be the first illustration of a truly converged network environment where wired and wireless communications will make use the same infrastructure, "driving further" the future networked society. It will offer virtually ubiquitous, ultra-high bandwidth, "connectivity" not only to separate users but also to (Internet-) connected objects.
Therefore, it is assumed that the future 5G infrastructure will "serve" a multiplicity of services/applications and domains/sectors also including professional uses (e.g. assisted driving, eHealth, energy management, possibly safety applications, etc.).

However, the next generation of mobile networks will not be just an extension of existing networks that will merely provide improved performance and speed.
It is expected that it will bring innovative and exceptional network and service capabilities, in common with modern applications and related services/facilities.
In addition, 5G will be a "key enabler" for the Internet of Things (IoT) by offering a suitable and modern "platform" to connect an enormous number of sensors, rendering devices and actuators -or other sort of equipment- with strict energy and transmission constraints.

Other features of 5G which are the reason why there is such a global interest in 5G research currently occurring are: Fixed-mobile convergence (FMC), i.e. seamlessness between the traditional fixed access network (e.g. fibre-to-the home, FTTH) and the mobile communications network; Device-to-device (D2D) communications and *ad-hoc* meshing (e.g. for sharing of content and social media in localised public spaces); Open Access.

Moreover, mission critical services requiring very high reliability, global coverage and/or very low latency, which are up-to-now handled by specific networks, typically public safety, will become natively supported by the 5G infrastructure.
It is thus evident that 5G will safeguard user quality of experience by providing continuity in challenging situations such as high mobility (e.g. in trains), very "dense" or "sparsely populated" areas, and journeys covered by heterogeneous technologies.

6.2 QoE Terminology

The meaning of QoS has been originally introduced in 6.7 as "*The collective effect of service performance which determine the degree of satisfaction of a user of the service*". In [2], this definition has been further expanded with technical aspects such as "*…service support performance, service operability performance, serveability performance, service security performance and other factors specific to each service*", but also with no technical such as *"…the term (QoS) is not used to express a degree of excellence in a comparative sense nor is it used in a quantitative sense for technical evaluations. In these cases, a qualifying adjective (modifier) should be used*". In 1999, Kalevi Killki stated that QoS "*is used to define the network's capability to meet the requirements of users and applications*" [3].
Current definitions for QoS requirements for multimedia services can be found in [4].

However, in later years the meaning of QoS has been moved towards pure technical issues. In **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**, authors stated that "*QoS provided by a transport service is defined in terms of the way a traffic stream is affected when it is transported through the network. This is typically in terms of the probability of cell loss, delay, and cell delay variation*".

The pure technical approach of QoS continued in recent years [6] with the definition of QoS "*as …. the ability of the network to provide a service at an assured service level*".
The loss of original meaning of QoS (despite its original definition) lead to an evolving definition of the new term Quality of EXperience ("QoE" or "QoX") aiming to address the need for a new end-to-end user-centred quality concept.

In [7], QoE has been defined "*…as an extension of the traditional QoS in the sense that QoE provides information about the delivered service from an end-user point of view*" while Soldani in [6] defines "*QoE is how a user perceives the usability of a service when in use –how satisfied he/she is with a service in terms of, e.g., usability, accessibility, retainability and integrity*".

Towards this direction ITU introduced the following definition of QoE: "*QoE is defined as overall acceptability of an application or service, as perceived subjectively by the end-user. It includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.). Overall acceptability may be influenced by user expectations and context*" [8].
In [4], "*QoE is … (the) user perceived experience of what is being presented by a communication service or application user interface*".

According to DSL Forum [9], "*QoE and QoS terminology are often used interchangeably but are actually two separate concepts*". QoE is the overall performance of a system from the point of view of the users. QoE is a measure of end-to-end performance at the services level from the user perspective and an indication of how well the system meets the user's needs. QoS is a measure of performance at the packet level from the network perspective. QoS also refers to a set of technologies (QoS mechanisms) that enable the network administrator to manage the effects of congestion on application performance as well as providing differentiated service to selected network traffic flows or to selected users. In order to deliver acceptable service quality, QoE targets should be established for each service and be included early on in system design and engineering processes. QoE is essential for the successful deployment of triple-play and beyond services and will be a key differentiator with respect to competing service offerings. Subscribers to network services don't care how service quality is achieved. What matters to them is how well a service meets their goals and expectations - their Quality of Experience (QoE).

It is widely acceptable that QoS acted as an inspiration to many researchers who introduced mechanisms to exploit QoS definitions towards the design and pricing of telecommunication systems and services, especially in the core network [10]. But nowadays, "*Quality of Experience includes everything that really matters*" and interesting questions are arising related to the access/home bandwidth management and business, especially as the access and home networks continue to remain actual bottlenecks.

6.3 Methodology and Goals

The methodology of building a strategic framework for QoE within SESAME is based on the Analytical Hierarchy Process (AHP). More specifically, the concept called the Technology Development Envelope (TDE) is capable to transform the interesting questions arisen, within a technology R&I project, to a dynamic, flexible and operational level. This can be further transformed to requirements, design specifications and technology roadmapping.

The AHP (Analytic Hierarchy Process) is applied as a part of the TDE framework with four levels (objective-criteria-factors-alternatives). AHP is constructed to decompose the complex decision problems and incorporate quantitative and qualitative aspects into the evaluation process.

The proposed approach applies a semi-absolute scale to quantify the values of services and technologies in conjunction with the determination of criteria priorities and the relative importance of factors under each criterion.
The impact of technologies on a company's objective is calculated as a composite index called Technology Value [11]. In Table 3, the main outcomes from the approach adopted in SESAME are presented.

Table 3: Methodology

	SESAME
Methodology Goals	Service and technology prioritization and road-mapping
Methodology Stages	1. Service and technology characterization 2. Hierarchical modelling 3. Feedback gathering 4. Data Analysis 5. Service and Technology Priorities and Evaluation
Methodology Strategies	- Experts discussions about criteria and factors - Questionnaires - Summarizing and Discussion
People Involved	Consortium members and people working in related fields
Kind of Services Tested	Mobility and service performance, economic and business aspects, operation, administration and maintenance issues, social acceptance and flexibility

6.4 Model Development

The referred model for the evaluation of emerging technologies is achieved in three discrete phases [11]:

1. Technology Characterization
2. Hierarchical Modelling
3. Technology Evaluation

Figure 23, below, presents the information flow within the model through the three steps as well as the strategic information, which is derived from experts and users, as inputs to the model.
The strategic information denotes the list of metrics describing the performance and physical characteristics of the technology under investigation.

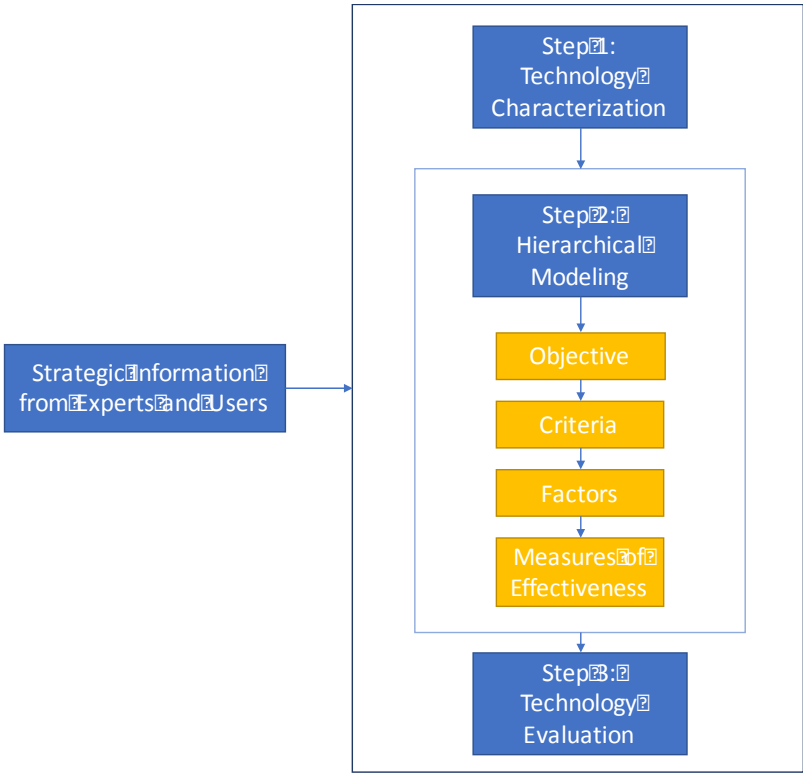

Figure 23: Information flow to and within the model

Strategic information on emerging technologies is obtained by applying Delphi (or interview, nominal group discussion) method[132] by gathering expert opinions.

The process of hierarchical decision modelling is followed to decompose the structure of complex issues into hierarchies and then apply the comparative judgments in order to synthesize the relative priorities of components in each hierarchy [12].

In each hierarchy, components and quantification of their relative priorities should be identified for the objective. For this purpose, a panel of some experts is involved in the implementation technologies.

[132] For further information about the Delphi method see, *for example*: https://en.wikipedia.org/wiki/Delphi_method.

6.4.1 Phase1: Technology and Service Characterization

Experts define and verify the objective for evaluating technologies. Then they decompose the decision complexity by determining criteria and technological factors, which contribute to the satisfaction of the objective.

A list of technological factors is identified under each area of criteria in order the contribution of technology to be directly measured. These factors could be either quantitative or qualitative parameters depending on the means which are used in measuring the contribution of technologies towards factors.

Identification of components placed in the criteria and technological factors level is accomplished based on the focus of their preferential independence even though some components can share their technical dependency.

6.4.2 Phase2: Hierarchical Modelling for the Evaluation of Emerging Technologies

A four-level hierarchical model is constructed according the following hierarchy, which is depicted below, as in Figure 24. This comprises of the following structural "modules":

1. Objective
2. Criteria
3. Factors
4. Alternatives

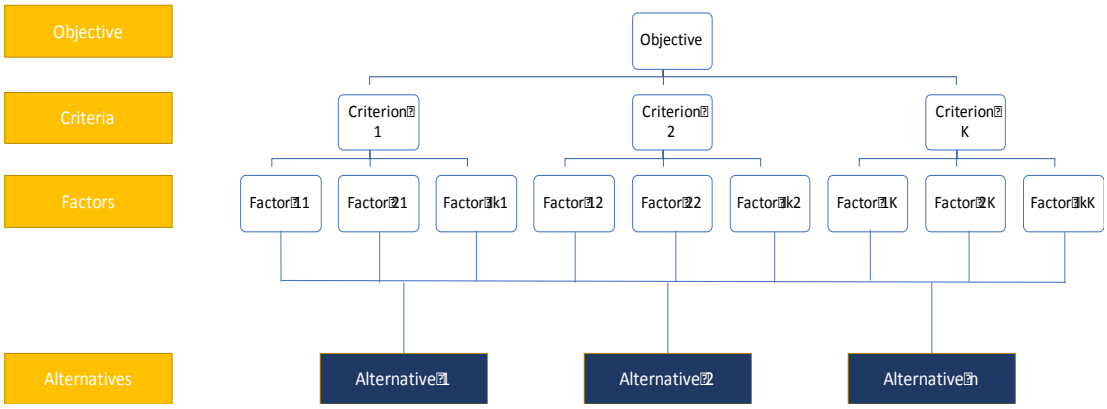

Figure 24: The generalized hierarchical model

According to this model, a hierarchical structure is developed, in which priorities of the criteria and the relative importance of factors on each criterion are determined. The relative values of components in each level are identified through a series of pairwise judgment qualification with respect to the elements in the next higher level.
In order to quantify the impact of a technology on the objective, a composite index called Technology Value is proposed.

For the new development, the model needs to be transformed to an operational model by adding an extra layer named measures of effectiveness (Figure 25).

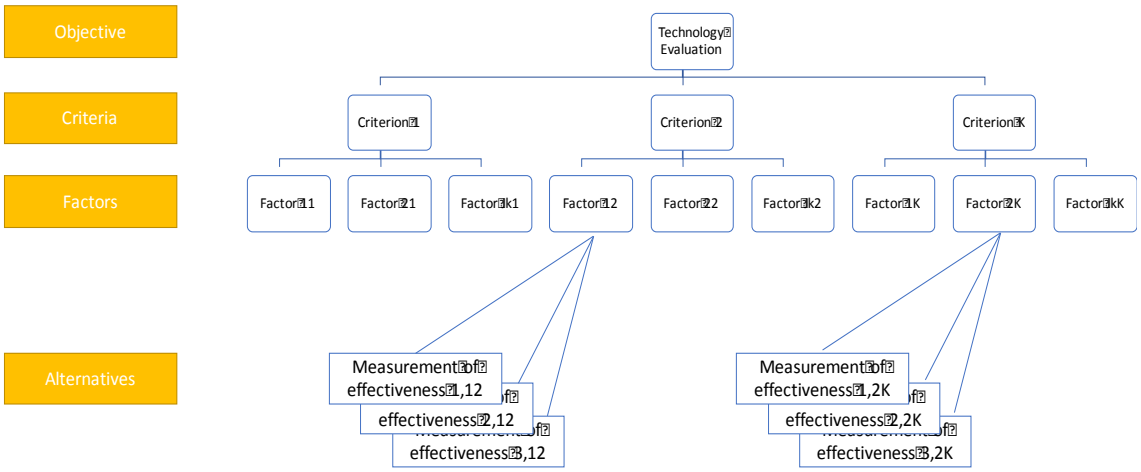

Figure 25: The operational hierarchical model developed

A set of measures of effectiveness (metrics) is defined for each technological factor so that the performance and physical characteristics of emerging technologies could be directly evaluated.
The impact relationships of metrics associated with each factor are determined through the quantification of judgments for the desirability of each measure of effectiveness.

6.4.3 Phase3: Technology Evaluation

The Technology Value of an emerging technology can be estimated based on the following equation [11]:

$$TV = \sum_{k=1}^{K} \sum_{j_k=1}^{J_k} w_k \cdot f_{j_k,k} \cdot V\left(t_{j_k,k}\right) \quad \text{(Eq.1)}$$

Where:

TV: Technology Value (QoE)

w_k : Relative Priority of Criterion (k) with respect to the objective.

$f_{j_k,k}$: Relative Importance of factor (jk) with respect to criterion (k).

$\sum_{j_k=1}^{J_k} w_k \cdot f_{j_k,k}$: Relative Importance of factor (jk) with respect to the objective.

$t_{j_k,k}$: Performance and physical characteristics

$V\left(t_{j_k,k}\right)$: Desirability Value of the performance and physical characteristics

The technology under study (SESAME solution) is evaluated by measuring how well the technological metrics meet users' (end-users or MNOs) desirability level and "weighting" that by the relative importance of factors as well as the relative priority of criteria.

For the above process two measurements need to be defined:

Measurement 1: For the calculation of relative priority of criterion k with respect to the objective, the below condition must be satisfied.

$$\sum_{k=1}^{K} w_k = 1 \ \ where \ w_k > 0 \quad \text{(Eq.2)}$$

Measurement 2: Estimation of relative importance of factor j_k requires the following condition:

$$\sum_{j_k=1}^{J_K} f_{j_k,k} = 1 \; for \; each \; criterion \; (k), where \; f_{j_k,k} > 0 \qquad \text{(Eq.3)}$$

6.4.4 Phase4: Formation of Technology Development Envelope (TDE)

According to TV estimated above a Technology Development Envelope (TDE) was formed.

6.5 Discrete Steps during the Survey

According to the above, the discrete steps of the AHP model procession will be presented below.

Step1: Estimation of relative priority of criterion k (w_k)

Experts are asked to allocate a point between 0-100 at each criterion with respect to the objective.
Then a series of comparative judgments are obtained from each expert on each pair of criteria and the judgments are converted to a normalized measure of priority in ration scale for the criteria through the method of "Pairwise Comparison".
All the weights (w) obtained from pairwise comparison for each expert are combined and the mean value is calculated to represent the group relative priority (w_k) for every criterion [13].

Step2: Determination of relative impact of factors j_k under each criterion k ($f_{jk,k}$)

Experts are asked to assign a point of importance between 0 – 100 to each factor with respect to the appropriate criterion every time.
A series of comparative pairwise judgments on technological factors follow and the relative importance of each factor associated with each criterion is estimated according to the approach described in step1.

Step3: Determination of the relative desirability ($V(j_k,k)$) of measures of effectiveness under each factor j_k and each criterion k.

The relative desirability of metrics [11] associated with each factor j_k and criterion k is identified through the steps 1-4, below:

1. Determination of the best and worst desirable limiting metrics for each factor (factor limits and nominal values).

2. Identification of metrics that their desirability value is linearly proportional to their arithmetic value between the above limits.

3. Creation of a semi-absolute scale assigning 0 to the worst and 100 to the best desirable limiting measures of effectiveness for each factor.

4. Estimation of relative desirability of the rest intermediate values based on one of the two ways that are described below:
   a. If metrics of a factor can be described as a linearly proportional function, then their desirability value can be developed linearly proportional to their arithmetic value between the above limits.
   b. If measure of effectiveness of a factor cannot be verified as a linearly proportional function, then the function of their desirability value needs to be developed.

$0 \le V(j_k,k) \le 100$, for each factor j_k and criterion k.

Finally, the relative desirability of measures of effectiveness for each factor is graphically depicted as a curve assigning the range of metrics value on x axis and the desirability value on y axis.

Step4: Mapping of technological metrics ($t_{jk,k}$) to desirability values ($V(t_{jk,k})$)

Experts are asked to assign values to each factor. These values are called technological metrics.

Then, technological metrics ($t_{jk,k}$) are being mapped under each factor to the desirability values (V($t_{jk,k}$)) through the relative desirability of metrics V(j_k,k) [11].

Step5: Estimation of Technology Value (TV)

The technology value TV is calculated according to the equation (1).
More specifically, matrix computations among the criteria priorities (step1), the relative importance of each factor on each criterion (step 2) and the desirability values of technological factors (step 4) can be calculated and have as an outcome the technology values of emerging technologies.

Due to the fact that this happens according the objective, it should be referred that the best technology could represent the technology value of 100.

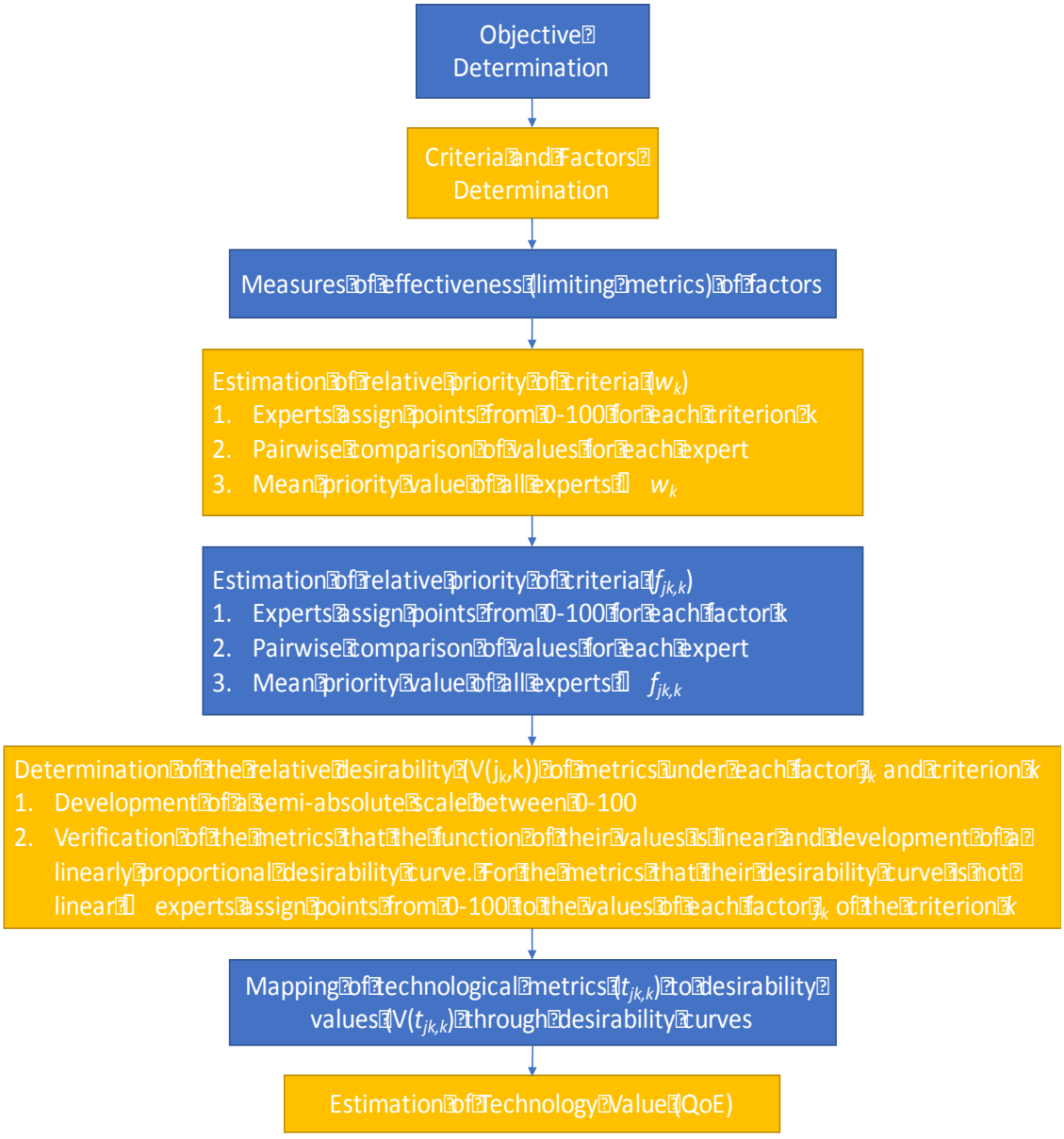

Figure 26: Flow diagram for the evaluation of Technology Value (QoE) using the AHP methodology

6.6 Design of SESAME Survey for QoE

In order to apply the above described methodology, a set of criteria and factors, under each criterion, have been defined as follows:

Criterion 1: System performance
- Coverage;
- Quality of the signal;
- Mobility;
- Handovers.

Criterion 2: Service Performance
- Achieved Bit Rate (from user perspective this is linked to the native quality of the service);
- Quality degradation of picture (image freeze/ pixellisation - i.e. cause by Packet loss rate);
- Quality degradation of sound (clicks/cut of sound - i.e. cause by Packet loss rate);
- Quality degradation of voice;
- Perceived delay;
- Adaptability to different conditions (Examines if the system is adaptable to environmental conditions.

Criterion 3: Flexibility
- Ease of use (The system is easy to be used. For example, it could offer auto detection and auto-installation);
- Ease of network deployment;
- Multitenancy;
- Ease of service chaining.

Criterion 4: Reliability
- Continuous operation;
- Durability under adverse conditions;
- Performance drop overtime.

A 5-point Likert or 6-point scale can be used to form the necessary questions in order to evaluate the technological metrics. For example:

Q1: How do you characterize the durability under adverse conditions of the system?

Durability under adverse conditions	Unacceptable	Poor	Acceptable	Good	Very Good	Excellent

6.7 Bibliographic References of Annex A

[1] ITU-T (1994, August): *ITU–T Recommendation E.800:* "*Terms and definitions related to quality of service and network performance including dependability*". ITU-T, Geneva.

[2] ITU-T (2001, November): ITU-T Recommendation G.1000: "*Communications Quality of Service: A framework and definitions*". ITU-T, Geneva.

[3] K. Kilkki (1999): Differentiated Services for the Internet. Macmillan Publishing.

[4] ETSI (2006): *ETSI TR 102 479 V1.1.1:* "*Review of available material on QoS requirements of Multimedia Services*".

[5] C. Courcoubetis and R. Weber (2003): *Pricing Communication Networks: Economics, Technology and Modelling.* Wiley.

[6] D. Soldani, M. Li, R. Cuny (2006): *QoS and QoE Management in UMTS Cellular Systems*. Wiley.

[7] D. Lopez, F. Gonzalez, L. Bellido, A. Alonso (2006, June): *Adaptive multimedia streaming over IP based on customer oriented metrics*. In IEEE Proceedings of the 2006 International Symposium of Computer Networks. IEEE, Istanbul, Turkey, June 16-18, 2006.

[8] ITU-T Special Group 12 on *Performance and Quality of Service (QoS)*. Available at *http://www.itu.int/ITU-T/studygroups/com12/index.asp*

[9] DSL Forum WT-126 v.0.5 (2006, February): "*Triple-play Services Quality of Experience (QoE) Requirements and Mechanisms for Architecture & Transport*".

[10] P. Reichl (2007, September): "From "Quality-of-Service" and "Quality-of-Design" to "Quality-of-Experience"": A Holistic View on Future Interactive Telecommunication Services, (Invited Paper), IEEE SoftCOM'07. IEEE, Split, Croatia, September 27-29, 2007.

[11] N. Gerdsri (2016): Strategic Planning: A Quantitative Model for the Strategic Evaluation of Emerging Technologies. In: Daim T. (eds.) *Hierarchical Decision Modeling. Innovation, Technology, and Knowledge Management*. Springer, Cham.

[12] M. Yurdaku (2004): AHP as a strategic decision-making tool to justify machine tool selection. *Journal of Materials Processing Technology 146*(2004), pp.365-376.

[13] T.L. Saaty (2003, February): Decision-making with the AHP: Why is the principal eigenvector necessary? *European Journal of Operational Research 145*(1), pp.85-91. Elsevier.