5Genesis

5TH GENERATION END-TO-END NETWORK, EXPERIMENTATION, SYSTEM INTEGRATION, AND SHOWCASING

[H2020 - Grant Agreement No. 815178]

Deliverable D2.3

Initial planning of tests and experimentation

Editor A. Díaz Zayas (UMA)

Contributors A. Díaz Zayas, B. García, P. Merino (UMA), A. Brunstrom (KAU), G. Caso (SRL), M. Emmelmann, R. Shrestha, Mikhail Smirnov (FhG), A. M. C Bosneag, E. Aumayr (LMI), J.O. Fajardo, E. Atxutegi (NEM), V. Koumaras, G. Theodoropoulos (INF), I. Pretel, A. Pineda (FON), Jesús Gutiérrez (IHP), D. Tsolkas (FOG), G. Xilouris, H. Koumaras, T. Anagnostopoulos (NCSRD), G. Gardikis (SHC).

Version 1.0

Date February 18th, 2019

Distribution PUBLIC (PU)

List of Authors

UMA	UNIVERSIDAD DE MALAGA
A. Díaz Zayas, B. García, P. Merino	
KAU	KARLSTADS UNIVERSITET
A. Brunstrom	
SRL	Simula Research Laboratory AS
G. Caso	
FhG	FRAUNHOFER GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V.
M. Emmelmann, Mikhail Smirnov, R. Shrestha	
LMI	L.M. ERICSSON LIMITED
A. M. C. Bosneag, E. Aumayr	
NEM	NEMERGENT
J.O. Fajardo, E. Atxutegi	
ATOS	ATOS SPAIN SA
Javier Melian	
INF	INFOLYSIS P.C.
V. Koumaras, G. Theodoropoulos	
FON	Fon Technology S.L.
I.Pretel, A. Pineda	
NCSRD	National Center for Scientific Research “Demokritos”
G. Xilouris, H. Koumaras, T. Anagnostopoulos	
IHP	IHP GMBH – INNOVATIONS FOR HIGH PERFORMANCE MICROELECTRONICS/LEIBNIZ-INSTITUT FUER INNOVATIVE MIKROELEKTRONIK
Jesús Gutiérrez	
FOG	Fogus Innovations & Services
D. Tsolkas	
SHC	Space Hellas Cyprus
G. Gardikis, N. Papadakis	

Disclaimer

The information, documentation and figures available in this deliverable are written by the 5GENESIS Consortium partners under EC co-financing (project H2020-ICT-815178) and do not necessarily reflect the view of the European Commission.

The information in this document is provided "as is", and no guarantee or warranty is given that the information is fit for any particular purpose. The reader uses the information at his/her sole risk and liability.

Copyright

The 5GENESIS Consortium consists of:

NATIONAL CENTER FOR SCIENTIFIC RESEARCH “DEMOKRITOS”	Greece
Airbus DS SLC	France
ATHONET SRL	Italy
ATOS SPAIN SA	Spain
AVANTI HYLAS 2 CYPRUS LIMITED	Cyprus
AYUNTAMIENTO DE MALAGA	Spain
COSMOTE KINITES TILEPIKOINONIES AE	Greece
EURECOM	France
Fogus Innovations & Services P.C.	Greece
FON TECHNOLOGY SL	Spain
FRAUNHOFER GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V.	Germany
IHP GMBH – INNOVATIONS FOR HIGH PERFORMANCE MICROELECTRONICS/LEIBNIZ-INSTITUT FUER INNOVATIVE MIKROELEKTRONIK	Germany
INFOLYSIS P.C.	Greece
INSTITUTO DE TELECOMUNICACOES	Portugal
INTEL DEUTSCHLAND GMBH	Germany
KARLSTADS UNIVERSITET	Sweden
L.M. ERICSSON LIMITED	Ireland
MARAN (UK) LIMITED	UK
MUNICIPALITY OF EGALEO	Greece
NEMERGENT SOLUTIONS S.L.	Spain
ONEACCESS	France
PRIMETEL PLC	Cyprus
RUNEL NGMT LTD	Israel
SIMULA RESEARCH LABORATORY AS	Norway
SPACE HELLAS (CYPRUS) LTD	Cyprus
TELEFONICA INVESTIGACION Y DESARROLLO SA	Spain
UNIVERSIDAD DE MALAGA	Spain
UNIVERSITAT POLITECNICA DE VALENCIA	Spain
UNIVERSITY OF SURREY	UK

Version History

Rev. N	Description	Author	Date
1.0	Full version	all	18/02/2019

LIST OF ACRONYMS

Acronym	Meaning
3GPP	Third Generation Partnership Project
5G PPP	5G Infrastructure Public Private Partnership
API	Application programming interface
CPU	Central Processing Unit
CQI	Channel Quality Indicator
C-RAN	Cloud-RAN
CSI	Channel State Information
DUT	Device Under Test
E2E	End To End
EaaS	Experimentation as a Service
EARFCN	Evolved-UTRA Absolute Radio Frequency Number
eMBB	Enhanced Mobile Broadband-5G Generic Service
eNB	eNodeB, evolved NodeB, LTE eq. of base station
ELCM	Experiment Lifecycle Manager
EU	European Union
EPC	Evolved Packet Core
ETL	Extract, Transform, and Load
ETSI	European Telecommunications Standards Institute
EUTRAN	Evolved Universal Terrestrial Access network
FDD	Frequency Division Duplexing
GPS	Global Positioning System
ICCID	Integrated Circuit Card Identifier
ICMP	Internet Control Message protocol
IMEI	International Mobile Station Equipment Identity
IMSI	International Mobile Subscriber Identity
IP	Internet Protocol
IOT	Internet of Things
KPI	Key Performance Indicator
LAC	Location Area Code
LTE	Long-Term Evolution
LTE-A	Long-Term Evolution - Advanced
MAC	Medium Access Control
MANO	NFV MANagement and Organisation
MCC	Mobile Country Code
MCS	Mission Critical Services
MCSI	Modulation and Coding Scheme Index
MEC	Mobile Edge Computing
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity
mMTC	Massive Machine Type Communications-5G Generic Service
MNC	Mobile Network Code

MOCN	Multiple Operator Core Network
MONROE	Measuring Mobile Broadband Networks in Europe.
NFV	Network Function Virtualisation
NGMN	Next generation mobile networks
NMS	Network Managment System
OFDM	Orthogonal Frequency Division Multiplexing
PoC	Proof of concept
PCRF	Policy and Charging Rules Function
PDCP	Packet Data Convergence Protocol (PDCP)
PDSCH	Physical Downlink Shared Channel
PoP	Point of Presence
P-GW	Packet Data Node Gateway
PLMN	Public Land Mobile Network
PMI	Precoding Matrix Indicator
PNF	Physical Network Functions
PRB	Physical Resource Block
RAN	Radio Access Network
REST	Representational State Transfer
RSCP	Received Signal Code Power
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSSI	Received Signal Strength Indicator
RTT	Round trip time
SCPI	Standard Commands for Programmable Instruments
SIM	Subscriber Identity Module
SIMO	Single input, multiple output
UDP	User datagram Protocol
UE	User Equipment
uRLLC	Ultra-Reliable, Low-Latency Communications
YAML	YAML Ain't Markup Language (human readable data serialisation language)

Executive Summary

This document is the third technical deliverable of the 5GENESIS project, and contains work conducted in the WP2 of the project, complementing the content of deliverables D2.1 and D2.2. Its major objective is to describe a common end-to-end testing and experimentation methodology that will be used to assess the targeted 5G Key Performance Indicators (KPIs) over the 5GENESIS platforms and define the schedule for the integration of this experimentation methodology in each one of the platforms. In that sense, this document serves as a reference point for WP3 activities, where the key components of this experimentation methodology will be implemented, and for WP6 activities, where the targeted 5G KPIs will be evaluated.

More precisely, the 5GENESIS experimentation methodology, introduced in this report, is expected to ensure the validity of the results, the repeatability of the experiments and the possibility of comparing the results between platforms. The lessons that will be learned from the implementation and testing to be conducted based on this methodology, will contribute to the appropriate refinement of the document and the release of the final experimentation methodology of the project (D2.4) that will drive the last evaluation cycle of the project. In parallel, the current document will be used as the 5GENESIS initial input to the 5G PPP Test, Measurements, and KPIs Validation Working Group (TMV WG).

To accomplish the goal of this document, the content has been organized as follows.

The introduction clarifies key concepts of the 5GENESIS objectives and architecture to facilitate the understanding of the approach followed for the definition of the experimentation methodology. Section 2 serves as reference section to assist the reader to familiarize with the terms associated with the 5GENESIS experimentation methodology and describes the experimentation workflows, regarding the 5GENESIS reference architecture. Section 3 specifies the 5GENESIS initial choices for the information that need to flow in the platform (Portal inputs, Experiment descriptor, Test case template) in order to set-up an experiment. Section 4 describes the main frameworks that are selected as the basis for the implementation of the experimentation methodology in 5GENESIS. Finally, Section 5 provides the initial planning for the realization of the experimentation methodology in the platforms that define the 5GENESIS facility.

This deliverable provides the initial specifications of the following key concepts in the 5GENESIS experimentation methodology:

- the experimentation scenarios, which details the end-to-end conditions for running the experiments, such as the mobility and the location of the User Equipment (UE).
- the experiment descriptor, which contains all the information required by the platforms to run the experiments.
- Slice configurations, which detail the end-to-end resources allocated for the execution of the experiments.
- the test cases, which define target of the experiment, the procedure and the measurements that have to be collected in order to validate the targeted KPIs.

Table of Contents

1. INTRODUCTION

5GENESIS project is building a distributed facility to support end-to-end experimentation over 5G networks in a controlled environment. The project has already identified the requirements of this facility (deliverable D2.1) and has also defined a reference architecture to be instantiated in all the platforms that compose the facility (deliverable D2.2).

A key aspect prior the instantiation of this architecture is the detailed design of the workflow needed to run the experiments in the platforms and in the global facility. Focussing on this specific aspect, the **5GENESIS experimentation methodology** is presented in this document. The methodology has been designed to facilitate the execution of a series of tests and to allow for the computation and validation of the Key Performance Indicators (KPIs) defined in the project.

As introduced in deliverable D2.1, 5GENESIS faces the challenge to validate the 5G network KPIs and verify the 5G technologies with an end-to-end approach as part of the on-going 5G PPP Phase 2 and 3 projects. Adopting the early releases of the 5G Architecture [26], 5GENESIS has proposed a 5G experimentation blueprint to serve as a common reference implementation architecture, including an openness framework, with APIs for exposing the 5GENESIS Facility to verticals for experimentation. Additionally, and in line with the project's Objective (Qualitatively assess and quantitatively validate business, performance, and societal 5G PPP KPIs in representative 5G use cases) a set of use cases with relevant KPIs have been identified to be used as the drivers in the development of the platforms (deliverable D2.1). These use cases will be the source of the first wave of experiments to be supported by the 5GENESIS consortium, however the facility is open to further experiments from vertical industries, and in particular to those coming from the ICT-19 project.

Considering the 5GENESIS reference architecture (Figure 1), the Platform Coordinator layer is responsible for the coordination of the platform, achieving overall supervision and E2E configuration for service deployment and management/monitoring. At this layer the interfaces towards the experimenter are exposed, the Experiment Lifecycle is managed, as well as security aspects, and analytics to fulfil the goals of KPI monitoring & validation are implemented. Moreover, this layer is responsible for exposing East/West interfaces to another Coordinator in other 5GENESIS platforms, facilitating interworking. Thus, the coordination layer carries the functionalities needed for defining an experiment and apply it to the other layers.

In that sense, the aim of this document is to define a methodology that expands the requirements identified in D2.1, fulfills the requirement COORD-5 (Table 1) and specifies the role of the architectural components in the experimentation process.

Table 1. COORD-5 requirement

COORD-5	Experiment execution
Priority	Essential
Description	The Coordination layer shall be able to execute an experiment experiment's workflow or control the operation of the experiment over the Facility infrastructure and collect appropriate results

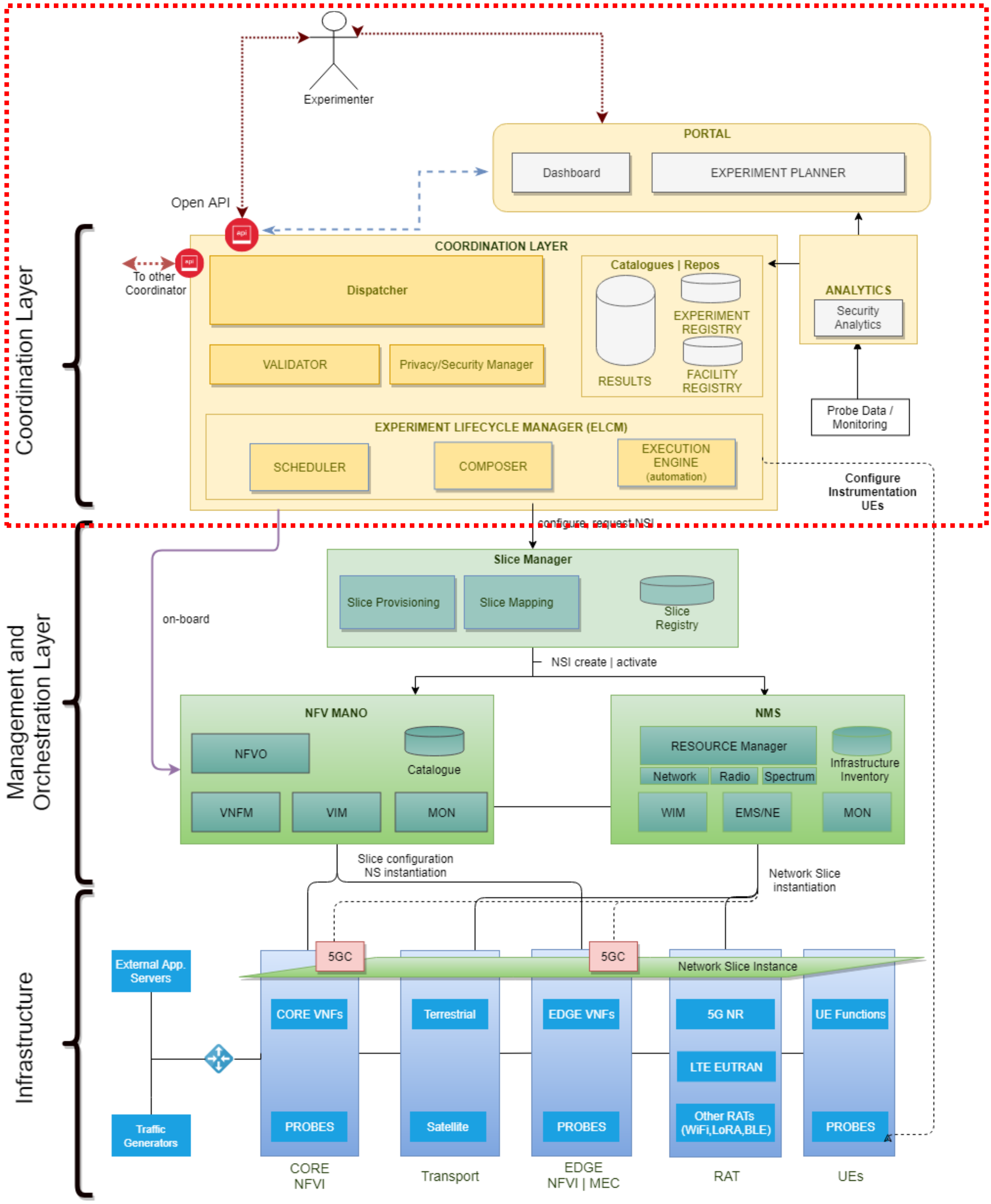

Figure 1: The coordination layer of the 5GENESIS reference architecture, reflecting the main architectural component that is involved in setting-up and applying for execution an experiment.

2. EXPERIMENTATION BASED ON TEST CASES

2.1. 5GENESIS Experimentation approach and alignment with 5GPP TMV WG

In order to ensure experiment repeatability, regardless of the test equipment and the entity performing the certification, the 5GENESIS experimentation methodology is based on the profiling of experiments including specific configurations and conditions. The latter in 5GENESIS terminology is referred to as a test case. Each test case apart from the description of the experiment and its measurable objectives includes a configuration file that specifies: i) the configuration of the environment (for all RATs and network); ii) the set of procedures; iii) the monitored metrics, and iv) formulas needed to calculate the KPIs. A test case provides uniformity and organization to run an experiment in a programmatic and structured way.

The software tool and the guidelines/rules for creating and executing test cases define a testing framework, which allows for an **automated testing process**. Hence, the testing framework provides (1) a well-defined format to specify inputs/test conditions and expected outputs, (2) an interface to plug in a scenario or experiment that is to be executed, (3) the actual test execution environment, and (4) a mechanism to report results.

In 5GENESIS a set of test cases will be used to define the data that needs to be measured to compute the KPIs and the procedure to run the measurements during the execution of the experiment. The template that includes all the information required for defining 5GENESIS test is presented in section 4.

The validation of 5G KPIs also implies testing under different network conditions, operation modes and so on, given that this heterogeneity can be easily found in 5G deployments. The large number of test conditions requires the definition of "**scenarios**" to ensure the repeatability and the coverage of all the relevant conditions that can impact the performance results of the experiment. The scenarios define the conditions of the experiments (signal strength, interference, UE mobility, etc.). This specification of the scenarios will be initiated in Task 3.3 and included in D3.5 and D3.6 deliverables.

Finally, the testing methodology takes also into account the 5G concept of the end-to-end slice. While the scenarios define the conditions of the experiment, the end-to-end slices define the specific resources allocated in each component of the end-to-end network for running an experiment. The definition of these slices will be started in tasks 3.2 and included in D3.3 and D3.4 deliverables.

To summarize, a 5GENESIS experiment is the execution of one of more test cases in different scenarios and using different slicing configurations. Thus, each of the test cases considered in the experiment will be executed for all possible combinations of the scenarios and the slicing configurations declared in the experiment.

The 5GENESIS experimentation methodology introduced in this document is oriented to the testing of KPIs in the context of the 5GENESIS, but the methodology proposed is flexible enough to be extended and cover new KPIs, new scenarios and new slicing configurations. Moreover, the 5GENESIS will participate actively in the TMV WG to reach a common approach and ensure the sustainability of the experimentation methodology promoted by 5GENESIS project.

In particular, the experimentation methodology proposed by 5GENESIS is aligned with the TMV WP approach. The 5GENESIS methodology is based on the definition of test cases, which is also a recurrent concept promoted by the TMV WG. In fact, the test case concept has been also used in the TRIANGLE project to define the testing methodology of applications and devices. UMA, the leader of T2.3, is one of the partners involved in the TRIANGLE project and has participated actively in the definition, implementation and execution of the test cases for applications and devices testing.

In the 5GENESIS project, and as well as in the TMV WG, the main target of the test cases is to test and validate 5G KPIs. 5GENESIS test cases will define the KPIs to be measured, the procedure to measure and compute them and the output provided after running the test case.

The measure and computation of the KPI is, again, a main activity that will covered in W3 and WP6.

Other important parts of the 5GENESIS experimentation methodology are the scenario and the slicing configurations. Slicing is also a key concept is 5G networks. The testing of slicing configurations is implicit in the test cases. The executions of the test cases under different slicing configurations will enable to characterize their performance based on the KPIs values obtained. The slicing configurations based on VNF and additional network configurations will be defined in Task 3.2 and will cover the requirements of the 5GENESIS use cases. Finally, we have the scenarios, the scenarios represent specific operation conditions of the network and of the terminal. Thus is, vehicular conditions, multi-cell deployment, indoor/outdoor deployment, … etc. These scenarios will capture the conditions were 5GENESIS uses cases are usually operated. These key concepts, test cases, scenarios, slicing topologies, have been also identified in the preliminary discussions of the TMV WG.

2.2. Defining an experiment

An experiment is a repeatable scientific procedure for testing a hypothesis. Experiments rely on the observation of change in a system's output while its inputs are manipulated. This way, the experimenter can investigate hypotheses about the cause-and-effect relationships between input and output. An experimental setup must be repeatable so that the validity of an experiment and its conclusions can be proven. In a typical experimental scenario, an experimenter might be interested in conditions that can produce a certain result, or in different results that can be achieved under given conditions.

Two different types of experiments are considered in the 5GENESIS project. In order to ensure measurement repeatability and KPI's comparability, defined are the so-called "standard experiments". This type of experiments specifies also the traffic sources to be used during an experiment. For these experiments, the 5GENESIS platforms will reproduce the same conditions, the same traffic, collect the same measurements and follow the same gathering procedure. Standard experiments are to check and to compare the results obtained in the same platform, and also between platforms.

In addition to the execution of standard experiments, the 5GENESIS platforms will also support the execution of "custom experiments"; in which the experimenters, e.g. verticals, have more degrees of freedom in configuring the services and the applications under test.

Custom experiments can be attended or unattended. The attended experiments include interaction with the user during the experimental flow. In such cases, the Experiment Life Cycle Manager (ELCM) configures the underlying components and provides periodic feedback about the information collected by the monitoring probes, intermediate KPIs computed by the analytic module and the status of the underlying components of the testbed. On the other hand, the unattended experiments do not require any user interaction during the experimental flow and hence provide a more automated experiment procedure. The Experiment Life Cycle Manager (ELCM) will configure traffic sources at both ends of a connection, or, in the case of experiments with real applications and services, the ELCM will interact with them to control its execution.

This section is devoted to the description of the experimentation workflow defined on top of the 5GENESIS architecture. Experimenters will use the 5GENESIS Portal to set up and execute an experiment since, as described in the previous section, the 5GENESIS Portal provides a set of configurations that guide the experimenters toward the definition of their experiment.

Once an experiment has been defined through the 5GENESIS Portal, its life cycle is composed of five stages:

- Check of available resources
- Configuration of the resource
- Running and collection of measurements, as well as computation of intermediate KPIs
- Computation of final KPIs
- Termination

The ELCM will supervise an experiment through these stages. This section will introduce how the Experiment Life Cycle runs and how it interacts with the rest of the components of the Coordination and underlying layers of the 5GENESIS architecture.

The workflow in Figure 2 shows how the different elements in the platform interact during the testing and validation of the KPIs.

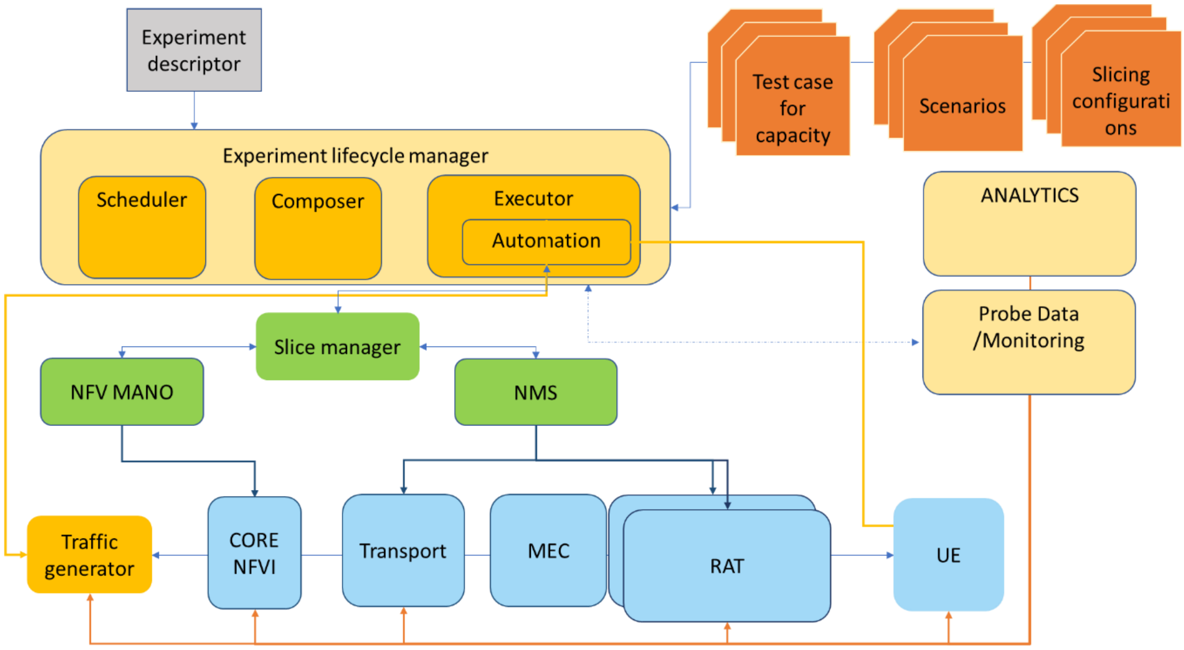

Figure 2 Experimentation workflow

The ELCM will compose an experiment plan using the experiment descriptor and the information of the test cases, the scenarios and the slicing configurations. The experiment plan includes the network scenario setup, the slices to be deployed, the measurements that have to be gathered in order to compute a given set of KPIs and, potentially, the servers or traffic sources. The experiment plan is built from existing templates in the platforms that enable the configuration of all components of the testbed. These templates are platform-specific because they depend on the components available in each platform.

This experiment plan is then executed. The slice manager communicates with the NFV MANO to deploy the network services in the virtual infrastructure. On the other hand, the Slice Manager will communicate with the NMS to configure the transport and the physical components. In case of an unattended test, the ELCM will execute the body of each test, acting directly onto the Servers/Applications/Traffic generators and UE and will gather results produced by measurement probes and tools.

The values obtained for the KPI(s) under test will be compared with (a set of) reference values, thus triggering the KPI(s) validation. To rate the performance, the KPI values will be compared against a set of reference values. Finally, the results will be presented to the experimenter in the Portal.

2.3. Sequence diagram of an experiment life cycle

Figure 3 shows the sequence of interactions between the different components of a 5GENESIS platform during the execution of a single experiment. The experimenter defines the experiment using the portal (the full list of steps representing the interaction between experimenter and the 5GENESIS Portal are omitted in the figure) and requests its execution. The Dispatcher obtains the experiment descriptor from the Portal, initiates the validation of the descriptor and sends the experimentation plan to the scheduler that enqueues the execution until all necessary resources are available.

Once the Management and Orchestration Layer confirms that the required resources are available then the execution of the experiment starts. The Scheduler creates an independent Executor for the experiment, which will handle the communication with the Monitor and Analytics system and the Management Layer.

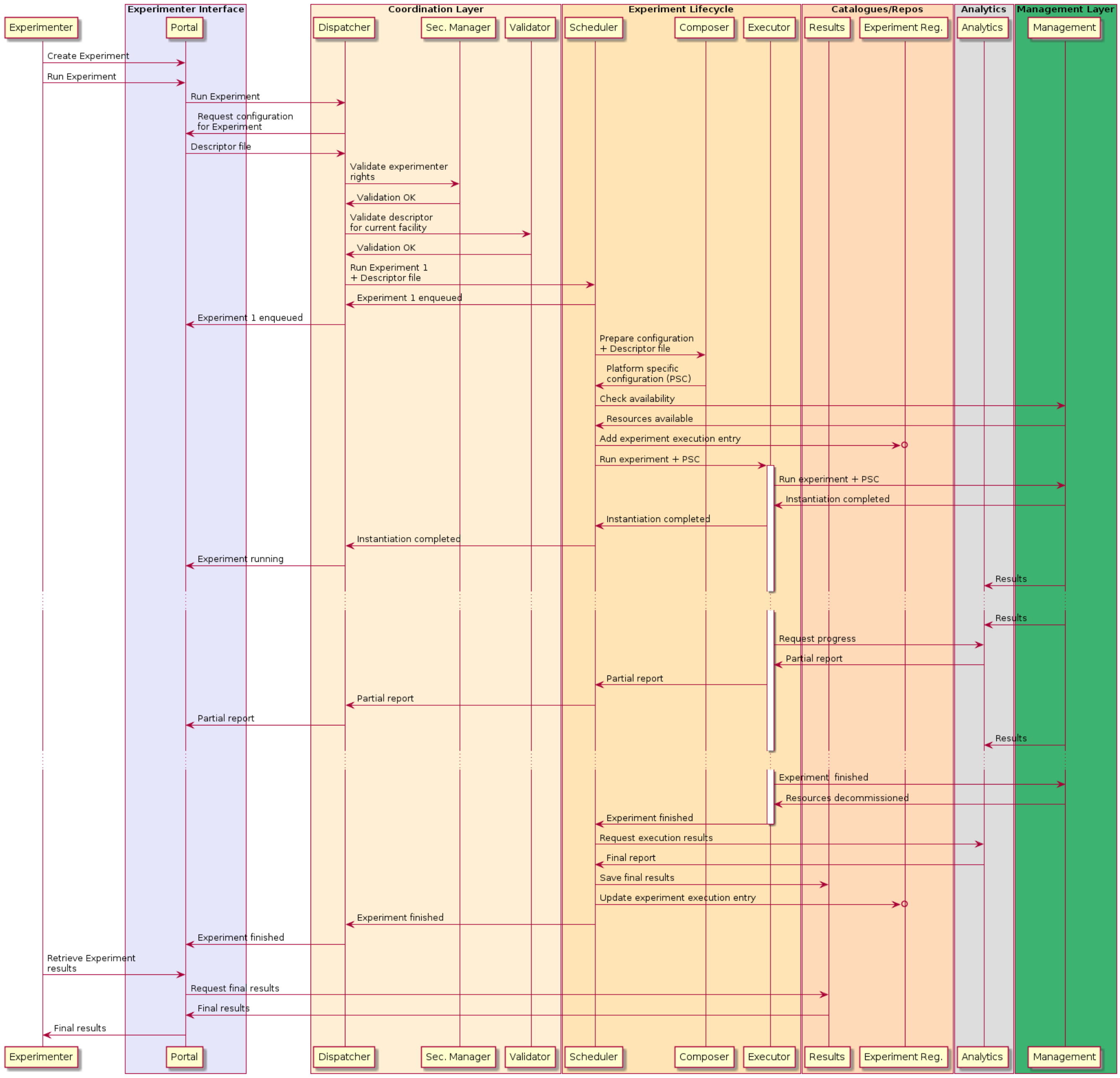

Figure 3 Sequence diagram of an experiment life cycle

When the experiment ends and all resources have been decommissioned the Scheduler initiates the computation of the final KPI's of the experiment using the Monitoring and Analytics components and saves all the generated data on the Results repository. Then, the Scheduler notifies the Dispatcher that the execution has finished.

Below there is a short description of the different components of the coordination layer of the 5GENESIS architecture introduced in D2.2 and the functions they perform.

- The **Portal** is the main interface toward the experimenter. It allows the creation and execution of experiments, the visualization of results and the on-boarding of new resources. The Portal communicates with the Coordination Layer (especially with the Dispatcher) using the 5GENESIS Open API.
- The **Dispatcher** is the entry point of the Facility. It can accept execution requests from the Portal or directly fetch an Experiment Descriptor. The Dispatcher initiates the validation of the Descriptor files and communicates with the Security Manager. When the experiment descriptor has been validated it offloads the execution to the Experiment Life Cycle Manager, while acting as the communication link between the Scheduler (ELCM component) and the Portal.
  The Dispatcher is also able to send part of an experiment descriptor to a Dispatcher on another 5GENESIS Platform for distributed execution of experiments.
- The **Security Manager** proves that the Experimenter has the required rights for the experiment execution and creates entries for each execution attempt for security logging.
- The **Validator** is responsible for validating the uploaded descriptors before they are saved on their corresponding repositories or used in the platform. During an experiment preparation, the Validator confirms that the experiment can be executed on the current Platform.
- The **Scheduler** orchestrates the execution of the different experiments. It is able to control the execution of multiple experiments at the same time, provided that there are enough resources on the Management Layer, or keep the experiments on an internal queue until these resources become available.
  The Scheduler also initiates the creation of the Platform specific configuration files using the Composer, and requests the post-processing of the experiment results and the creation of the final KPIs on the Analytics system once an experiment execution has been finalized. The execution of the experiment is off-loaded to one Executor (the platform can support the execution of concurrent experiments; in this case, there are different instances of the Executor, one per experiment in execution).
- The **Composer** is able to generate the Platform specific configuration files required for the experiment, by using the information contained on the Facility Registry (not shown in Figure 3). The Facility Registry contains information about the resources available on the Platform.
- The **Executor** (one for each running experiment) handles the creation and decommission of the required experiment resources on the Management and Orchestration Layer and communicates the intermediate results of the experiment to the upper layers. If **Automation** is enabled, it will also handle the execution of different actions on the servers deployed at the infrastructure layer and/or on the user equipment.

- The **Results** and **Experiment Registry** contains information about previously executed experiments. The Experiment Registry contains entries for each experiment execution. These entries include information such as the start and end time of the execution, the global result (successful or failed) and other metadata. The Results repository contains all the data generated from the experiment, including the raw results and the obtained computed KPIs.
- The **Analytics** will take as input the measurements from the both application monitoring probes (e.g. MONROE probes), instantiated at different vantage points in the platform (UEs, access network, servers), and network monitoring tools (e.g. tools to monitor the status of physical/virtual platform components, e.g. Ceilometer (if the platform is based on OpenStack), or others). The measurements will serve for the computation of KPIs, as well as for the analysis of the overall status of the platform during the experiment.

2.4. Concurrent execution of experiments

Figure 4 shows an example of the sequence of actions performed by the different components on a 5GENESIS platform when multiple experiments are executed concurrently. For the sake of simplicity, the validation of the requests and the handling of results have been omitted (can be seen in Figure 3).

In this example, one or multiple users try to run 3 different experiments on the same platform. The first experiment (shown in red arrows) is executed immediately (the platform is idle at the start and the Management Layer reports that all resources are available). Once the second experiment (shown in blue arrows) arrives, the Scheduler asks the Management Layer if it is possible to run it. Also, in this case the Management Layer reports that there are enough free resources and the execution can start.

The third experiment (shown in green arrows) requires resources that are being used by Experiment 1. Because of this, the Management Layer reports that it is not possible to run the experiment. The Scheduler keeps Experiment 3 in an internal queue, waiting for Experiment 1 to end.

Once Executor 1 reports that its experiment has been finished, the Scheduler asks again if it is possible to run Experiment 3 to the Management Layer. At this point all the resources from Experiment 1 have been decommissioned, and the execution of Experiment 3 can start.

2.5. Experiments with human interaction

If an experiment involves real users, the ELCM configures all the underlying components and provides periodic feedback about the information collected by the monitoring probes, intermediate KPI's computed by the Analytics module and the status of the underlying components of the platform. During the experiment execution time, the experimenter will be in charge of executing the application under test.

2.6. Unattended experiments

In the case of unattended experiments, the ELCM will also configure traffic sources at both edges of the connection or, in the case of experiments with real apps, the ELCM will interact with the servers deployed at the core or MEC servers and the apps running in the UE.

2.7. Measurements collection for the computation of KPIs

The ELCM will periodically ask the Analytics system about the intermediate values of computed KPI's and will deliver these values to the 5GENESIS Portal to show these values to the experimenter. At the end of the experiment, the ELCM will retrieve the final KPIs and will send them again to the 5GENESIS Portal. All these values (intermediate and final) will be also stored in the 5GENESIS Results repository.

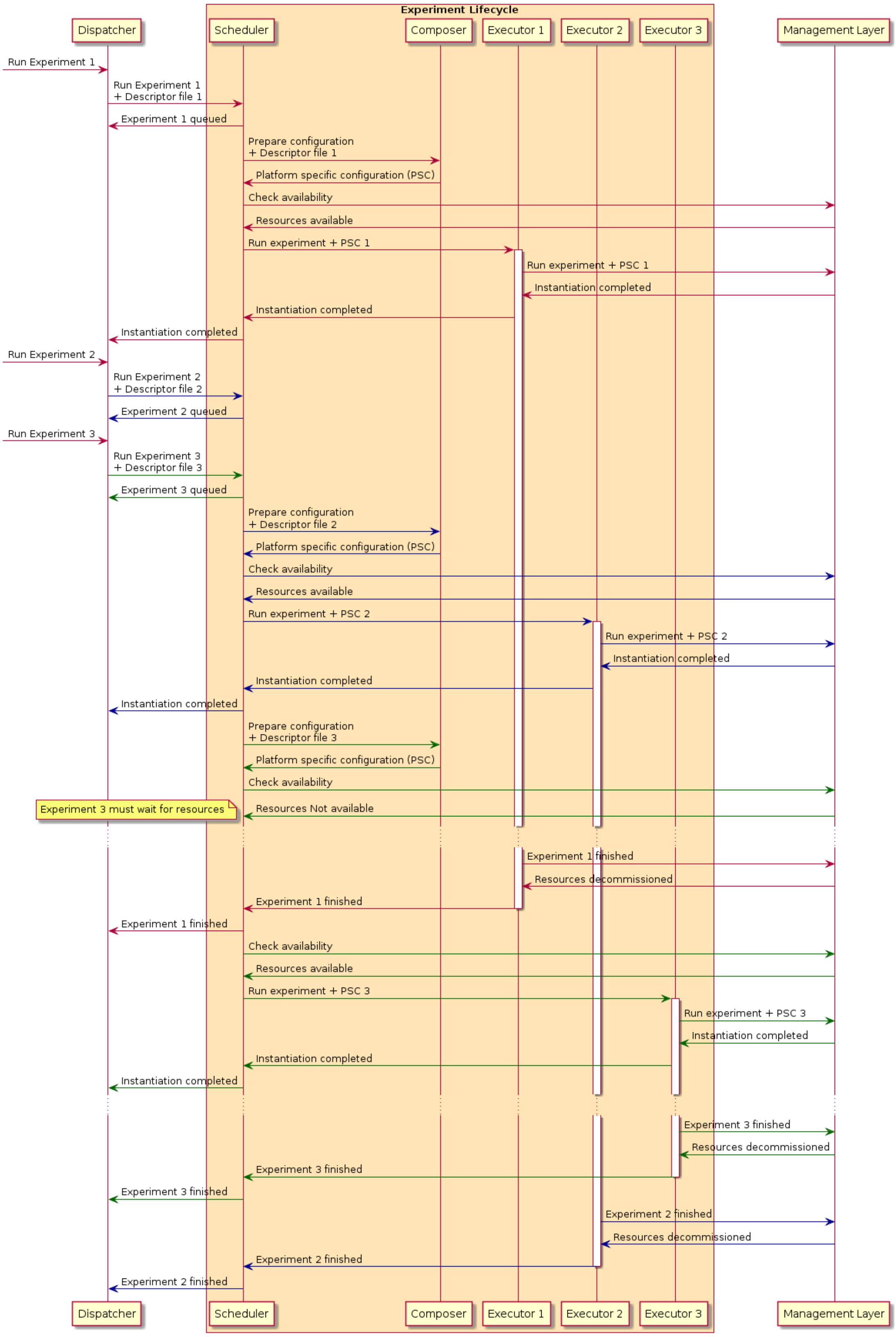

Figure 4 Sequence diagram of the execution of concurrent experiments

3. Experiment Set-Up In 5GENESIS Platforms

3.1. The 5GENESIS Experiment descriptor

An experiment descriptor is a file that describes all the information needed to execute the experiment and compute the KPIs associated with the experiment. An experiment descriptor can imply the execution of one or more test cases, in one or more scenarios and in one or more slicing configurations.

In 5GENESIS the experiment descriptor contains the following information:

- **Experimenter details:** experiment ID, owner ID (person defining and running the experiment), organization, platform ID, etc. This information will be managed by the Security Manager. (Mandatory)
- **Type of experiment:** custom or standard. (Mandatory)
- **List of test cases to be executed** (Mandatory): Defines the list of KPI's that will be computed as well as the procedure and the measurements. *e.g. test case Latency, test case Throughput, etc.*
- **Slicing configuration**: slice eMBB_1
- **Service** (only for custom experiments)**:** Service configuration
    - NS's: Network services that are used in the experiment including the configuration for each of its VNF's (Virtual Network Functions) including regions of deployment (core or edge deployment)
    - PNF's: configuration for the physical network functions
- **Traffic sources** (only for custom experiments to define traffic profiles different from those defined in the standard experiments): that can be used to emulate the traffic from real applications or to reproduce background traffic conditions.
- **Application under test** (only for unattended custom experiments)
- **UE identification** (only for unattended custom experiments): UE1, UE2, etc.
- **Scenario** (Mandatory): can define Multi-cell, SNR, fading profile, backhaul link, etc. (depends of the features available at the testbed, these features are storage at the platform registry. This component is part of the 5GENESIS platform architecture defined in D2.2).
- **Duration** (only for custom experiments)
- **Number of repetitions (only for custom experiments)**
- **Intermediate reporting of KPI's**: yes/no
- **Time between intermediate reports**

Figure 5 shows an example of a custom experiment descriptor and Figure 6 shows an experiment descriptor for a standard experiment using the YAML syntax. The final experiment descriptor format of 5GENESIS will be defined in Task 3.8 and included in deliverable D3.15 “Experiment and Lifecycle Manager”.

```
ExperimentDescriptor:
  Experimenter:
    ID: Experimenter1
    Organization: Organization1
  Experiment:
    ID: Experiment1
    Platform: 5GenesisPlatform1
    Type: Custom
    Unattended: True
    Scenario: Dense Urban
    SlicingConfig: 'Slice eMBB_1'
    TestCases:
        - 'Latency'
        - 'Throughput'

    NSs:
      Services:
        - Id: Service1
          Configuration:
            - Parameter1: False
          VNFs:
            - Id: CustomVNF1
              Location: CORE
              Configuration:
                  - Parameter1: True
                  - Parameter2: 3.5

          PNFs:
            - Type
            Provider
            Configuration:
                - Parameter 1
                  Parameter 2
                  Parameter 3
    Application:
      - com.test.application.apk
    UE:
      - OS: Android
        ID: edb90ee0f59f8bf2d4e4
      - OS: Android
        ID: f17c5fc1d2b3bdf09f16
    Duration: '3:30'
  Reporting:
    Enabled: True
    Period: '0:15'
```

Figure 5 Custom experiment descriptor

```
ExperimentDescriptor:
  Experimenter:
    ID: Experimenter2
    Organization: Organization2
  Experiment:
    ID: Experiment2
    Platform: 5GenesisPlatform2
    Type: Standard
    Unattended: True
    Scenario: Dense Urban
```

```
    SlicingConfig: 'Slice eMBB_1'

    TestCases:
      - 'Latency'
      - 'Throughput'
  Reporting:
    Enabled: True
    Period: '0:15'
```

Figure 6 Standard experiment descriptor

3.2. 5GENESIS Portal

The experiments will be defined through the 5GENESIS Portal, as it is defined in the 5GENESIS architecture. This section refers to the functionalities that the 5GENESIS Portal will offer.

The main functionality offered by the Portal will provide the following features:

- Login process to the experimenter
- Information about the platform and its supported features
- Provide and guide through the definition of an experiment
- Request for the experiment execution
- Access to the experiment results (per execution)
- Access to the current status of the experiment (scheduled, running, error, finished)

The information introduced through the Portal will be used to generate the Experiment Descriptor. The 5GENESIS Portal will offer a dashboard that gives access to the entire set of functionalities offered by the corresponding platform. From the dashboard, an experimenter, e.g. a vertical, can define new experiments, view the results obtained on previous executions, schedule the execution, make reservations for the platform and manage deployed VNF's. Through the experimenter will also have access to system-wide notices and a direct access to information about the latest performed activities. Moreover, the experimenter will be able to perform basic setup activities such as setting a name for the experiment, as well as selecting its type. Depending on whether the experiment is, Standard or Custom, the experimenter will be able to select different configuration values.

3.3. Specifications for the 5GENESIS experimentation

In 5GENESIS three main documents have been identified as the key sources for specifying the experimentation process. Two of them belong to the 3GPP, organisation with a long trajectory in the definition of a test case for conformance testing. In these tests it is also an important part, the definition of the conditions of the test. In particular, TR 37.901 specifies the test procedure to run throughput tests at the application level in a set of scenarios which covers a wide range of test conditions focused on LTE radio parameters. As part of the 5GENESIS project we will go a step further and will define a set of 5G end-to-end scenarios for the testing of 5G KPI's. For doing this we will use also as reference the TS 28.554. TS 28.554 is focused on 5G KPI's and network slicing, but, at this moment, it only covers three KPI's. We will try to extrapolate the guidelines provided in this specification to cover the list of KPI's considered in 5GENESIS (see D2.1) and define the slicing configurations.

The 3rd document is a document, which is in the same direction that the approach adopted 5GENESIS, defining a list of representative 5G scenarios and an initial list of test cases oriented to the testing of 5G KPIs. The scenarios defined also cover radio configuration; we aim at extending these definitions to provide end-to-end scenarios covering the configuration of all the components of an end-to-end 5G connection.

In order to evaluate how a) network conditions and configurations impact on the performance of applications and services under test, and b) 5G systems will manage the issues derived from them, it is needed to define a set of network scenarios that reproduce realistic conditions and ensure the relevance of the obtained results.

The scenarios have to mimic real operation conditions and covering the most representative situations: high-dense areas, rural areas, urban-macro, static conditions, vehicular and high-speed conditions, low to high traffic loads, etc. These scenarios will specify the working conditions of all the components of the infrastructure layer: radio conditions, transport conditions, core network, servers, etc., and will be based on the specifications previously defined.

An example of scenario can be the following:

```
Single cell
     -Average Signal Received: -95 dBm

UE1 Location:
     -UE1: outdoor cell edge
UE1 mobility:
     -speed: 3 Km/h
RAN-Core link:
    -Type: Fiber connection

PoD1-PoD2 link:
    -Type: Fiber connection
```

Figure 7 Urban-macro scenario

In order to evaluate the performance of applications/services when using a network slicing approach, it is also required to have the definitions of end-to-end slices covering the most relevant 5G uses cases: eMBB, mMTC and URLLC. The slice manager component of the 5GENESIS architecture and the end-to-end slices will be defined in Task 3.2 and presented in more detail through the corresponding deliverables. The 5GENESIS slicing configurations will be based on 3GPP specifications analyzed previously.

3.3.1. Reference testing specifications

To specify the information needed for a test case and the requirements for the scenarios and the slice configuration the reference documents mentioned above are briefly reported here.

- *Definition of the testing framework for the NGMN 5G Pre-commercial network trials, NGMN ALLIANCE, (January 2018), Version 1* [1]. The main goal of this document is to evaluate the performance of 5G new radio (NR) based on the 3GPP standard, and to assess and benchmark its performance. The initial focus is on Phase 1 (Release 15) of the 5G NR standardization, which primarily considers eMBB and some aspects uRLLC use cases.

The NGMN document defines a set of test cases for the list of KPIs considered. The structure of these test cases has been used as reference for the definition of the 5GENESIS test case template. Moreover, the NGMN document identifies a set of scenarios based on 3GPP 38.913 [15], which will be taken as a reference for the definition of 5GENESIS scenarios for the testing and validation of the 5G KPI's. More details about the NGMN document can be found in Annex 1.

- *3GPP TR 37.901, Technical Specification Group Radio Access Network; User Equipment (UE) application layer data throughput performance* [3]. The objective of study item TR 37.901 is to define test procedures to measure UE data throughput performance at the application-layer. It is one of the few specifications from 3GPP that covers performance measurement at the application level. the project should adopt from this specification the detailed scenarios specified for the testing of the throughput. However, this study item is limited to LTE and W-CDMA (Rel-5) (HSDPA). 5GENESIS will use this specification to extend the scenarios identified in [1]. This recommendation is analyzed in detail in Annex 2.

- *3GPP TS 28.554, Management and orchestration; 5G end to end Key Performance Indicators (KPI)* [3]. 3GPP has released its specifications for end to end Key Performance Indicators (KPI) for the 5G network and network slicing. The current version that is reviewed here is 3GPP TS 28.554 v15.0.1. The document specifies the following KPI categories: i) accessibility; ii) integrity; iii) utilization, while some 5G KPIs are left to be defined in the future releases (i.e. retainability, availability and mobility). The list of KPIs under this specification document is more generic than the 5G KPIs that are advertised as the main KPIs that need to be validated (list of 5G KPIs) and are somehow included into the ones discussed in the 3GPP document. In 5GENESIS the list of relevant KPI's considered to be validated are presented in D2.1. More details about the NGMN document can be found in Annex 3.

3.4. The 5GENESIS test case template

This section defines the concept of the test case as adopted in 5GENESIS project partially based on the NGMN document analyzed previously [1]. As already introduced, test cases specify the procedure to compute 5G KPI's, and will be fully defined in WP6. In 5GENESIS, a test case comes in the form of a document or **test case template**, that specifies the set of procedures and measurements needed to measure and validate a KPI.

Since the test cases are part of the experiment descriptor, the information provided by the test cases will be used to configure and control the components of the testbed for the execution of the experiment. In particular, the 5GENESIS test case template is initially formed by the following descriptions:

- **Description of the test case.** Target of the test case, e.g., the definition of the targeted KPI's.

- **Test procedure:** Description of the execution of the test
    - Test sequence to run the test case: sequence of actions for running the experiment and collect the measurements needed to compute the KPI's.
    - Experiment duration (per iteration).
    - Number of iterations of the experiment. Number of repetitions of the experiments to obtain relevant statistical results.
    - Measurements collected to compute the KPI's.
- **Procedure for KPI's calculation.** Description of the procedure applied (statical aggregation, algorithm, etc, …) to compute the KPI's based on the raw measurements collected.

- **Expected output of the test case**
    - Results provided

Below we provide a test case template to be used as a reference for 5GENESIS test cases.

Test case ID	
Description:	Description of the context of the test. Definition of the targeted KPIs. Definition of the measurements needed to compute the KPIs.
Test procedure:	Test sequence: Sequence of actions for running the test.
	Duration
	Number of iterations
	Metrics to be collected
	Example *1. Start logging at UE* *2. Start injecting traffic in the downlink* *3. Wait for 2 minutes* *4. Stop traffic injection* *5. Logout* *6. Calculate and record the amount of bytes/seconds received at the UE* *8. Repeat 5 times step 1 to 7*
KPI computation procedure	Average amount of bytes/seconds transmitted in the downlink in the 5 iterations.
Test case output	Raw results, counters, KPIs
	Example: *Collected logs and metadata* *KPI: Measured downlink data rate*

The final report after the execution of a test case in a particular scenario and slicing configuration should follow the structure bellow:

Test Case ID		
Description		
Executed by	Partner	Date
Purpose		
Scenario		
Slicing configuration		
Components involved		
Metrics collected		
Tools involved		
Results and KPI's		
Comments		

4. IMPLEMENTATION OF THE METHODOLOGY

4.1. Candidate frameworks

This section reviews the experimentation frameworks adopted as the basis for the definition of the 5GENESIS experimentation methodology. In particular, the MONROE project [17] has been considered as the starting point for the 5GENESIS monitoring functionalities, devoted to the collection of parameters for the KPI's evaluation, while the management framework developed as part of the TRIANGLE project [16] is the core of the Coordination layer and of the ELCM, as defined in the 5GENESIS architecture.

4.1.1. MONROE

MONROE ([23], [24], [25]) is the first transnational, open platform for independent, large-scale, end-to-end (E2E) experimentations in commercial Mobile BroadBand (MBB) networks. It was developed in the scope of the MONROE EU H2020 project (RIA-644399) and is now operated and maintained by the MONROE Alliance.

As shown in Figure 8, the MONROE platform is comprised of three main software/hardware components:

1. The MONROE nodes, on which the experiments (measurement probes) run;
2. The User Access and Scheduling system, that controls the access to the nodes and schedules the experiments;
3. The Back-end system, which enables large-scale data collection from the experiments.

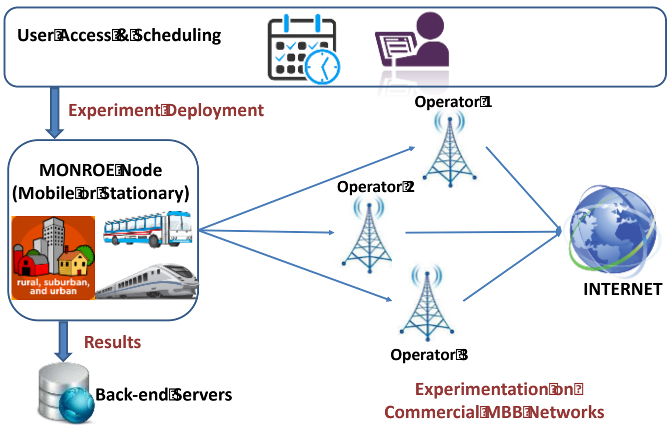

Figure 8 Main building blocks of the MONROE measurement system.

The current set of MONROE nodes counts about 150 stationary and mobile physical entities distributed in four European countries, each one providing multi-homing connectivity to 4G/LTE and WiFi networks. In particular, each node is connected to three different MBB providers using commercial grade subscriptions, and allows to run experimentation software on top of its hardware.

Adopting an innovative Experimentation as a Service (EaaS) paradigm, the MONROE platform allows experimenters and users to design an experiment, access the nodes via the User Access and Scheduling system, run the experiment within the nodes, and, finally, collect the experiment results on Back-end servers, together with other, possibly useful, context information (e.g., node location and overall status, type of subscription, connection status, and so on). User experiments are designed and run inside virtualized environments, formally Docker containers, in order to ensure separation and containment of processes, and provide a reliable and secure experimentation environment.

The MONROE platform enables users to conduct a wide range of experiments and measurements, thus forming the basis for a complete MBB monitoring and analysis function, that is expected to play a crucial role in future 5G design and deployment. Long-term large-scale measurements can be conducted in order to identify key MBB performance parameters, and capture the relationship between them, the underlying network status and configuration, and the final user experience. As a result, this can be used to optimize network protocols for traffic management, energy efficiency, security, and reliability, just to mention a few. Moreover, accurate and meaningful monitoring and assessment of MBB networks can trigger an optimized design and deployment of innovative and heterogeneous services and applications built on top of 4G/5G networks, e.g. massive Machine Type Communication (mMTC) and Internet of Things (IoT).

Besides user-specific experiments (see https://github.com/MONROE-PROJECT/Experiments for the full set) MONROE nodes are currently designed in order to:

1. Periodically retrieve context information (metadata), e.g. signal strength and network configuration data for each connection interface, GPS data, and several node parameters, such as CPU load and temperature, memory usage, and so on. Full set of available metadata is provided in Table 1. Collected metadata reports a timestamp field that enables correlating these entries with experiment results.
2. Periodically run connectivity measurements for each connection interface, via ICMP ping to a fixed external server (currently, Google's DNS at 8.8.8.8). The IP round-trip-time (RTT) values, which represent basic connectivity and delay information for each network under test, are collected and transferred to the Back-end servers.
3. Provide analysis of traffic patterns at network and transport levels, via *Tstat*, the passive *mPlane* monitoring probe that collects, for each interface, detailed flow level statistics.

In addition, a measurement tool suitable for evaluating data rate, latency and application-level performance for applications such as Web and video is also available. Within the suit, for example, MONROE-Nettest is a configurable tool for MBB speed measurements, while MONROE-browsertime is a container that collects a number of web performance metrics while visiting a target webpage using either Firefox or Chrome browser with different web protocols, such as HTTP1.1/TLS, HTTP2, and QUIC.

Finally, it can be also observed that the multihoming capabilities of nodes can be leveraged to investigate multipath multimedia transmission over 4G/LTE, WiFi, and future 5G in an E2E fashion, with the goal of maximizing user-side Quality of Experience (QoE). This in turn can be obtained through the definition of advanced, intelligent mechanisms for the selection of the best available technology, path, protocol, scheduling policy, etc, thus optimizing E2E resource allocation, management, and service instantiation, with particular emphasis on 5G-specific dynamic slice management.

TOPIC	FIELDS
Modem	Interface name, operator name, cell ID, physical cell ID, device mode, submode and state, EC/IO, evolved base station ID, ICCID, IMSI, IMEI, IP address, MCC, MNC, LAC, RSRP, frecuency, RSRQ, band, RSCP and RSSI.
GPS	Longitude, latitutde, altitude, speed, satellite, count, NMEA raw string.
Node sensors	CPU temperature, boot counter (session ID), start time, uptime, CPU usage counters, memory and swap space usage.

Table 2 Main “MONROE” metadata topics: each entry includes the Node ID from which it originates, a timestamp and a node-local monotonically increasing sequence number.

4.1.1.1. MONROE insights to 5GENESIS

As introduced above, the MONROE experimentation framework provides an important starting point for the 5GENESIS Monitoring and Analytics system, devoted to the collection of parameters for 5G KPI's analysis. MONROE provides a plethora of SW measurement probes for several QoS/QoE KPI's, up to the application layer. As a matter of fact, the analyses carried out during the MONROE project highlights the importance of providing an extensive evaluation for each considered KPI, comparing different measurement solutions under several environmental scenarios, yet adopting a uniform measurement methodology. The MONROE experience in designing and using a vast set of probes is the key for a proper definition of 5GENESIS monitoring procedures, in order to accurately and meaningfully measure the KPI's.

While MONROE provides a complete HW/SW experimentation framework, the MONROE probes are transparent to the underlying HW. Within 5GENESIS the MONROE probes will be adapted to run on top of virtual MONROE nodes that can be deployed as VNF's within the 5GENESIS facility. While the main focus of MONROE is on E2E performance, this will allow the probes to be deployed at both user and network sides, thus providing the same measurement methodology at different vantage points in all 5GENESIS platforms. The measurement probes are Docker-containerized, and thus isolated from the management SW, aiming to contain the activity of the experimenters, and thus avoid security risks, HW/SW corruption and thus unreliable results. Overall, these aspects target consistency, reliability and repeatability of experiments, and naturally fit well the 5GENESIS objectives.

As mentioned above, the base MONROE SW is dedicated to the collection of UE-side parameters, up to application layer “metadata”, e.g., GPS locations, signal strengths, and basic connectivity/performance tests. Metadata are stored in a NoSQL distributed database, to ensure scalability and availability. Such experiences are key inputs for the 5GENESIS Monitoring and Analytics system, which requires the collection of large amounts of heterogeneous network parameters for a complete KPI's validation. Additionally, MONROE metadata is made available to experimenters through a standardized interface. Within 5GENESIS, 5G network metadata provided by the ECM UE will be integrated with the MONROE virtual node offering the possibility to use such metadata in a structured way in custom experiments.

4.1.2. TRIANGLE Experimentation framework

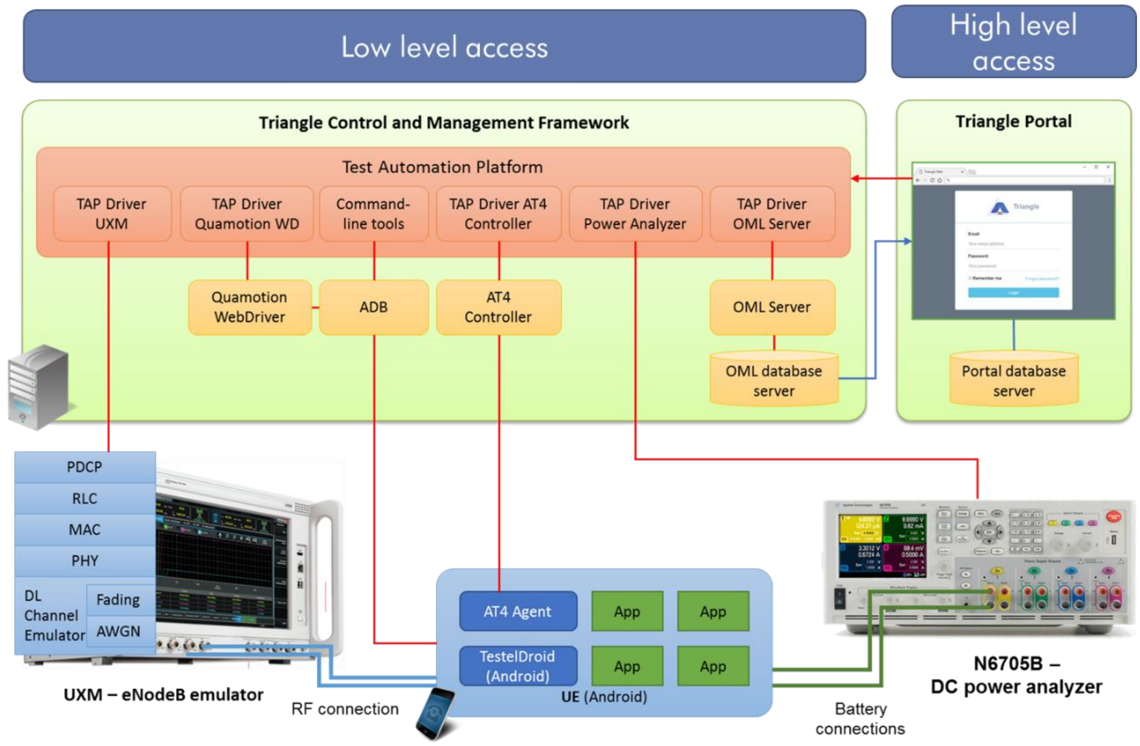

Figure 9 TRIANGLE experimentation framework

TRIANGLE offers a complete framework for automating the testing of mobile applications and mobile devices in realistic network conditions [19]. In this context, the main target of the TRIANGLE testbed is to simplify the testing procedures not to overwhelm experimenters with all the configuration details, or just to prevent them from dealing with a very time-consuming task. The adopted solution is the provision of a pool of high-level scenarios, which wrap the complexity of the configurations of the all elements of the testbed. The testbed also offers a low-level access for users with different profiles such as researchers, devices manufacturers or operators that could be interested in accessing the configuration details of the testbed.

This framework has been developed as part of the project "5G Applications and Devices Benchmarking" (TRIANGLE) funded within the framework of the "Horizon 2020" (H2020) initiative under FIRE/FIRE+ objective devoted to Future Internet Research and Experimentation. More details about the framework can be found in [10].

The TRIANGLE testbed integrates mobile networks emulators, commercial user equipment and other lab instruments such as power analyzer. To configure and control the different components, the testbed uses the Test Automation Platform (TAP) from Keysight. Each testbed component is controlled through a TAP driver, which serves as a bridge between the TAP engine and the actual component interface.

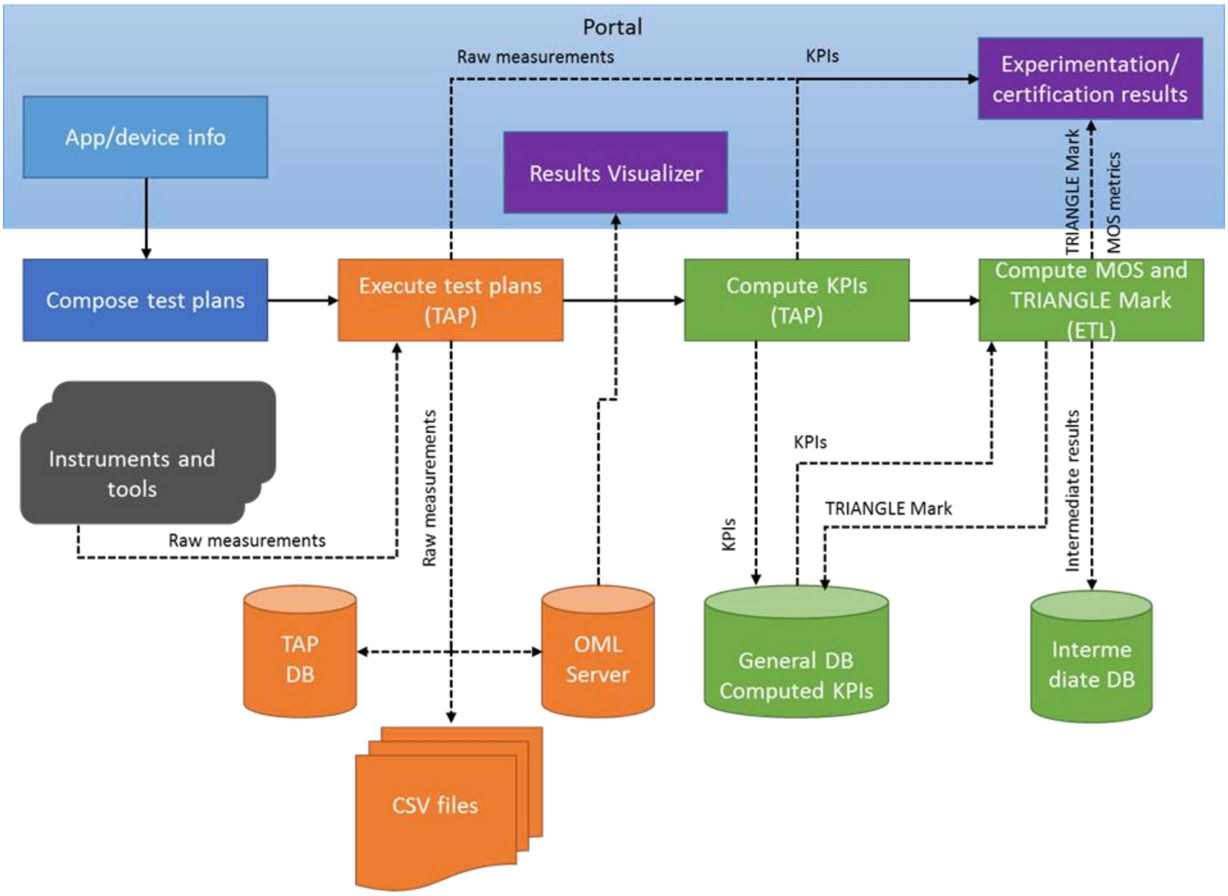

Figure 10 Post-processing chain in TRIANGLE

The TRIANGLE framework includes a post-processing chain that enables the conversion of the raw results gathered at the different components of the testbeds into MOS (Mean Opinion Score) values. The TAP test plans (scripts) configure any equipment needed to run the experiment, execute the body of the test, and collect results produced by any of the measurement probes and tools integrated into the testing framework. The results are stored in different formats and data bases (TAP data base, csv files and OML data base). The next step is to compute the KPI's from these measurements.

The measurements stored in the TAP database serve as the source material for extracting and computing the KPI values. The KPIs computation is done by TAP through the execution of specific KPI computation steps implemented in TAP. The atomic KPI's are stored in the general database.

Next step in the post-processing chain is the ETL (Extract, Transform and Load) framework. The ETL framework has as an input the computed KPI's and implements all the required steps for the assessment of MOS metrics (the process computed a separated MOS per domain (power consumption, app user experience, device resources, etc. and aggregates all of them to provide a final MOS). The full methodology to compute MOS values is described in [18], [11].

The TRIANGLE testbed includes also an extension based on model-based testing ([20], [21], [22]). Model-based testing is a testing technique that uses a model of the system under test to extract the test cases. In TRIANGLE, the model describes the interaction of the user and the mobile app, and the objectives are a) to automatically produce a pool of app user flows that can be used in the test campaigns and b) to activate (in many different ways) the required application functionality required to evaluate the corresponding KPI.

4.1.2.1. TRIANGLE insights to 5GENESIS

During the TRIANGLE project a testing methodology for the benchmarking of mobile applications and devices has been built and refined. This methodology was also based on the definition of a set of representative network scenarios and test cases that cover the different domains, which influence the performance of mobile and devices.

As part of the development process, the procedures and tools for implementing the methodology were identified. Moreover, the developed methodology has tackled the problem of dealing with the control and management of different equipment, tools and probes. An effective way of controlling all the heterogeneous components of a testbed and correlating and aggregating the measurements from different sources was also developed.

The problems addressed by TRIANGLE are common to those identified in 5GENESIS. While the developed tools cannot be directly applied to 5GENESIS, the acquired knowledge has influenced the definition of the upper layer of the 5GENESIS architecture and will also be applied in the development of the Experiment Lifecycle Manager (ELCM).

4.2. TAP approach to implement the experimentation methodology

TAP is the technology adopted in the TRIANGLE project for the implementation of the experimentation framework. Based on the experience acquired on this software UMA has proposed to use it for the implementation of the experimentation methodology in 5GENESIS.

TAP (Test Automation Platform) provides a flexible and extensible test sequence and test plan creation. An experiment plan will be composed by one or several test plans, for example a test plan for configuring the radio and a different test plan per each one of the probes to be configured. TAP is a Microsoft .NET-based application that can be used stand-alone or in combination with higher-level test executive software environments developed by Keysight. Leveraging C# and Microsoft Visual Studio, TAP is a platform upon which it is possible to build test solutions.

TAP test plans consist of a sequence of test steps, which specify the set of actions that are performed on different TAP instruments (the logical entities that control the physical equipment on the Platform) or control the execution flow of the test plan (for instance, repeating certain steps or executing different actions depending on the results of a previous test step). TAP defines a set of basic test steps that are suited for most user requirements, and it is possible to develop custom steps in C# for complex or very specific needs.

TAP Instruments (and the functionally equivalent Device Under Test) define the logic for interacting with other entities in the facility. TAP includes generic instrument for controlling SCPI compatible equipment, while custom TAP Instruments can be developed using C#. Each TAP instrument encapsulates the configuration and management logic of the equipment, as well as defining the actions that can be performed on it.

Overall, TAP plays a key role in the implementation of the ELCM as it allows controlling and automating all the components present in the testbed.

TAP allows the definition of a set of parameters within a test plan (any if the configuration values available on the test steps) to be grouped and marked as *External Parameters*. Once a configuration value is made available externally, it is possible to modify this value before the execution of a test plan, making it possible to orchestrate the execution of TAP using an external entity. In this way, it is possible to use TAP for controlling the fine-grained interaction with the different equipment of the Platform, while keeping a separate, upper layer for general experiment orchestration.

For such cases, Python can be the best alternative to implement the rest of the functionality needed in the ELCM. Python is a modern programming language with a large community of developers that continues to improve with the inclusion of new built-in modules and features on each subsequent release. Additionally, there is a large amount of ready to use, open source, frameworks and libraries suited for common needs. This can facilitate the development of the components of the Platform, for example, by using Python's standard libraries on the definition of the REST API that supports the communication between different components.

Since Python is an interpreted language it is also possible to modify the code rapidly, with no need for a separate compilation stage on a development machine, which can drastically reduce the complexity of testing and the time needed for the resolution of simple issues.

5. Integration Plan

The methodology introduced in this document ensures the repeatability and the comparability of the experiments in the same platform and also between platforms.

In this section, we have identified the key component of the 5GENESIS methodology, and we provide an estimation of the delivery dates for each feature and for each platform.

5GENESIS scheduling is to be made in 3 iteration cycles (see Figure 11) to integrate features per platform. The *Main Phase* refers to the core work of the project, which is the development of the Facility and the interim tests and experiments. In order to achieve a smooth evolution and concurrent validation, this phase includes three main integration cycles, each one lasting for 6 months. The scope of each integration activity/cycle is to upgrade each testbed with the latest technical achievements and results from all relevant R&D activities as well as to align it with the most current edition of the 5G standards. Each integration cycle will be followed by a 3-month testing phase, during which: a) the current release of the facility will be fully verified, ii) selected experiments will be implemented and relevant KPI's will be validated, and iii) the technical specs of the facility and the implementation plan will be updated, according to the lessons learnt from the integration and testing activities.

The implementation of the cycles is done in Work package 4, where each platform has one specific task.

Figure 11 Initial plan for the 5GENESIS integration cycles

Table 3 and Table 4 provides an estimation of the integration of the experimentation features per platform. The estimation is aligned with the 5GENESIS integration cycles shown in Figure 11. The components enumerated in the tables are part of the coordination layer of the 5GENESIS reference architecture described in D2.2, this layer is in charge of offering the northbound interface for the experimenters and coordinating the execution of the experiment. The rest of the concepts included in the tables have been defined in the 5GENESIS experimentation methodology introduced in this document. The combination of all off them will have as result the complete implementation of the 5GENESIS experimentation methodology in cycle 3.

Table 3 Experimentation methodology components in the first integration cycle

	Athens platform	Berlin platform	Limasol platform	Málaga platform	Surrey platform
Open API's + Dispatcher	No	No	No	No	No
Experiment Life Cycle	No	No	No	POC	No
Portal	No	No	No	POC	No
Custom experiments	POC	POC	No	POC	POC
Standard experiments	No	No	No	No	No
End-to-end slices	POC	No	No	No	No
VNF's	Yes	Yes	Yes (LBO, IoT interoperability, WAN acceleration)	No	Yes
Scenarios	MOCN	POC (mmWave backhauling)	LTE	POC	POC (LoRa + NB-IOT)
Un-attended experiments	Yes Network performance measurements	POC: Delay and Throughput evaluations of the core testbed infrastructure	No	POC: Video 5G-NR indoor (only DL)	No
Attended experiments	Yes POC Delay, throughput (4G)	POC: 4G Delay and Throughput	POC: IoT scenarios over satellite backhaul	POC: MCS 4G	POC
Security Manager	NA	NA	No	No	No

*POC=Proof of Concept

Table 4 Experimentation methodology components in the second integration cycle

	Athens platform	Berlin platform	Limasol platform	Málaga platform	Surrey platform
Open API's + Dispatcher	Preliminary	No	No	First version	No
Experiment Life Cycle	POC	POC	First version	Yes	Yes
Portal	No	No	No	Yes	No
Custom experiments	Yes	Yes	Yes(UC2: Rural scenarios)	Yes	Yes
Standard experiments	Yes	Yes (Initial set of standard experiments)	Yes (UC2: Rural scenarios)	Initial set of standard experiments	Yes (Initial set of standard experiments)
End-to-end slices	Yes in the core and the transport. However still preliminary for the radio	POC	Limited features	Yes	POC (5GC + RAN)
VNF's	Yes	Yes	Yes	Yes	Yes
Scenarios	Yes	Yes	Yes	Yes	Yes
Un-attended experiments	Yes	Yes	Limited features	Yes	Yes
Attended experiments	Yes	Yes	Yes	Yes	Yes
Security Manager	Yes	NA	Yes	Yes	NO

*POC=Proof of Concept

The status of the development and the integration of the experimentation features will be updated in WP3 and WP4 deliverables respectively.

6. REFERENCES

[1] Definition of the Testing Framework for the NGMN 5G Pre-Commercial Network Trials, NGMN, Version 1

[2] 3GPP TR 37.901, Technical Specification Group Radio Access Network; User Equipment (UE) application layer data throughput performance

[3] 3GPP TS 28.554 Management and orchestration; 5G end to end Key Performance Indicators (KPI)

[4] 3GPP TS 28.554 v15.0.1 - Management and orchestration; 5G end to end Key Performance Indicators (KPI) (Release 15), Technical Specification, Sept. 2018, 3GPP.

[5] 3GPP TS 28.552 V16.0.0 - Management and orchestration; 5G performance measurements (Release 16), Technical Specification, Dec. 2018, 3GPP.

[6] 3GPP TS 28.553 V0.4.0 - Management and orchestration of networks and network slicing; 5G Core Network (5GC) performance measurements and assurance data (Release 15), Technical Specification, Dec. 2018, 3GPP.

[7] 3GPP TS 32.425: "Performance Management (PM); Performance measurements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN)".

[8] 3GPP TS 32.451: "Key Performance Indicators (KPI) for Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Requirements".

[9] 3GPP TS 32.404: "Performance Management (PM); Performance measurements - Definitions and template".

[10] EU H2020 TRIANGLE project, "Deliverable D3.5 Implementation report on the testing framework Rel 4, 2019: https://www.triangle-project.eu/project-old/deliverables/

[11] EU H2020 TRIANGLE project, "Deliverable D2.2 Final report on the formalization of the certification process, requirements and use cases", 2017, available online: https://www.triangle-project.eu/project-old/deliverables/

[12] 3GPP TS 36.521-1, Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) conformance specification Radio transmission and reception Part 1: Conformance Testing

[13] 3GPP TS 34.121-1, User Equipment (UE) conformance specification; Radio transmission and reception (FDD); Part 1: Conformance specification

[14] 3GPP TS 34.122, Terminal conformance specification; Radio transmission and reception (TDD) Test Automation Platform: https://www.keysight.com/en/pc-2873415/test-automation-platform-tap?cc=US&lc=eng

[15] 3GPP TR 38.913, V14.4.0, "Study on Scenarios and Requirements for Next Generation Access Technologies (Release 14)", Oct. 2017.

[16] TRIANGLE Project – 5G Applications and Devices Benchmarking, www.triangle-project.eu

[17] MONROE Project Measuring Mobile Broadband Networks in Europe, www.monroe-project.eu

[18] Almudena Díaz-Zayas, Laura Panizo, Janie Baños Polglase, Carlos Cárdenas, Michael Dieudonne: QoE Evaluation: The TRIANGLE Testbed Approach. Wireless Communications and Mobile Computing 2018: 6202854:1-6202854:12 (2018)

[19] Almudena Díaz-Zayas, Andrea F. Cattoni, Carlos Cárdenas, Oscar Castañeda, Janie Baños Polglase, Pedro Merino, Michael Dieudonne, TRIANGLE: a Platform to Validate 5G KPIs in End to End scenarios. IEEE CAMAD 2018: 1-7

[20] Ana Rosario Espada, María-del-Mar Gallardo, Alberto Salmerón, Pedro Merino: Performance Analysis of Spotify® for Android with Model-Based Testing. Mobile nformation Systems 2017: 2012696:1-2012696:14 (2017)

[21] Laura Panizo, Alberto Salmerón, María-del-Mar Gallardo, Pedro Merino: Guided test case generation for mobile apps in the TRIANGLE project: work in rogress. SPIN 2017: 192-195

[22] Laura Panizo, Almudena Díaz, Bruno García: An Extension of TRIANGLE Testbed with Model-Based Testing. SPIN 2018: 190-195

[23] Mancuso, Vincenzo, Miguel Peón Quirós, Cise Midoglu, Mohamed Moulay, Vincenzo Comite, Andra Lutu, Özgü Alay et al. "Results from running an experiment as a service platform for mobile broadband networks in Europe." Computer Communications 133 (2019): 89-101.

[24] Alay, Özgû, Vincenzo Mancuso, Anna Brunstrom, Stefan Alfredsson, Marco Mellia, Giacomo Bernini, and Hakon Lonsethagen. "End to End 5G Measurements with MONROE: Challenges and Opportunities." In 2018 IEEE 4th International Forum on Research and Technology for Society and Industry (RTSI), pp. 1-6. IEEE, 2018.

[25] Alay, Özgü, Andra Lutu, Miguel Peón-Quirós, Vincenzo Mancuso, Thomas Hirsch, Kristian Evensen, Audun Hansen et al. "Experience: An open platform for experimentation with commercial mobile broadband networks." In Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking, pp. 70-78. ACM, 2017.

[26] 3GPP TS 23.501 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; System Architecture for the 5G; Stage 2 (Release 15), 2018

Annex 1 Definition of the testing framework for the NGMN 5G Pre-Commercial Network Trials

The NGMN document provides the definition of the KPIs and the test case associated to each one of them. As commented before, each test case includes a detailed definition, considered scenarios, the target values and most importantly a clear testing methodology.

The list of KPIs included in the NGMN document is the following:

- Latency
- User Throughput
- Cell capacity
- Spectral efficiency
- Coverage
- Mobility
- Reliability and Retainability
- User Experience
- Energy efficiency
- Inter-RAT procedures
- RAN architecture split

The structure of the test case defined in the NGMN document includes a section devoted to the formal definition of the KPI targeted by the test case.

The next portion of the structure of the test case defined in the NGMN document reports the first part of the "Test environment", that is the "Test setup". In the case of UP Latency, the Test setup includes information on number of users, and percentage of network resources in usage.

The second part of the "Test environment" is the "Test configuration" that provides the critical radio configurations settings for running the KPI validation experiments.

The "Logging part" specifies the metrics that need to be collected in order to assess test conditions and evaluated the specific KPI. Then, the Test procedure (Figure 3) details the procedure for the activation of the probes able to monitor the metrics to be collected, as well as of traffic generators (if needed). In addition, the number of experiment repetitions and, possibly, the duration of each, is specified.

Finally, in the "Success criteria" section, advanced procedures to compute the KPIs (average, rate, number of samples, etc.) as well as KPI target values are reported.

ANNEX 2 3GPP TR 37.901, TECHNICAL SPECIFICATION GROUP RADIO ACCESS NETWORK; USER EQUIPMENT (UE) APPLICATION LAYER DATA THROUGHPUT PERFORMANCE

The objective of study item TR 37.901 is to define test procedures to measure UE data throughput performance at the application-layer.

The test environment defined in TR 37.901 includes the specification of signal levels, fading profiles, noise and interference levels, to simulate typical network conditions with re-use of radio conditions already specified in TS 34.121-1 [13], TS 34.122 [14] , and TS 36.521-1 [12]. In particular, MAC scheduling and downlink radio conditions to simulate typical network conditions are provided. In relation to signal levels, a power sweep test is proposed to characterize the receiver under a range of signal levels in a relatively low-noise environment where the UE noise floor may be the dominant factor in determining SNR. Typical transmission modes used in operation based on the downlink signal levels presented to the UE are also considered (Mode 1, Mode 2, Mode 3,…,Mode7).

The common RF test conditions for performance testing are defined in TS 36.521 and may vary depending on the transmission mode being used.

In a typical test, two separate phases will be present. In the first, default non-fading conditions are used to set the device under test into known favorable initial conditions. Then, the actual impaired channel conditions will be applied as required. The test procedure is thus orchestrated with an initial step for basic cell configuration.

Within this scope, the following parameters are set:

- Power = -85dBm. This is the typical initial connection power for the devices to attach to the network, and it is also the default power for static propagation conditions when neither noise nor fading are required.
- FDD. Only frequency division has been implemented so far. TDD is currently pending, and will be implemented in the future. However, the core of the methodology is exactly applicable as it is.
- Cell ID = 0. For the purpose of these tests, any cell identifier could have been chosen but, to ensure consistency between different test implementations, the 3GPP specifications and recommendations tend to explicitly state all the parameters in very high detail.
- TM1. This is the Transmission Mode used when the cell transmits using a single antenna without beamforming. As a baseline, all LTE UEs use two receiving antennas even when the base station is transmitting using a single transmitting antenna.
- TX 1x2 ->SIMO. SIMO states for single transmission antenna and two (multiple) receiving antennas.

- BW = 20MHz. The maximum applicable channel bandwidth per band is used as recommended by 37.901. Typically, the bandwidth will be 20MHz, but note that not all the bands allow all the bandwidths defined in the specifications. As an example, some bands may have 10MHz as a bandwidth, hence only channels with a lower bandwidth fit into them.
- Frequency: Mid EARFCN. Following the same spirit as when choosing the channel bandwidths, only a single frequency is used. Selecting only the central frequency helps to reduce the number of test combinations, and thus keeping the test time and cost in control.
- Security parameters, as required by the SIM card being used. Typically, for conformance testing, special programmable test SIMs are used as the devices are not connected to real operators but to a simulated operator in the test equipment.

During the initial attach phase, a different scheduler is used than those required for the test procedures. The following scheduling configuration should be used during the attach:

- PRB Allocation: 6PRBs. This common configuration helps isolating from the actual channel bandwidth used in the tests, as 6PRBs will fit into the lowest possible channel bandwidth of 1.4MHz.
- Allocation centered in the channel bandwidth. In a channel bandwidth of 20 MHz with 100 PRBs (0 to 99), the 6 allocated PRBs will start in PRB 47. Please note that it will have significant impact on the coding rate in subframes 0 and 5, as there will be a lower amount of resources available for PDSCH transmission because of the collision with other physical channels.
- Using a low MCS (Modulation and Coding Scheme) Index value (of 5). Selecting a low value for the modulation and coding scheme index helps achieving a robust communication. It will force QPSK modulation, which has the highest constellation distance, and a high redundancy as each bit of user meaningful information will be coded within multiple physically transmitted bits.

  During the actual test operation, instead of a fixed static scheduler, TR 37.901 specifies the use of an adaptive scheduler. For that purpose, the Channel State Information (CSI) is configured to be periodically reported. Instead of the default CQI to IMCS mapping defined in TS 36.213, a different mapping has been defined in TR 37.901. Note that the actual mapping will depend on the transmission scheme.

LTE radio channel models are also specified in TR 37.901. In order to implement these models, the use of Faders and AWGN is required to emulate the radio environment scenarios. TR 37.901 is oriented to laboratory testing using a system simulator supporting LTE radio technology with IP connectivity. However, the list of specified parameters is very useful for the definition of the radio conditions in 5GENESIS test cases. For field testing the project will define a set of radio conditions to reproduce realistic conditions of mobile networks.

To finalize the study of this specification we have elaborated a list of the configuration parameters specified by the TR 37.901:

- Cell power level
- Carrier aggregation (a minimum of two cells is needed to support carrier aggregation)

- Duplex FDD (Frequency Division Duplex) and TDD (Time Division Duplex)
- Physical Layer Cell Identity
- Downlink transmission modes: [TM1] [TM2] [TM3] [TM4] [TM6] [TM7SISO] [TM7MIMO] [TM8] [TM9]
- Transmit diversity: [TX 1X1] [TX 1X2] [TX 2X2] [TX 4X2]
- E-UTRAN radio band FDD band range: [1..14] [17..31] TDD band range: [33..44]
- LTE Bandwidth: [1.4MHz] [3MHz] [5MHz] [10MHz] [15MHz] [20MHz]
- DL and UL frequency setting mode based on EARFCN
- Downlink radio bearer parameters and scheduling configuration: Number of resource blocks, resource block start, subframe allocation.
- CSI reporting configuration. Select type of CSI Reporting (Periodic / Aperiodic)
- CSI to MCS mapping. Determine the MCS based on the received CQI reports from the UE.
- Manual CQI table mapping
- CQI/PMI Configuration Index. CQI/PMI report configuration index to be used for each cell.
- RI Configuration Index. RI report configuration index to be used for each cell
- CSI Feedback Mode. CSI feedback mode to be used for each cell
- All subframes Set all Aperiodic CSI Request
- All Aperiodic CSI Request Set all Aperiodic CSI Request
- Individual Aperiodic CQI Setting Aperiodic CSI Request
- Codebook subset restriction. Mask used to define the allowed set of precoding matrices.
- Power control PUCCH target power. It specifies the target power in dBm for PUCCH.
- Power control PUSCH target power. It specifies the target power in dBm for PUSCH.
- Power control tolerance. Maximum accepted deviation from the target power.
- Power boosting per channel. Power boosting in LTE is mainly performed on the cell-specific reference signals (RS). As the RS are embedded into the overall signal bandwidth at certain resource elements (RE), in order to have a constant power for all OFDM symbols to avoid power variations at the UE receiver, different powers are allocated to each downlink channel for the OFDM symbols where RS is present or absent.
- Downlink max HARQ transmissions. Downlink maximum Hybrid Automatic Repeat Request (HARQ) transmission controls the number of times the LTE Application attempts to transmit a downlink transport block before it is discarded. Successful transmission of a downlink transport block is met with the UE sending an ACK or NACK.
- Uplink max HARQ transmissions. Uplink maximum Hybrid Automatic Repeat Request (HARQ) transmission controls the number of times the LTE Application expects the UE to attempt to transmit a transport block before it is discarded.
- Downlink HARQ redundancy version sequence. A redundancy version (or RV) is used during the rate matching process when a transport block is being encoded. If a transport block is retransmitted, a different redundancy version from the original transmission is generally used, this enables incremental redundancy to improve the chances of the transport block being decoded.
- Downlink HARQ retransmission allocation mode. This setting defines the algorithm used for PRB (Physical Resource Block) allocation during retransmissions.

The list of specified measurements is the following:

- UL/ACK count. Cumulative number of Transport Blocks (TBs) that have been ACKed.
- UL/NACK count. Cumulative number of TBs that have been NACKed
- Downlink ACK count. Cumulative number of TBs that have been ACKed
- Downlink NACK count. Cumulative number of TBs that have been NACKed
- Downlink StatDTX count. Cumulative number of TBs that have neither been ACKed or NACKed
- IP throughput

ANNEX 3 3GPP TS 28.554, MANAGEMENT AND ORCHESTRATION; 5G END TO END KEY PERFORMANCE INDICATORS (KPI)

TS 28.554 specifies a KPI definition template in order to allow categorization of the KPIs and the methods, tools, calculation that are used in order to measure and validate the KPIs. For reference we include the template that will also be used by the project in order to validate the KPIs.

The template comprises of the following information:

- Long name (mandatory)- *the name of the KPI*
- Description (mandatory)
- Logical formula definition (mandatory)
- Physical formula definition (optional) - *KPI formula description using the 3GPP defined counter names, it is only applicable if the counters used for the KPI formula are defined in 3GPP TS for performance measurements TS 28.552 and TS 28.553*
- Measurement names used in KPI (optional) - *defined only when underlying measurements for the KPI formula can be defined, i.e the previous field is filled*
- KPI Object (mandatory) - *can be any of NR and NG-RAN, 5GC, and 5GS*
- KPI Category (mandatory) - classification *of the KPI into one of the KPI categories listed above*
- Unit of KPI (mandatory) - *percentage, time interval, Erlang, kbps*
- Type of KPI (mandatory) - *MEAN, RATIO, CUM*
- Remark (optional) - *further information*

The document also presents a number of KPI that relate to the three categories that are analysed in this specifications document. Finally, the document presents example use cases for the end to end KPIs validation.

3GPP specifications provide two more documents that focus on 5G performance measurements (TS 28.552) [5] and 5G Core Network (5GC) performance measurements and assurance data (TS 28.553) [6]. The first document provides specifications for the performance measurements of 5G networks including network slicing. Performance measurements for NG-RAN as well as for 5GC are defined in this document. The latter provides specifications for the performance measurements and assurance data for 5GC Network Functions.

The performance measurements for NG-RAN applies also to NR option 3 in many cases, but not to the RRC connection related measurements which are handled by E-UTRAN for NR option 3 (those are measured according to TS 32.425 [7] and related KPIs in TS 32.451 [8]). The performance measurements are defined based on the measurement template as described in TS 32.404 [9].

Both documents provide more information on the measurement of specific metrics for the performance of the 5GC and 5G-RAN. The documents focus more on the performance measurement of the each 5G component (i.e. 5GC or 5G-RAN) while TS 28.554 focuses more on end-to-end KPI validation. 5GENESIS take into account the aforementioned documents in order to:

- Define the experimental descriptor (template)
- Configure the Monitoring and Analytics platform
- Implement and deploy probes that allow the collection of the metrics
- Facilitate the collection of performance measurements for specific metrics from the infrastructure components
- Define the report template to be provided to the experimenter